BRITISH LIBRARY

Christina Belisario
Research Coordinator
Morrison Foerster
12531 High Bluff Drive
San Diego, CA 92139-2040
USA

14 October 2025

Dear Christina

Springer Handbook of Acoustics editied by Thomas D. Rossing – Springer Science & Business Media LLC, 2007. Shelf mark: Vm07/50571.

According to our records, this item was receipted by The British Library on **23 July 2007** and would have been available for public use from that date.

A copy of the title page showing the date stamp indicating the date of availability, the copyright page and the contents pages has been attached. We have also attached the requested pages. This covers 3 pdf files.

Please note that we can only provide the date that the British Library made this item available for public use; for the actual date of publication, please contact the publisher.

Yours sincerely

Tony Corsini

British Library Research Service
Tel: +44 (0) 20 7412 7903 bipc-research@bl.uk

Springer Handbook of Acoustics



CD-ROM

Rossing
Editor

W

PLEASE CHECK FOR CD TAPE DISK ON RETURN

removed

THIS ITEM MUST BE RETURNED BY RECORDED DELIVERY

CAGE

EXPENSIVE ITEM

This item must be returned by recorded delivery

m07

VE

wrapper

Springer Handbook of Acoustics

Thomas D. Rossing (Ed.)

With CD-ROM, 962 Figures and 91 Tables

BRITISH LIBRARY
DOCUMENT SUPPLY CENTRE
23 JUL 2007
VmO7/.58571

Springer

Editor:

Thomas D. Rossing
Stanford University
Center for Computer Research in Music and Acoustics
Stanford, CA 94305, USA

Editorial Board:

Manfred R. Schroeder, University of Göttingen, Germany
William M. Hartmann, Michigan State University, USA
Neville H. Fletcher, Australian National University, Australia
Floyd Dunn, University of Illinois, USA
D. Murray Campbell, The University of Edinburgh, UK

Library of Congress Control Number: 2006927050

ISBN: 978-0-387-30446-5 e-ISBN: 0-387-30425-0
Printed on acid free paper

© 2007, Springer Science+Business Media, LLC New York
All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer Science+Business Media, LLC New York, 233 Spring Street, New York, NY 10013, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden. The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

The use of designations, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Product liability: The publisher cannot guarantee the accuracy of any information about dosage and application contained in this book. In every individual case the user must check such information by consulting the relevant literature.

Production and typesetting: LE-TeX GbR, Leipzig
Handbook Manager: Dr. W. Skolaut, Heidelberg
Typography and layout: schreiberVIS, Seeheim
Illustrations:
schreiberVIS, Seeheim & Hippmann GbR, Schwarzenbruck
Cover design: eStudio Calamar Steinen, Barcelona
Cover production: WMXDesign GmbH, Heidelberg
Printing and binding: Stürtz AG, Würzburg

SPIN 11309031 100/3100/YL 5 4 3 2 1 0

Contents

List of Abbreviations XXI

1 **Introduction to Acoustics**
Thomas D. Rossing 1
1.1 Acoustics: The Science of Sound 1
1.2 Sounds We Hear 1
1.3 Sounds We Cannot Hear: Ultrasound and Infrasound 2
1.4 Sounds We Would Rather Not Hear: Environmental Noise Control 2
1.5 Aesthetic Sound: Music 3
1.6 Sound of the Human Voice: Speech and Singing 3
1.7 How We Hear: Physiological and Psychological Acoustics 4
1.8 Architectural Acoustics 4
1.9 Harnessing Sound: Physical and Engineering Acoustics 5
1.10 Medical Acoustics 5
1.11 Sounds of the Sea 6
References 6

Part A Propagation of Sound

2 **A Brief History of Acoustics**
Thomas D. Rossing 9
2.1 Acoustics in Ancient Times 9
2.2 Early Experiments on Vibrating Strings, Membranes and Plates 10
2.3 Speed of Sound in Air 10
2.4 Speed of Sound in Liquids and Solids 11
2.5 Determining Frequency 11
2.6 Acoustics in the 19th Century 12
2.7 The 20th Century 15
2.8 Conclusion 23
References 23

3 **Basic Linear Acoustics**
Alan D. Pierce 25
3.1 Introduction 27
3.2 Equations of Continuum Mechanics 28
3.3 Equations of Linear Acoustics 35
3.4 Variational Formulations 40
3.5 Waves of Constant Frequency 45
3.6 Plane Waves 47

Amazon EX1041 Page 5

3.7 Attenuation of Sound 49
3.8 Acoustic Intensity and Power 58
3.9 Impedance 60
3.10 Reflection and Transmission 61
3.11 Spherical Waves 65
3.12 Cylindrical Waves 75
3.13 Simple Sources of Sound 82
3.14 Integral Equations in Acoustics 87
3.15 Waveguides, Ducts, and Resonators 89
3.16 Ray Acoustics 94
3.17 Diffraction 98
3.18 Parabolic Equation Methods 107
References 108

4 **Sound Propagation in the Atmosphere** 113
Keith Attenborough 113
4.1 A Short History of Outdoor Acoustics 114
4.2 Applications of Outdoor Acoustics 115
4.3 Spreading Losses 116
4.4 Atmospheric Absorption 116
4.5 Diffraction and Barriers 120
4.6 Ground Effects 129
4.7 Attenuation Through Trees and Foliage 131
4.8 Wind and Temperature Gradient Effects on Outdoor Sound 142
4.9 Concluding Remarks 143
References 143

5 **Underwater Acoustics** 149
William A. Kuperman, Philippe Roux 151
5.1 Ocean Acoustic Environment 155
5.2 Physical Mechanisms 165
5.3 SONAR and the SONAR Equation 167
5.4 Sound Propagation Models 177
5.5 Quantitative Description of Propagation 179
5.6 SONAR Array Processing 185
5.7 Active SONAR Processing 195
5.8 Acoustics and Marine Animals 201
5.A Appendix: Units 201
References 201

Part B Physical and Nonlinear Acoustics

6 **Physical Acoustics** 207
Mack A. Breazeale, Michael McPherson 209
6.1 Theoretical Overview 219
6.2 Applications of Physical Acoustics .....

Amazon EX1041 Page 6

6.3 Apparatus 226
6.4 Surface Acoustic Waves 231
6.5 Nonlinear Acoustics 234
References 237

7 **Thermoacoustics**
Gregory W. Swift 239
7.1 History 239
7.2 Shared Concepts 240
7.3 Engines 244
7.4 Dissipation 249
7.5 Refrigeration 250
7.6 Mixture Separation 253
References 254

8 **Nonlinear Acoustics in Fluids**
Werner Lauterborn, Thomas Kurz, Iskander Akhatov 257
8.1 Origin of Nonlinearity 258
8.2 Equation of State 259
8.3 The Nonlinearity Parameter *B/A* 260
8.4 The Coefficient of Nonlinearity β 262
8.5 Simple Nonlinear Waves 263
8.6 Lossless Finite-Amplitude Acoustic Waves 264
8.7 Thermoviscous Finite-Amplitude Acoustic Waves 268
8.8 Shock Waves 271
8.9 Interaction of Nonlinear Waves 273
8.10 Bubbly Liquids 275
8.11 Sonoluminescence 286
8.12 Acoustic Chaos 289
References 293

Part C Architectural Acoustics

9 **Acoustics in Halls for Speech and Music**
Anders Christian Gade 301
9.1 Room Acoustic Concepts 302
9.2 Subjective Room Acoustics 303
9.3 Subjective and Objective Room Acoustic Parameters 306
9.4 Measurement of Objective Parameters 314
9.5 Prediction of Room Acoustic Parameters 316
9.6 Geometric Design Considerations 323
9.7 Room Acoustic Design of Auditoria for Specific Purposes 334
9.8 Sound Systems for Auditoria 346
References 349

Amazon EX1041 Page 7

10 **Concert Hall Acoustics Based on Subjective Preference Theory** 351
Yoichi Ando 353
10.1 Theory of Subjective Preference for the Sound Field 361
10.2 Design Studies 370
10.3 Individual Preferences of a Listener and a Performer 377
10.4 Acoustical Measurements of the Sound Fields in Rooms 384
References 384

11 **Building Acoustics** 387
James Cowan 387
11.1 Room Acoustics 400
11.2 General Noise Reduction Methods 403
11.3 Noise Ratings for Steady Background Sound Levels 405
11.4 Noise Sources in Buildings 407
11.5 Noise Control Methods for Building Systems 419
11.6 Acoustical Privacy in Buildings 424
11.7 Relevant Standards 425
References 425

Part D Hearing and Signal Processing

12 **Physiological Acoustics** 429
Eric D. Young 429
12.1 The External and Middle Ear 434
12.2 Cochlea 449
12.3 Auditory Nerve and Central Nervous System 452
12.4 Summary 453
References 453

13 **Psychoacoustics** 459
Brian C. J. Moore 460
13.1 Absolute Thresholds 461
13.2 Frequency Selectivity and Masking 468
13.3 Loudness 473
13.4 Temporal Processing in the Auditory System 477
13.5 Pitch Perception 483
13.6 Timbre Perception 484
13.7 The Localization of Sounds 485
13.8 Auditory Scene Analysis 494
13.9 Further Reading and Supplementary Materials 495
References 495

14 **Acoustic Signal Processing** 503
William M. Hartmann 504
14.1 Definitions 505
14.2 Fourier Series

Amazon EX1041 Page 8

14.3 Fourier Transform 507
14.4 Power, Energy, and Power Spectrum 510
14.5 Statistics 511
14.6 Hilbert Transform and the Envelope 514
14.7 Filters 515
14.8 The Cepstrum 517
14.9 Noise 518
14.10 Sampled data 520
14.11 Discrete Fourier Transform 522
14.12 The z-Transform 524
14.13 Maximum Length Sequences 526
14.14 Information Theory 528
References 530

Part E Music, Speech, Electroacoustics

15 **Musical Acoustics** 533
Colin Gough
15.1 Vibrational Modes of Instruments 535
15.2 Stringed Instruments 554
15.3 Wind Instruments 601
15.4 Percussion Instruments 641
References 661

16 **The Human Voice in Speech and Singing** 669
Björn Lindblom, Johan Sundberg
16.1 Breathing 669
16.2 The Glottal Sound Source 676
16.3 The Vocal Tract Filter 682
16.4 Articulatory Processes, Vowels and Consonants 687
16.5 The Syllable 695
16.6 Rhythm and Timing 699
16.7 Prosody and Speech Dynamics 701
16.8 Control of Sound in Speech and Singing 703
16.9 The Expressive Power of the Human Voice 706
References 706

17 **Computer Music** 713
Perry R. Cook
17.1 Computer Audio Basics 714
17.2 Pulse Code Modulation Synthesis 717
17.3 Additive (Fourier, Sinusoidal) Synthesis 719
17.4 Modal (Damped Sinusoidal) Synthesis 722
17.5 Subtractive (Source-Filter) Synthesis 724
17.6 Frequency Modulation (FM) Synthesis 727
17.7 FOFs, Wavelets, and Grains 728

Amazon EX1041 Page 9

17.8 Physical Modeling (The Wave Equation) 730
17.9 Music Description and Control 735
17.10 Composition 737
17.11 Controllers and Performance Systems 737
17.12 Music Understanding and Modeling by Computer 738
17.13 Conclusions, and the Future 740
References 740

18 **Audio and Electroacoustics** 743
Mark F. Davis 744
18.1 Historical Review 747
18.2 The Psychoacoustics of Audio and Electroacoustics 751
18.3 Audio Specifications 757
18.4 Audio Components 768
18.5 Digital Audio 775
18.6 Complete Audio Systems 778
18.7 Appraisal and Speculation 778
References 778

Part F Biological and Medical Acoustics

19 **Animal Bioacoustics** 785
Neville H. Fletcher 785
19.1 Optimized Communication 787
19.2 Hearing and Sound Production 788
19.3 Vibrational Communication 788
19.4 Insects 790
19.5 Land Vertebrates 795
19.6 Birds 796
19.7 Bats 797
19.8 Aquatic Animals 799
19.9 Generalities 799
19.10 Quantitative System Analysis 802
References 802

20 **Cetacean Acoustics** 805
Whitlow W. L. Au, Marc O. Lammers 806
20.1 Hearing in Cetaceans 813
20.2 Echolocation Signals 821
20.3 Odontocete Acoustic Communication 827
20.4 Acoustic Signals of Mysticetes 830
20.5 Discussion 831
References 831

21 **Medical Acoustics** 839
Kirk W. Beach, Barbrina Dunmire 841
21.1 Introduction to Medical Acoustics 841

Amazon EX1041 Page 10

21.2 Medical Diagnosis; Physical Examination 842
21.3 Basic Physics of Ultrasound Propagation in Tissue 848
21.4 Methods of Medical Ultrasound Examination 857
21.5 Medical Contrast Agents 882
21.6 Ultrasound Hyperthermia in Physical Therapy 889
21.7 High-Intensity Focused Ultrasound (HIFU) in Surgery 890
21.8 Lithotripsy of Kidney Stones 891
21.9 Thrombolysis 892
21.10 Lower-Frequency Therapies 892
21.11 Ultrasound Safety 892
References 895

Part G Structural Acoustics and Noise

22 **Structural Acoustics and Vibrations**
Antoine Chaigne 901
22.1 Dynamics of the Linear Single-Degree-of-Freedom (1-DOF) Oscillator 903
22.2 Discrete Systems 907
22.3 Strings and Membranes 913
22.4 Bars, Plates and Shells 920
22.5 Structural–Acoustic Coupling 926
22.6 Damping 940
22.7 Nonlinear Vibrations 947
22.8 Conclusion. Advanced Topics 957
References 958

23 **Noise**
George C. Maling, Jr. 961
23.1 Instruments for Noise Measurements 965
23.2 Noise Sources 970
23.3 Propagation Paths 991
23.4 Noise and the Receiver 999
23.5 Regulations and Policy for Noise Control 1006
23.6 Other Information Resources 1010
References 1010

Part H Engineering Acoustics

24 **Microphones and Their Calibration**
George S. K. Wong 1021
24.1 Historic References on Condenser Microphones and Calibration 1024
24.2 Theory 1024
24.3 Reciprocity Pressure Calibration 1026
24.4 Corrections 1029

Amazon EX1041 Page 11

24.5 Free-Field Microphone Calibration 1039
24.6 Comparison Methods for Microphone Calibration 1039
24.7 Frequency Response Measurement with Electrostatic Actuators 1043
24.8 Overall View on Microphone Calibration 1043
24.A Acoustic Transfer Impedance Evaluation 1045
24.B Physical Properties of Air 1045
References 1048

25 **Sound Intensity** 1053
Finn Jacobsen 1054
25.1 Conservation of Sound Energy 1055
25.2 Active and Reactive Sound Fields 1058
25.3 Measurement of Sound Intensity 1068
25.4 Applications of Sound Intensity 1072
References 1077

26 **Acoustic Holography** 1077
Yang-Hann Kim 1077
26.1 The Methodology of Acoustic Source Identification 1079
26.2 Acoustic Holography: Measurement, Prediction and Analysis 1092
26.3 Summary 1092
26.A Mathematical Derivations of Three Acoustic Holography Methods and Their Discrete Forms 1095
References 1101

27 **Optical Methods for Acoustics and Vibration Measurements** 1101
Nils-Erik Molin 1101
27.1 Introduction 1105
27.2 Measurement Principles and Some Applications 1122
27.3 Summary 1123
References 1127

28 **Modal Analysis** 1127
Thomas D. Rossing 1127
28.1 Modes of Vibration 1128
28.2 Experimental Modal Testing 1133
28.3 Mathematical Modal Analysis 1136
28.4 Sound-Field Analysis 1137
28.5 Holographic Modal Analysis 1138
References 1139

Acknowledgements 1141
About the Authors 1147
Detailed Contents 1167
Subject Index 1167

Amazon EX1041 Page 12

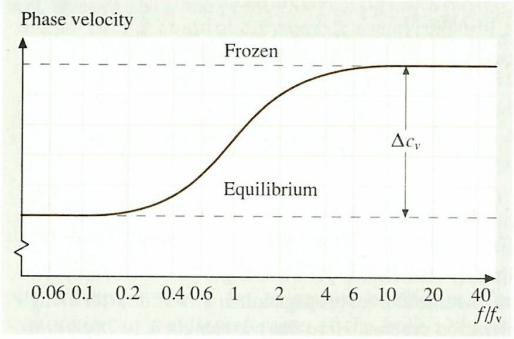


Fig. 3.12 Change in phase velocity as a function of frequency for propagation in a medium with a single relaxation process. The asymptote at zero frequency is the equilibrium sound speed c, and that in the limit of infinite frequency is the frozen sound speed $c+\Delta c_\nu$

when the frequency is increased from a value small compared to the relaxation frequency to one large compared to the relaxation frequency. Here the term relaxation frequency is used to denote the reciprocal of the product of 2π with the relaxation time.

3.7.3 Continuously Distributed Relaxations

It is sometimes argued that, for heterogeneous media such as biological tissue and rocks, the attenuation is due to a statistical distribution of relaxation times, so that the sum in (3.224) should be replaced by an integral, yielding

$$k=\frac{\omega}{c}\left(1+\mathrm{i}\frac{\omega\delta_{\mathrm{cl}}}{c^2}+\frac{1}{c}\int_0^\infty\frac{\mathrm{d}(\Delta c)}{\mathrm{d}\tau}\frac{\mathrm{i}\omega\tau}{1-\mathrm{i}\omega\tau}\,\mathrm{d}\tau\right), \tag{3.229}$$

and so that the attenuation constant due to the relaxation processes becomes

$$\sum_\nu\alpha_\nu\to\frac{1}{c^2}\int_0^\infty\frac{\mathrm{d}(\Delta c)}{\mathrm{d}\tau}\frac{\omega^2\tau}{1+(\omega\tau)^2}\,\mathrm{d}\tau\,. \tag{3.230}$$

Here, the quantity

$$\frac{\mathrm{d}(\Delta c)}{\mathrm{d}\tau}\Delta\tau \tag{3.231}$$

is interpreted as the additional increment in phase velocity at high frequencies due to all the relaxation processes whose relaxation times are in the interval $\Delta\tau$.

Such a continuous distribution of relaxation processes is intrinsically capable of explaining a variety of experimentally observed frequency dependencies of $\alpha(\omega)$. A simple example is that where

$$\frac{\mathrm{d}(\Delta c)}{\mathrm{d}\tau}=\frac{K}{\tau^q}\,, \tag{3.232}$$

with K being a constant independent of τ, and with q being a specified exponent, with $0<q<2$. In such a case the integral in (3.230) becomes

$$\frac{1}{c^2}\int_0^\infty\frac{\mathrm{d}(\Delta c)}{\mathrm{d}\tau}\frac{\omega^2\tau}{1+(\omega\tau)^2}\,\mathrm{d}\tau=\frac{K}{c^2}\omega^q\int_0^\infty\frac{u^{1-q}}{1+u^2}\,\mathrm{d}u \tag{3.233}$$

giving a power-law dependence on ω that varies as ω^q. Although the hypothesis of (3.232) is probably unrealistic over all ranges of relaxation time τ, its being nearly satisfied over an extensive range of such times could yield predictions that would explain a power-law dependence. (The integral that appears here is just a numerical constant, which is finite provided $0<q<2$. The integral emerges when one changes the integration variable from τ to $u=\omega\tau$.)

The corresponding increment in the reciprocal $1/v_{\mathrm{ph}}=k_{\mathrm{R}}/\omega$ of the phase velocity v_{ph} is

$$\Delta\left(\frac{1}{v_{\mathrm{ph}}}\right)=-\frac{1}{c^2}\int_0^\infty\frac{\mathrm{d}(\Delta c)}{\mathrm{d}\tau}\frac{\omega^2\tau^2}{1+(\omega\tau)^2}\,\mathrm{d}\tau\,. \tag{3.234}$$

Insertion of (3.232) into this yields

$$\Delta\left(\frac{1}{v_{\mathrm{ph}}}\right)=-\frac{K}{c^2}\omega^{q-1}\int_0^\infty\frac{u^{2-q}}{1+u^2}\,\mathrm{d}u\,, \tag{3.235}$$

which varies with ω as ω^{q-1}. In this case, however, the integral exists only if $1<q<3$, so the overall analysis has credibility only if $1<q<2$.

3.7.4 Kramers–Krönig Relations

In attempts to explain or predict the frequency dependence of attenuation and phase velocity in materials, some help is found from the principle of causality [3.56]. Suppose one has a plane wave traveling in the $+x$-direction and that, at $x=0$, the acoustic pressure is a transient given by

$$p(0,t)=\int_{-\infty}^\infty\hat{p}(\omega)\mathrm{e}^{-\mathrm{i}\omega t}\,\mathrm{d}\omega=2\mathrm{Re}\int_0^\infty\hat{p}(\omega)\mathrm{e}^{-\mathrm{i}\omega t}\,\mathrm{d}\omega. \tag{3.236}$$

Amazon EX1041 Page 13

Then the transient at a distant positive value of x is given by

$$p(x,t) = 2\mathrm{Re}\int_0^\infty \hat{p}(\omega)\,\mathrm{e}^{-\mathrm{i}\omega t}\,\mathrm{e}^{\mathrm{i}kx}\,\mathrm{d}\omega\,, \tag{3.237}$$

where $k = k(\omega)$ is the complex wavenumber, the imaginary part of which is the attenuation coefficient.

The causality argument is now made that, if $p(0,t)$ vanishes at all times before some time t_0 in the past, then so should $p(x,t)$. One defines $k(\omega)$ for negative values of frequency so that

$$k(-\omega) = -k^*(\omega)\,. \tag{3.238}$$

This requires that the real and imaginary parts of the complex wavenumber be such that

$$k_\mathrm{R}(-\omega) = -k_\mathrm{R}(\omega)\,; \qquad k_\mathrm{I}(-\omega) = k_\mathrm{I}(\omega)\,, \tag{3.239}$$

so that the real part is odd, and the imaginary part is even in ω. This extension to negative frequencies allows one to write (3.237) in the form

$$p(x,t) = \int_{-\infty}^\infty \hat{p}(\omega)\,\mathrm{e}^{-\mathrm{i}\omega t}\,\mathrm{e}^{\mathrm{i}kx}\,\mathrm{d}\omega\,. \tag{3.240}$$

The theory of complex variables assures one that this integral will indeed vanish for sufficiently early times if: (i) one can regard the integrand as being defined for complex values of ω, and (ii) one can regard the integrand as having certain general properties in the upper half of the complex ω plane. In particular, it must be analytic (no poles, no branch cuts, no points where a power series does not exist) in this upper half of the comp;ex plane. Another condition is that, when t has a sufficiently negative value, the integrand goes to zero as $|\omega| \to \infty$ in the upper half-plane. The possibility of the latter can be tested with the question of whether

$$\mathrm{Re}\,[-\mathrm{i}\omega t + \mathrm{i}kx)] \to -\infty \qquad \text{as} \quad \omega \to \mathrm{i}\infty\,, \tag{3.241}$$

with $t < 0$ and $x > 0$. In any event, for all these conditions to be met, the complex wavenumber $k(\omega)$, considered as a complex function of the complex variable ω, has to be analytic in the upper half-plane. One cannot say at the outset, without having a well-defined causal mathematical model, just how it behaves at infinity, but various experimental data suggests it approaches a polynomial with a leading exponent of not more than two. Also, analysis on specific cases suggests it is always such that k/ω is analytic and, moreover, finite at $\omega = 0$.

Contour Integral Identity

Given these general properties of the complex wavenumber, one considers the contour integral

$$I_\mathrm{A}(\zeta_1,\zeta_2,\zeta_3,\zeta_4) = \oint \frac{k/\omega}{(\omega-\zeta_1)(\omega-\zeta_2)(\omega-\zeta_3)(\omega-\zeta_4)}\,\mathrm{d}\omega\,, \tag{3.242}$$

where the contour is closed, the integration proceeds in the counterclockwise sense, the contour is entirely in the upper half plane, and where the contour encloses the four poles at ζ_1, ζ_2, etc., all of which are in the upper half-plane.

(The number of poles that one includes is somewhat arbitrary, and four is a judicious compromise for the present chapter. The original [3.57,58] analysis directed toward acoustical applications included only two, but this led to divergent integrals when the result was applied to certain experimental laws that had been extrapolated to high frequencies. The more poles one includes, the less dependent is the prediction on details of extrapolations to frequencies outside the range of experimental data.)

The integral can be deformed to one which proceeds along the real axis from $-\infty$ to ∞ and which is then completed by a semicircle of infinite radius that encloses the upper half-plane. Because of the supposed behavior of k at ∞ in the upper half-plane, the latter integral is zero and one has

$$I_\mathrm{A}(\zeta_1,\zeta_2,\zeta_3,\zeta_4) = \int_{-\infty}^\infty \frac{k/\omega}{(\omega-\zeta_1)(\omega-\zeta_2)(\omega-\zeta_3)(\omega-\zeta_4)}\,\mathrm{d}\omega\,. \tag{3.243}$$

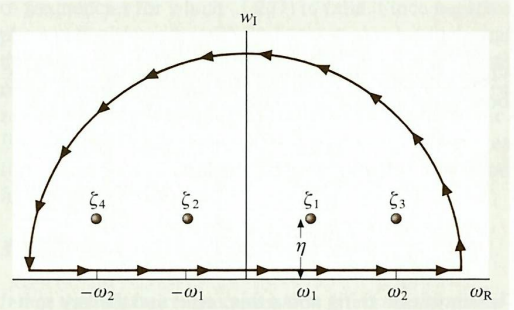


Fig. 3.13 Contour integral in the complex frequency plane used in the derivation of one version of the Kramers–Krönig relations

Amazon EX1041 Page 14

selected). A distance different from 10 m can be used, if a correction for the distance attenuation is applied as well.

The expected value of G according to diffuse field theory becomes a function of T as well as of the room volume, V:

$$G_{\mathrm{exp}} = 10\log_{10}\left(\frac{T}{V}\right) + 45\,\mathrm{dB}\,. \qquad (9.12)$$

The subjective difference limen for G is about 1.0 dB. The definition of G is illustrated in Fig. 9.7.

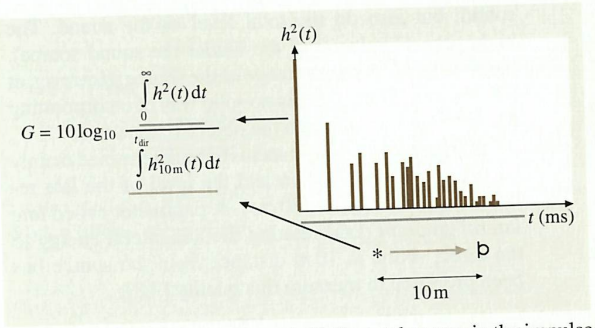


Fig. 9.7 The definition of strength G: the total energy in the impulse response measured relative to the direct sound level at 10 m distance from the source

9.3.4 Measures of Spaciousness

Spaciousness is the feeling that the sound is arriving from many different directions in contrast to a monophonic impression of all sound reaching the listener through a narrow opening. It is now clear that there are two aspects of *spaciousness*, both of which are attractive, particularly when listening to music:

- *apparent source width* (ASW): the impression that the sound image is wider than the visual, physical extent of the source(s) on stage. This should not be confused with localization errors, which of course should be avoided.

and

- *listener envelopment* (LEV): the impression of being inside and surrounded by the reverberant sound field in the room.

Both aspects have been found to be dependent on the direction of incidence of the impulse response reflections. When a larger portion of the early reflection energy (up to about 80 ms) arrives from lateral directions (from the sides) the ASW increases. When the level of the late, lateral reflections is high, strong LEV results.

The lateral components of the impulse response energy can be recorded using a figure-of-eight microphone with the sensitive axis held horizontal and perpendicular to the direction towards the sound source (so that the source lies in the deaf plane of the microphone). For measurement of the lateral energy fraction (LEF), the early part (up to 80 ms) of this lateral sound energy is compared with the energy of the direct sound plus all early reflections picked up by an ordinary omnidirectional microphone:

$$\mathrm{LEF} = \int_{t=5\,\mathrm{ms}}^{t=80\,\mathrm{ms}} h_1^2(t)\,\mathrm{d}t \Big/ \int_{t=0\,\mathrm{ms}}^{t=80\,\mathrm{ms}} h^2(t)\,\mathrm{d}t\,, \qquad (9.13)$$

where h_1 is the impulse response pressure recorded with a figure-of-eight microphone, whereas h is captured through the (usual) omnidirectional microphone.

It is mainly the energy at low and mid frequencies that contribute to the *spaciousness*. Consequently, LEF is normally averaged over the four octaves 125–1000 Hz. The higher the value of LEF, the wider the ASW. The literature contains many data on LEF values in different, existing concert halls.

LEF does not have an expected value related to T. In a completely diffuse field, LEF would be constant with a value of 0.33, which is higher than that normally found in real halls. The subjective difference limen for LEF is about 5%.

The definition of LEF is illustrated in Fig. 9.8.

The ASW aspect of *spaciousness* is not only dependent on the ratio between early lateral and total early

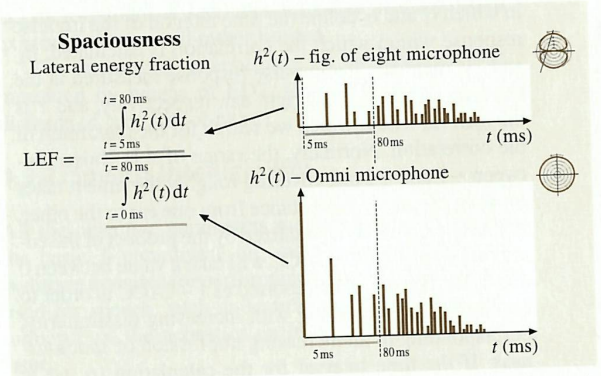


Fig. 9.8 The definition of lateral energy fraction LEF: the ratio between early reflection energy arriving from the sides and total early energy

Amazon EX1041 Page 15

sound; but also on the total level of the sound. The higher the G value (and the louder the sound source), the broader the acoustic image of the source. However, at the time of writing, there is no solid way of incorporating the influence of level into the objective measure.

Listener envelopment seems to be determined mainly by the spatial distribution and the level of the late reflections (arriving after 80 ms). A parameter called *late lateral strength* (LG) relating the late lateral energy to the direct sound at 10 m distance from the source has been proposed to measure this quality [9.8]:

$$\mathrm{LG} = \int_{t=80\,\mathrm{ms}}^{t=\infty} h_{\mathrm{l}}^2(t)\,\mathrm{d}t \Big/ \int_{t=0\,\mathrm{ms}}^{t=t_{\mathrm{dir}}} h_{10\,\mathrm{m}}^2(t)\,\mathrm{d}t\,. \qquad (9.14)$$

LG is likely to be included in the ISO 3382 standard (from 2007); but so far LG has only been found correlating with subjective responses in laboratory experiments using synthetic sound fields. Therefore, it may be wise not to put too much emphasis on this parameter until its value has also been confirmed by experiments in real halls.

In contrast to reflections arriving in the median plane, lateral reflections will cause the instantaneous pressure at the two ears of the listener to be different. Such a dissimilarity of the signals at the two ears can also be measured by means of the interaural cross-correlation coefficient IACC:

$$\mathrm{IACC}_{t_1,t_2} = \max \left| \frac{\int_{t_1}^{t_2} h_{\mathrm{L}}(t) h_{\mathrm{R}}(t+\tau)\,\mathrm{d}t}{\sqrt{\int_{t_1}^{t_2} h_{\mathrm{L}}^2(t)\,\mathrm{d}t \int_{t_1}^{t_2} h_{\mathrm{R}}^2(t)\,\mathrm{d}t}} \right| \qquad (9.15)$$

in which t_1 and t_2 define the time interval of the impulse response within which the correlation is calculated, h_{L} and h_{R} represent the impulse response measured at the entrance of the left and right ear, respectively, and τ is the interval within which we search for the maximum of the correlation. Normally, the range of τ is chosen between -1 and $+1$ ms, covering roughly the time it takes for sound to travel the distance from one ear to the other. As the correlation is normalized by the product of the energy of h_{L} and h_{R}, the IACC will take a value between 0 and 1. Results are often reported as $1 - \mathrm{IACC}$ in order to obtain a value increasing with increasing dissimilarity, corresponding to an increasing impression of *spaciousness*. If the time interval for the calculation (t_1, t_2) is (0 ms, 100 ms), then the IACC will measure the ASW, while a later interval $(t_1, t_2) = (100\,\mathrm{ms}, 1000\,\mathrm{ms})$ may be used to describe the LEV. The literature on measured values in halls mainly contain data on IACC related to the 0–100 ms time interval.

Although the LEF and IACC parameters relate to the same subjective quality, they are not highly correlated in practice. Another puzzling fact is that LEF and IACC emphasize different frequency regions being of importance. LEF is primarily measured in the four lowest octaves, 125 Hz, 250 Hz, 500 Hz and 1000 Hz while IACC should rather be measured in the octave bands above 500 Hz. IACC values would always be high in the lower octaves, because the distance between the ears (< 30 cm) is small compared to 1/4 of the wave length (≈ 70 cm at 125 Hz).

9.3.5 Parameters Relating to Timbre or Tonal Color

Timbre describes the degree to which the room influences the frequency balance between high, middle and low frequencies, i. e. whether the sound is *harsh*, *bright*, *hollow*, *warm*, or whatever other adjective one would use to describe *tonal color*. Traditionally, a graph of the frequency variation of T (per 1/1 or 1/3 octave) has been used to indicate this quality; but a convenient single-number parameter intended to measure the warmth of the hall has been suggested [9.9]: the bass ratio (BR) given by:

$$\mathrm{BR} = \frac{T_{125\,\mathrm{Hz}} + T_{250\,\mathrm{Hz}}}{T_{500\,\mathrm{Hz}} + T_{1000\,\mathrm{Hz}}}\,. \qquad (9.16)$$

Likewise, a *treble ratio* (TR) can be formed as:

$$\mathrm{TR} = \frac{T_{2000\,\mathrm{Hz}} + T_{4000\,\mathrm{Hz}}}{T_{500\,\mathrm{Hz}} + T_{1000\,\mathrm{Hz}}}\,. \qquad (9.17)$$

However, in some halls, a lack of bass sound is experienced in spite of high T values at low frequencies. Therefore EDT or perhaps G versus frequency would be a better – and intuitively more logical – parameter for measurement of *timbre*. Likewise, BR or TR could be based on G rather than T values.

Besides the subjective parameters mentioned above, a quality called *intimacy* is also regarded as being important when listening in auditoria. *Beranek* [9.9] originally suggested this quality to be related to the initial time delay gap ITDG; but this has not been experimentally verified (At the time of writing, Beranek no longer advocates this idea but mentions the observation that, if the ITDG exceeds 45 ms, the hall has no intimacy [9.10]). It is likely that *intimacy* is related to a combination of some of the other parameters already mentioned, such as a high sound level combined with a clear and envelop-

Amazon EX1041 Page 16

ated by the architects. However, one needs to break up eventual curved surfaces into plane elements and often it is also advisable to simplify details in the geometry provided by the architect. Alternatively, the acoustician can build the room model by writing a file of coordinate points or mathematical expressions for all surface corners. Of course this will be more time consuming in cases of complex room geometry, but some programs offer intelligent programming routines for the modeling of geometric elements.

When the geometry has been completed, the absorption values for each octave band, the scatter, and the eventual degree of acoustic transparency have to be assigned to each surface. In addition, the source and receiver characteristics (power and directivity versus frequency, position and orientation) must also be entered.

In the calculation of sound propagation, most models disregard phase properties and use an energy approximation. The sound propagation is studied by means of rays (up to millions depending on the room complexity) being traced from the source position through reflections until their energy has been reduced below the dynamic range of interest (often determined by T or the signal-to-noise ratio in auralizations). Reflection directions are chosen so that angle of reflection equals angle of incidence or with more or less random angles depending on the scatter value attributed to the surface. For the calculation of early reflections, some models also apply image source theory, which actually makes phase considerations possible if the complex impedances of the various surface materials and not just their energy absorption coefficients are known. However, the late reverberation part is normally calculated by some kind of ray method, as image source calculations become impractical after a few orders of reflection. Recently, however, attempts to carry out complete calculations inspired by finite-element (FEM) or boundary-element (BEM) methods have been reported even for rather complex rooms [9.23], whereby phase and diffraction phenomena can be handled.

As a result of tracing the history of all rays emitted in the room model, source and receiver specific impulse responses are produced with detailed information about the direction of incidence, delay and level versus frequency of each reflection component. From these, all the parameters mentioned in Sect. 9.3 can be calculated. Most programs can generate a grid of receiver point over selected audience areas and code the numeric results using a color scale to facilitate a fast overview of the detailed position results for any of the parameters. Also, automatic plots of relevant statistics regarding the distribution with frequency and position are provided.

Impulse responses suitable for use in *auralization* can be generated by applying available head-related transfer functions, which have been recorded as a function of direction of sound incidence on an artificial head. Such transfer functions not only represent the directivity of the human receiver, but are also important for correct perception of direction of incidence of sound field components when listening through headphones. Alternatively, the directional information of the individual reflections making up the impulse response can be coded into a surround-sound format for listening via a multiple-loudspeaker setup.

Figure 9.14 shows a view of a room model as well as two binaural impulse responses and a color mapping of a calculated parameter (generated by the ODEON software).

9.5.5 Scale Model Predictions

Acoustic scale models have been used since the 1930s [9.24] to study the behavior of sound in rooms. With proper care in detailing the geometry and the choice of materials for building the model, this technique can provide quite accurate predictions of impulse responses and acoustic parameters. Actually, scale modeling is still regarded as more reliable than computer models in cases where the room contains a lot of irregular surfaces or objects that diffuse or diffract sound waves. However, the use of physical scale models is limited to larger, acoustically demanding projects, as building and testing the model is far more time consuming and costly than computer modeling.

The diffusion and diffraction caused by an object or surface in a room depends on the ratio between the linear dimensions and the wavelength of the sound: l/λ. Therefore, in a $1:M$ scale model, the diffraction/diffusion conditions will be modeled correctly if the wavelengths of the measurement signals in the model are chosen such that:

$$\frac{l}{\lambda} = \frac{l_\mathrm{M}}{\lambda_\mathrm{M}} = \frac{l/M}{\lambda/M} \quad \Rightarrow \quad f_\mathrm{M} = fM \,. \tag{9.28}$$

This means that we should increase the frequency used in the model measurements by the factor M. Fortunately, sound sources and microphones exist that are capable of handling frequencies up to around 100 kHz, whereby the frequency range up to 10 kHz can be handled in a $1:10$ scale model or 2 kHz in a $1:50$ scale model. The sources may be small loudspeaker units (piezoelectric or

Amazon EX1041 Page 17

electro magnetic) and the microphones 1/4" or even 1/8" condenser types.

Ideally, the absorption coefficient α_M of the materials chosen for building each surface in the hall model should fulfil the equation:

$$\alpha_M(fM) = \alpha(f) \,. \tag{9.29}$$

In practice, however, we need only to distinguish between mainly reflecting and mainly absorptive materials. The reflecting surfaces are quite easy, as these can be made of any hard material (plywood, plaster) with a couple of layers of varnish if necessary. It is far more important to build the seats/audience so that both head diffraction and absorption is correct for the scale chosen. In the end it is often necessary to fine-tune the reverberation time in the model by applying absorption to some secondary surfaces (surfaces that do not generate primary early reflections).

Another important issue is to compensate for the excessive air absorption at high frequencies in the model. This can be achieved either by drying the air in the model, by exchanging the air for nitrogen, or simply as a calculated compensation since the attenuation is a simple linear function of time once the, frequency, humidity and temperature have been specified. However, in the case of compensation the signal-to-noise ratio at high frequencies will be severely limited, which reduces the decay range for the estimation of T and the sound quality if auralization is attempted. Figures 9.15 and 9.16 show an example of a 1:10 scale model and its audience respectively.

9.5.6 Prediction from Empirical Data

In the 1980s, when most of the current objective acoustic parameters had been defined, several researchers started making systematic measurement surveys of concert, opera and multipurpose halls. Many of these measured data were published in [9.3]. Combining acoustic data with data on the geometry of the halls has made it possible to derive some simple rules of thumb regarding how, and how much, acoustic parameters are affected by the choice of gross room shape and dimensions.

Through statistical multiple linear-regression analysis of data from more than 50 halls, a number of relationships between acoustical and geometrical properties have been found [9.25]. A number of the linear regression models derived are shown in Table 9.4. In these models, the variations in the acoustic parameters are described as functions of the expected value according to diffuse field theory (G_{exp} and C_{exp}, Sect. 9.3) plus various geometrical variables such as: average room height H, average width W, room volume V, number of seats N, number of rear balcony levels 'no. rear balc.' etc. When the models were derived, the measured reverberation time (T) was used to calculate the expected values of strength G_{exp} and clarity C_{exp}. When using these models for prediction of the values in a hall not yet built, the expected values could be based on a Sabine calculation of T instead. It should be mentioned that, apart from $\Delta G(10\,\text{m})$, the attenuation in strength per 10 m increase in source–receiver distance, the predicted values should be considered as representing the position average of values to be found in the hall.

The three rightmost columns in Table 9.4 contain information about how well each prediction formula matched the measured data. When comparing the amounts of variance explained by the different models, it is seen that the diffuse field prediction is responsible for the largest portion of the variance. This may be interpreted as hall volume and total absorption area being the most important factors governing the behavior

Table 9.4 Regression models for the relationships between room acoustic parameters and room design variables as derived from a database containing data from more than 50 halls (after [9.25])

Room acoustic parameter	Regression models: $f(\text{PAR}_{\text{expected}}, \text{geometry})$	Correlation coefficient	% of variance	STD residuals
C (dB)	$-0.1+1.0C_{exp}$	0.76	58	1.0 dB
	$-1.4+0.95C_{exp}+0.47W/H+0.031$ floor slope	0.83	68	0.9 dB
	$-1.2+1.03C_{exp}+0.43W/H+0.013$ wall angle	0.84	70	0.9 dB
	$-1.77+1.10C_{exp}+0.055W+0.027$ stage ceil. angle	0.86	74	0.8 dB
G (dB)	$-2.0+0.94G_{exp}$	0.91	83	0.9 dB
	$-5.61+1.06G_{exp}+0.17V/N+0.04$ distance	0.94	89	0.9 dB
$\Delta G(10\,\text{m})$ (dB)	$-1.85+0.42$ no. rear balc.	0.50	25	0.7 dB
	$-1.41+0.35$ no. rear balc. -3.93 distance$/(HW)$	0.55	31	0.6 dB
LEF (–)	$0.39-0.0061$ width	0.70	49	0.05
	$0.37-0.0051$ width -0.00069 wall angle	0.72	53	0.05

Amazon EX1041 Page 18

wonder at the effect it might have had on the course of history, not to mention the effect it may yet have [18.2].

Sound is hardly the easiest quantity to work with. As a vibration in air, it is invisible, dynamic, and three-dimensional (3-D), not counting the dimensions of time and frequency. Human interaction with a sound field is often dynamic: natural movements of the head and body alter the perceived sounds in a way that is difficult to predict or reproduce. Despite over a century of progress, the transducers available for recording and reproducing sound, microphones and loudspeakers, are still limited to being point-in-space devices, and are not necessarily well suited to interacting with a 3-D physical phenomenon. And yet, human beings make do with just two ears; so how hard can it be to capture and reproduce all the characteristics and nuances of what the ear perceives in nature?

The intent of this chapter is to review and evaluate the performance of audio systems based on the known performance of the human auditory system. Starting with a brief historical overview, the salient aspects of human auditory performance will be described. These in turn will be used to derive a set of general audio specifications that may be expected to apply to a broad range of audio systems. Then, some of the more common audio components will be examined in light of these specifications. Finally, the performance of complete systems will be considered, in the hope of gaining some overall perspective of progress to date and possibly fertile directions for further investigation.

18.1 Historical Review

The first golden age of audio research arguably started in the late 1800s with the inventions of the telephone by A. G. Bell and the phonograph by T. A. Edison [18.3], and ended in 1982 with the introduction of digital audio as a practical reality as embodied in the audio compact disc (CD). The inventions that inaugurated this age established that an audio channel could indeed be transmitted and recorded, as a continuous mechanical, magnetic, optical, or electrical analog of the sound vibrations, but brought with them a raft of problems, including inadequate recording sensitivity and playback volume, noise, distortion, speed irregularities such as wow and flutter, nonlinear and limited frequency response, and media wear. Although steady incremental progress was made on these problems in the decades that followed, they were all largely resolved once and for all with digital audio. Efforts since then have focused on reduction of the required data rate (audio coding) and on extending the realm of audio from single isolated channels to full three-dimensional systems. Ongoing parallel efforts to understand the physiology and psychoacoustics of the human ear more fully have continued to extend the knowledge base, and in some cases have yet to be exploited by commercial audio systems.

Why wasn't sound recording invented sooner? It is hard to say. Bell's telephone may have required a steady electric current, which presumably was not available in ancient times, but Edison's phonograph was entirely mechanical. The key to the latter invention seems to have been the recognition that sound is a vibration and that the vibration is mechanically transferable, but this notion that is readily demonstrable by putting your fingers to your lips or throat, and humming. The inventions seem to have been part of a general flourishing of physical sciences and mathematics, one that has continued to the present day. Prior to that, however, there is certainly evidence of some understanding of the acoustic processes involved, as seen in the creation of musical instruments and the design of concert halls.

Actually, both Edison and Bell were driven in part by a desire to improve telegraph systems. The telegraph was popularized by Samuel Morse, starting around 1837, although the notion of an electrical telegraph harks back at least to a demonstration of sending a simple signal over a long wire by William Watson in 1747 [18.4]. The cause was greatly helped by the invention of the electric battery by Alessandro Volta in 1799, and the storage battery in 1859 by Gaston Plante [18.5].

Many of the theoretical and mathematical foundations of early audio technologies were established in the 1800s, with work in the electrical sciences by Ohm, Faraday, Henry and others, together with research in acoustics and hearing physiology by Helmholtz, Lissajous and others. Although Edison is credited with the invention of the first device to reproduce sound, the first audio recording device is generally credited to Leon Scott, who on March 25, 1857, patented the phonautograph [18.6]. This device combined a diaphragm and a stylus to inscribe a trace of a sound on a visual medium, such as a smoke-blackened glass or cylinder, but it had no way to reproduce the recording audibly. John Tyndall seems to have followed Scott's work on visualizing sound vibrations, using a variety of different devices.

Amazon EX1041 Page 19

Alexander Graham Bell filed his patent for the telephone on February 14, 1876, just two hours before Elisha Gray filed for a similar device. Aside from its intrinsic value, the invention was significant for employing acoustical to electrical and electrical to acoustical transduction, principles that are still almost universally employed in telephony, broadcast, and audio recording.

Transduction between acoustical and electrical energy was initially less important to Edison's invention of the phonograph on December 8, 1877, since a physical medium to store the recording was necessary in any case, and there neither was nor is a ready way to directly store a dynamic electrical signal directly on such a physical medium. So it is perhaps not too surprising that the original invention employed purely mechanical transduction for both recording and playback.

With these two inventions, the foundations for analog audio were established, and the race was on to refine and commercialize the systems. In 1887, Emile Berliner patented a recording format that used lateral groove modulation on a flat disk [18.8], in contrast to Edison's use of vertical modulation on a cylinder. The underlying principles may have been much the same, but Berliner's arrangement qualified as a separate patent, and the practical aspects of mass-producing, playing, and storing disks instead of cylinders led to the grooved disk becoming the commercial recording medium of choice for over half a century.

Lacking the means to record an electrical signal directly, but seeking to avoid the wear accompanying the playback of grooved recordings with a stylus, other media were soon being explored, such as the use of magnetic recording by Valdemar Poulson in 1898 and optical recording, by Leon Gaumont in 1901. Meanwhile, the eventual foundation of the broadcast industry was begun in 1894, with the invention of the wireless telegraph by Guglielmo Marconi.

These systems were initially handicapped by a lack of viable amplification means, and although there were some imaginative attempts at mechanical amplification, the key breakthrough occurred in 1906, with Lee de Forest's invention of the *audion* vacuum-tube triode amplifier. Vacuum tubes could and would be used to amplify all manner of electrical signals, but de Forest's choice of the name *audion* seems derived directly from the word audio. It is therefore a matter of some speculation as to whether the tube amplifier would have come about when it did had it not been for the formative invention of sound recording just 29 years earlier.

Aside from resolving the problems of providing better recording sensitivity and adequate playback volume, the vacuum tube also enabled the wireless transmission of audio, giving rise to the radio receiver and broadcast industries.

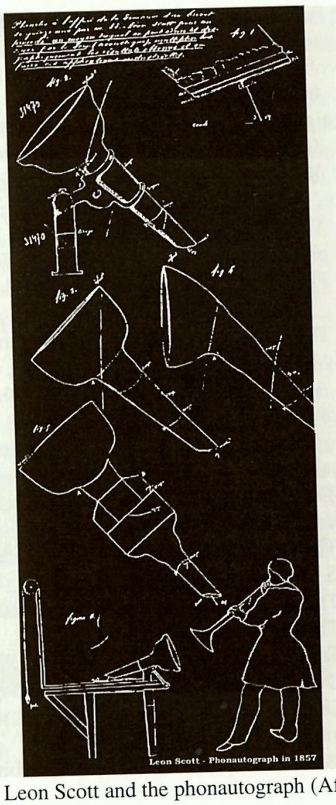


Fig. 18.1 Leon Scott and the phonautograph (After [18.7])

So by 1910, the basic principles, physics, and devices of audio and electroacoustics were in place, resulting in widespread continuing efforts to develop, patent, and refine components and systems, establish standards where required, and commercialize a growing variety of formats and services. Amidst competitive struggles and numerous patent fights, the following years saw a dramatic series of audio innovations, including [18.9, 10]:

- 1913: An early *talking* movie by Edison, using a cylinder synchronized with the picture.
- 1916: A superheterodyne radio circuit patented by Edwin F. Armstrong.
- 1917: A condenser microphone developed by E. C. Wente at Bell Laboratories.
- 1921: The first commercial amplitude-modulated (AM) broadcast, KDKA, Pittsburgh, PA.

Amazon EX1041 Page 20

- 1925: An electrical record cutter developed at Bell Laboratories, followed by the commercial release of 78 revolutions per minute (RPM) records.
- 1926: Iron-oxide-coated paper tape for magnetic recoding developed by O'Neill.
- 1927: *The Jazz Singer*, the first commercial sync-sound film, using audio on disk synchronized to the picture.
- 1929: The sampling theorem, the basis of digital signal processing, developed by Harry Nyquist.
- 1935: Plastic-based magnetic recording tape, from BASF.
- 1936: The first tape recording of a symphony concert, by Sir Thomas Beecham.
- 1936: A cardioid condenser microphone, by Von Braunmuhl and Weber.
- 1941: Commercial frequency-modulated (FM) broadcasting inaugurated in the US.
- 1947: The Williamson high-fidelity power amplifier developed.
- 1948: The 33 1/3-RPM long-play record, developed by CBS laboratories.
- 1951: The transistor developed at Bell Laboratories.
- 1954: The transistor radio, Sony.
- 1963: Tape cassette format developed by Philips.
- 1965: Dolby *A* noise-reduction system developed.
- 1969: Dr. Thomas Stockham experiments with digital recording.
- 1976: Stockham produces the first 16-bit digital recording, of the Santa Fe Opera.
- 1982: The digital compact disc is released commercially.

Much of the thrust of these and countless other developments in this period was to improve the quality of a given audio channel, by extending the frequency response, reducing noise, etc. As noted earlier, once digital recording came of age, these concerns were rendered somewhat moot, as adequate quality became a matter of using a sufficient number of bits per sample and adequately sampling rate to avoid audible degradation. Of course, certain parts of the audio chain, notably transducers such as microphones and loudspeakers, and associated interface electronics, have remained analog, as is likely to continue to be true at least until a direct link is established to the human nervous system, bypassing the ears altogether.

18.1.1 Spatial Audio History

Perhaps the one area of audio that was not immediately improved with the adoption of digital recording techniques was spatial reproduction. It is unlikely that the importance of the spatial aspect of audio was lost on early audio developers, nor the fact that most humans do in fact, possess two ears. It was probably just a case of priorities. Spatial audio systems would almost certainly require more channels and/or other information, which was hard to accommodate with early audio media, and the emphasis was, quite reasonably, on refining the quality of the single-channel monophonic format.

Still, there were some promising early experiments, notably a demonstration in 1881 by Clement Ader, who set up a series of microphones in front of the stage of the Paris Opera, and ran their outputs to earphones set up in a nearby room [18.9, 10]. Whether by accident or design, some of the listeners elected to use two headphones, one on each ear, and were rewarded with a crude but effective binaural spatial presentation.

Despite isolated experiments such as Ader's, there was little organized investigation of spatial audio until 1931, when Alan Blumlein single-handedly developed a range of devices and protocols to support stereo sound, including techniques to upgrade the entire audio broadcast and recording infrastructure of the time. Blumlein filed a far-ranging patent covering the bulk of his work; unfortunately, it seems to have been about a quarter century ahead of his time, as it would be the late 1950s before stereophonic sound became a commercial reality.

In the 1930s, Bell Laboratories also became interested in spatial audio, albeit with more of an emphasis on audience presentation. They concluded that stereo was ineffective for such uses, as the virtual imaging of just two channels would not work for listeners removed from the centerline. They further concluded that adding a center channel produced a far more satisfactory result, as demonstrated by Fletcher, Steinberg, and Snow in 1933.

With consumer media still limited to single-channel operation, it fell for a time to film studios to engage in occasional commercial forays into spatial audio. A notable experiment in the venue was the film *Fantasia* (1940), the road show productions of which used a synchronized three-channel soundtrack, plus control track, to feed as many as 10 loudspeakers [18.9, 10].

The use of multiple film channels accelerated somewhat in the 1950s, as film studios sought ways to differentiate their product from television. One popular approach was to use wider-than-standard 70 mm film stock, instead of the normal 35 mm, which provided enough room along the edge for several audio channels

Amazon EX1041 Page 21

18.3.8 Dynamic Range

Dynamic range is the amplitude range from the largest to the smallest signal an audio device or channel can handle. In cases where there is no minimum signal, other than zero level, the minimum useable level may be taken as the noise level of the channel.

If the noise level of the channel is substantially independent of signal level, the dynamic range of the channel may alternately be referred to as the signal-to-noise ratio (SNR). If the noise is signal dependent, the audibility of the noise will depend on the instantaneous signal-to-noise ratio, which is likely to be quite different from the dynamic range.

Since both the noise level and the overload point of a channel are likely to be functions of frequency, so too will the dynamic range. A common characteristic of the noise of traditional analog recording systems has been rising noise levels at the ends of the audio band, coupled with diminishing maximum signal levels, a double whammy that tends to limit the resulting dynamic range, in some cases rather severely.

Historically, unadorned analog channels have tended to have overall dynamic-range specifications on the order of 45–60 dB. This is true of such media as FM stereo, grooved records, optical film soundtracks, and magnetic tape. Perhaps the most limited hi-fi medium in this regard has been cassette tape, which manages about 55 dB at midband and on the order of 30 dB dynamic range at high frequencies.

Happily, DSP systems usually have flat overload characteristics and flat, or at least white, noise floors. The dynamic range of a DSP system is largely a question of the number of bits used to represent the signal. A rough rule of thumb is that each additional bit in the representation of sample values adds 6 dB to the dynamic range. The common 16-bit format used on digital compact discs yields a theoretical dynamic range of 96 dB. In practice, the combined effect of real-world imperfections and converter noise limits the practical attainable dynamic range with 16-bit audio to a little short of 96 dB, perhaps about 93 dB. A 24-bit converter can approach 144 dB dynamic range, which not only allows coverage of the full human auditory dynamic range, but allows margin for error in setting recording levels and subsequent mixing and processing operations.

18.4 Audio Components

With the common audio specifications now delineated, consideration can be given to some of the common audio components, their typical specifications, and specific characteristics and limitations.

18.4.1 Microphones

For any acoustical recording, the type, pickup pattern, position, and characteristics of the microphone are likely to have the greatest effect on the resulting quality of the recording. A microphone is, of course, an analog transducer, generally used to convert acoustic signals to proportional electrical signals. It is the functional analog of the eardrum in the human auditory system.

The term microphone was coined by Sir Charles Wheatstone in 1827, although at that point it referred to a purely acoustical device, such as a stethoscope. Its start as a transducer was pretty much coincident with Bell's invention of the telephone. Bell's initial *transmitter* used a vibrating needle connected to a diaphragm and immersed in an acid bath to implement a variable resistance responsive to acoustic waves [18.23]. This was followed in short order by more practical, rugged designs by Bell, Edison, Berliner, Hughes, and others.

Common to virtually all microphones ever made is a diaphragm. Most microphones do not directly sense the instantaneous pressure deviations of an airborne acoustic wave. Instead, the sound waves impinging on a diaphragm cause movement of the diaphragm, and it is this movement that is sensed and converted to an electrical audio signal.

The methods used for the mechanical-to-electrical transduction are many and varied. Bell's telephone patent seems to describe an odd type of electromagnetic microphone, in which the diaphragm was mechanically coupled to an armature in a magnetic circuit. Acoustic vibrations caused vibration of the armature, in turn inducing changes in the reluctance of the magnetic circuit, which induced a dynamic current in a coil of wire.

Later dynamic microphones simply attached a fine coil of wire to the diaphragm, and suspended it in a magnetic field. Movement of the diaphragm induces a voltage in the coil of wire, presumably proportional to the velocity of the diaphragm, and said voltage can

Amazon EX1041 Page 22

then be amplified as necessary. Dynamic microphones do not require an external source of power. Their popularity was limited prior to World War II by a lack of powerful permanent magnets.

Many early telephone transmitters used carbon in one form or another, typically carbon granules coupled to the diaphragm such that vibration of the diaphragm would dynamically squeeze the granules, causing a variation in resistance that could be electrically sensed by an applied voltage. This arrangement could be scaled up in size and voltage enough to produce a sizeable signal, of particular importance prior to the advent of electrical amplification.

Another type of self-generating microphone is the crystal microphone, based on the piezoelectric effect discovered by the Curies in 1880. Here, the diaphragm constricts a thin crystal made of Rochelle salt, ceramic, or similar material, directly producing a voltage. The low compliance of the piezoelectric element typically limits the performance of such a microphone, especially the extent of its frequency response.

Indeed, the performance of most of the microphone types mentioned so far is somewhat limited by the need to attach some object or structure to the diaphragm to detect its motion, increasing the mass of the moving elements, which in turn tends to limit the high-frequency response.

Two types of microphones avoid any mechanical attachment to the diaphragm: the ribbon microphone and the condenser microphone.

A popular variant of the dynamic microphone, the ribbon microphone, uses a thin metal ribbon diaphragm suspended in a magnetic field. Movement of the diaphragm induces a current in the diaphragm, which can then be amplified. In effect, the diaphragm acts as a single-turn moving coil. The ribbon microphone was developed by Schottky and Gerlach in 1923, and considerably refined by Harry Olsen of RCA in 1930 [18.26].

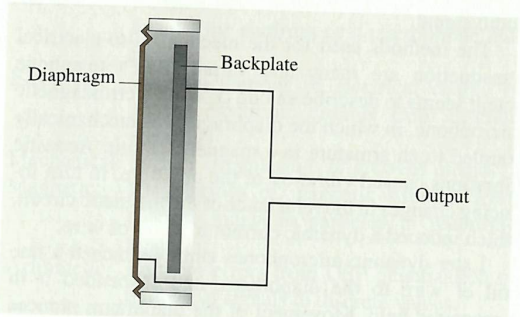


Fig. 18.9 A condenser microphone consists of a very thin diaphragm suspended parallel to a backplate (After [18.24])

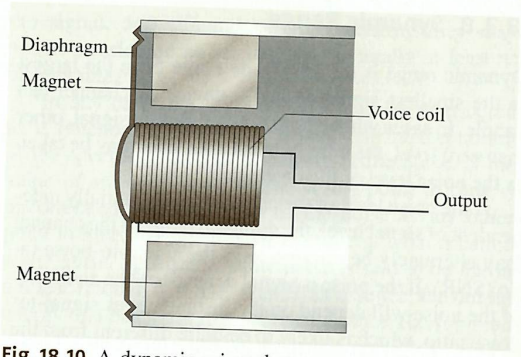


Fig. 18.10 A dynamic microphone uses a coil of wire attached to a diaphragm, suspended in a magnetic field (After [18.24])

The condenser microphone is based on the notion of electrical capacitance. A metal, or other conducting, diaphragm is placed close to a second, fixed conducting plate, forming a capacitor, and a polarizing voltage is applied between the plates. Movement of the diaphragm causes alterations in the capacitance, in turn causing corresponding alteration of the charge on the plates, which is converted to a dynamic voltage by a resistor, and amplified as necessary. The invention of the condenser

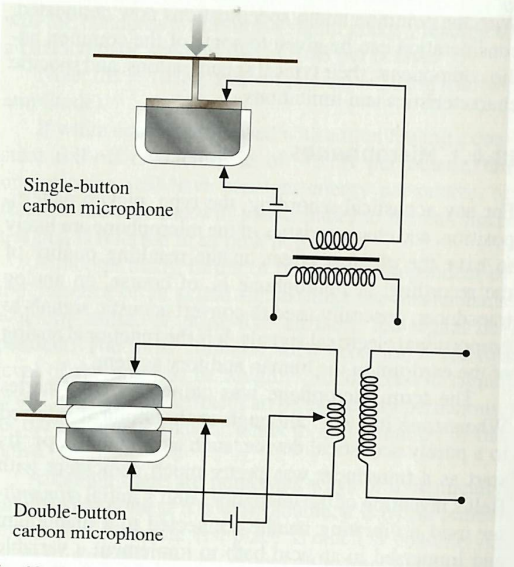


Fig. 18.11 A carbon microphone (After [18.25])

Amazon EX1041 Page 23

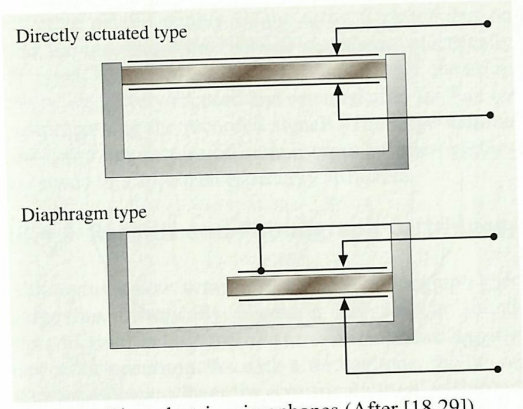


Fig. 18.12 Piezoelectric microphones (After [18.29])

microphone is generally credited to E. C. Wente of Bell Laboratories, in 1917 [18.27]. A popular modern variant is the electret condenser microphone, perfected in 1962 by Sessler and West of Bell Laboratories [18.28]. An electret is an object possessing a permanent electrostatic charge, making it rather like the electrical analog of the permanent magnet. Use of an electret to polarize the capacitor in a condenser microphone eliminates the need to supply a polarizing voltage for the capacitor. Since electret condenser microphones are used almost universally in cell phones, as well as other applications, they are probably the most popular type in current use. The diaphragm of electret condenser microphones is usually made of a thin sheet of flexible plastic which has had a metal coating deposited on it, a notably lightweight arrangement that permits good sensitivity and frequency response. Such microphones also usually exhibit good sample-to-sample consistency, making them practical for multi-microphone arrays.

A more recent variant of the condenser microphone is the silicon microphone. This is a microscopic condenser microphone constructed using micro electromechanical system (MEMS) technology. The required electrostatic polarizing field is provided by a charge pump, and the diaphragm can be free floating, being held in place by electrostatic attraction.

In recent years, the motion of a diaphragm has been measured using optical means, including laser interferometry, which may in time become a practical alternative to the capacitor microphone.

Microphone Dynamic Range

The noise level of a microphone is limited by Brownian motion of air molecules impinging on the diaphragm. Of course, it can be further limited by electrical noise of the amplification which follows it, but even discounting those effects, the so-called self-noise of a microphone is likely to be 10–20 dB higher than the threshold of hearing [18.30]. Lower self-noise is possible with larger diaphragms, say on the order of one inch, but this comes at a price of making the microphone more directional at high frequencies.

The loud signal limit of a microphone is often set by the maximum signal capacity of the associated preamplifier. Using large rail voltages can be of some aid in this regard. Eventually the diaphragm will reach the limits of its travel, and experience soft clipping. There has been some work on using force feedback to linearize the characteristics of the diaphragm [18.31, 32]. With a little effort and care, it is possible to make microphones that can handle high-level signals about as well as the human ear can.

Finally it should be noted that there has been some work on creating a microphone that does not require a diaphragm. Blondell and Chambers did some work on a flame microphone in 1902 and 1910 [18.23]. More recent efforts have used lasers to detect index of refraction changes due to acoustic pressure waves [18.33].

Microphone Pickup Pattern

Interacting with a real, three-dimensional sound field, a microphone is an audio device inextricably tied to spatial considerations. Immersed in the sound field, the microphone responds to the sound at essentially just one point, so its output represents a spatial sample of the total sonic event. This is not to say that it necessarily responds equally to sounds arriving from all directions; on the contrary, most microphones exhibit some directional bias, whether or not intended, and it is probably not possible to specify a single universally preferred pickup pattern.

The simplest arrangement is to put a sealed chamber behind the diaphragm, allowing only a slow leak for stabilization of quiescent atmospheric pressure. With short-term pressure behind the diaphragm fixed, the microphone responds to the sound pressure at the front of the diaphragm. Since pressure is a spatial scalar, it might be expected that a pressure microphone might respond equally to sound from all directions, and indeed this might be the case if the diaphragm had zero dimension. As it is, a finite-dimension diaphragm will spatially integrate the sound pressure across its surface. This will tend to increase output for normal-incident waves and waves with long wavelength, but will have less boosting effect for random incidence sounds, such as Brownian

Amazon EX1041 Page 24

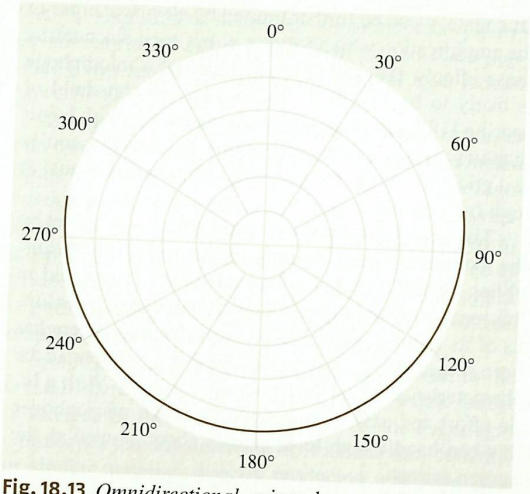


Fig. 18.13 *Omnidirectional* microphone pickup pattern. Note the slight attenuation of sounds arriving from the rear (After [18.35])

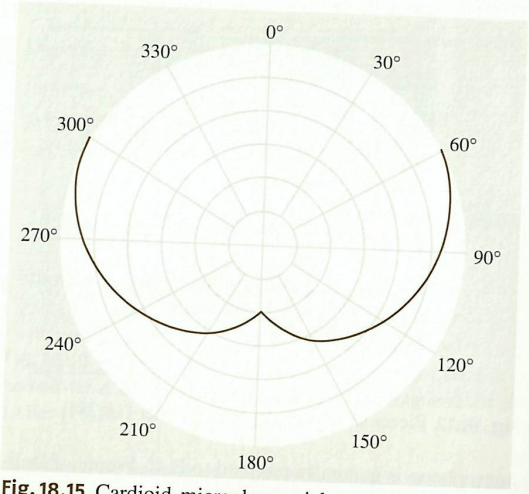


Fig. 18.15 Cardioid microphone pickup pattern, obtained by combining omnidirectional and bidirectional pickup patterns (After [18.35])

noise, and may reduce response to sounds with oblique incidence and wavelengths shorter than twice the diameter of the diaphragm. The resulting pickup pattern may therefore become directional at high frequencies, while remaining omnidirectional at lower frequencies. To remain omnidirectional up to 20 kHz, a diaphragm will have to have a maximum diameter of half the wavelength at the upper frequency limit, or about 8 mm. Whether by accident or design, this happens to be the approximate diameter of the human eardrum [18.34]. The fact that the eardrum is so small, and is coupled to the structures of the inner ear, and yet the ear still manages to exhibit a maximum sensitivity of 0 dB SPL or better has to be considered impressive performance, especially considering that quiet sounds just fade into perceived silence, rather than some inherent background perceptual noise.

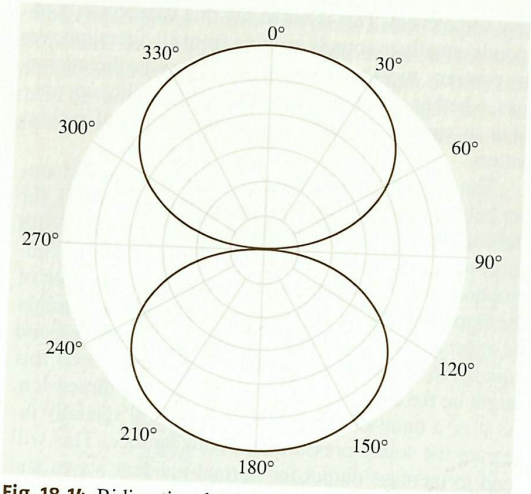


Fig. 18.14 Bidirectional microphone pickup pattern, typical of an open-back ribbon microphone (After [18.35])

More elaborate microphone pickup patterns (Figs. 18.13–18.15) can be achieved by increasing the number and complexity of acoustic paths to the diaphragm. For example, simply opening the back surface of the diaphragm to the atmosphere, the arrangement used in the classic open ribbon microphone, results in a microphone sensitive to the difference between front and back pressure, so sounds arriving from the side produce no response, and the pickup pattern usually resembles that of Fig. 18.8, albeit with the front arrival signal of opposite phase to the rear arrival signal.

If a pickup pattern as in Fig. 18.8 is summed appropriately with an omnidirectional pickup pattern, the rear arrival sound can be made to cancel, producing a cardioid front-directional pickup pattern.

With more elaborate channeling of the audio reading the diaphragm, more specialized or directional pickup patterns may be obtained.

Greater flexibility in the choice of pickup pattern can be obtained by using arrays of microphones, typically arranged in fixed physical orientations, where the

Amazon EX1041 Page 25

effective pickup pattern of the array is dependent on how the individual microphone signals are electrically summed. By recording the individual signals, the array can be effectively focused and oriented after the fact by postprocessing the recorded signals. The large-volume low-cost manufacture of certain types of microphones has made this approach especially attractive.

18.4.2 Records and Phonograph Cartridges

Although by now semi-obsolete, the phonograph cartridge has historically played a key role in sound reproduction, and is fairly close to a microphone in many aspects of operation. As with a microphone, the intent is to act as a transducer for acoustical vibration, in this case the vibration originating from a stylus following the undulations of a groove, rather than a diaphragm responding to vibrations in air.

Although early phonographs employed direct mechanical-to-acoustical transduction, the advent of electronic amplification made the electromechanical phono cartridge the default device for record playing.

Most of the methods of transduction used for microphones have been tried for phono cartridges as well, with piezoelectric and magnetic being the two most popular. Whereas most magnetic microphones have been of the moving-coil type, the magnetic cartridge design has employed both moving magnet and moving coil designs. Perhaps the use of a cantilever design of the stylus to reduce the effective mass of the coil or magnet seen by the stylus tip, coupled with the lack of a diaphragm, led to the preference of magnetic transduction for phonograph cartridges, while the use of capacitive transduction in condenser microphones has tended to be preferred in that transducer.

Despite the popularity of the grooved phonograph record for well over half a century, it has always been a difficult medium with which to achieve full fidelity.

For one thing, the stylus assembly is fraught with high-frequency resonance problems, much like the microphone, which have tended to limit the high-frequency response. In many cases, the high-frequency resonance has been underdamped, resulting in a serious peak in the mid-high-frequency response. The material used to mount the stylus in a way to secure it but still allow it to vibrate, often little bits of rubber, can sometimes degrade with age, impairing the characteristics of the cartridge. Then there is the need for the cartridge to drag a tonearm across the record with it.

In theory, the tonearm should provide a rock-stable platform for the cartridge at audio frequencies, so that only the stylus vibrates. Below the audio band, the tonearm should move as freely as possible, to follow the slowly spiraling groove and miscellaneous movements induced by the rotating disc, including horizontal eccentricity and vertical warp. In practice, this behavior cannot be achieved perfectly (since a perfect *brick-wall* filter is impossible), but at least the cartridge and tonearm should be matched, and the low-frequency resonance peak resulting from the compliance of the stylus and the effective mass of the tonearm should be properly damped, to avoid a peak either in the low-frequency audio response, or in the subsonic region where a resonance could make the system overly sensitive to record warp. Even in the best arrangement of properly damped resonance centered an octave or so below the audio band, the system is only second order, so some low-frequency tonearm vibration and some subsonic stressing of the stylus from pulling the tonearm is inevitable [18.36], which can lead to distortion and uneven frequency response. To avoid the build up of acoustic traveling waves in the tonearm, it is desirable to have its construction and material chosen to damp out such vibrations; not all tonearms have done this well.

Another problem that has hampered record playback, although not strictly the province of the phono cartridge, is that the playback stylus usually does not exactly follow the same path and orientation as the cutting stylus. This is because discs are cut on a cutting lathe, in which the cutter slowly traverses a straight-line path from the outer grooves to the inner grooves. The playback cartridge, in turn, is usually mounted on a pivoted arm, so its path is an arc. (The longer the tonearm, the smaller the disparity between record and playback orientations.) In order to minimize the path mismatch, a right dogleg bend has often been incorporated into the shape of the tonearm. The cartridge has to be carefully mounted on the tonearm so that the stylus is a specified distance from the pivot if this feature is to be properly exploited. Unfortunately, the bent tonearm traversing an arc results in a steady-state force from friction between the moving record groove and stylus tip, one which does not pull straight down the length of the tonearm, but instead generates a sideways force, called the skating force, which pulls the stylus tip against one or the other of the groove walls, which is another potential source of distortion. The skating force varies with groove speed and stylus tip size and shape, among other things, making it challenging to fully compensate.

At various times straight-line tracking tonearms have been produced, usually with the back of the tonearm mounted via bearings to a platform intended to move

Amazon EX1041 Page 26

attempts to apply complementary compression and expansion to record audio to improve the system dynamic range, notably discs produced by dbx Inc. in the 1970s.

Successful record playback can also run afoul of a problem known as acoustic feedback, if sound from the loudspeakers causes the disc itself to vibrate as a diaphragm. Solutions include the use of an inert turntable mat, remote positioning of the turntable, and limiting the playback level.

Of course, dust, dirt, scratches and wear will tend not only to degrade the recorded signal but also to raise the noise level, including the possible introduction of clicks and pops.

Given the litany of compromises and limitations in disc recording and playback, the level of performance that has been achieved with this medium is laudable. Although largely eclipsed by digital audio, some refinement of grooved record playback continues, including efforts to achieve the holy grail of disc playback, contactless playback via light or laser beam, notably by Fadeyev and Haber at Lawrence Berkeley National Laboratory [18.38]. The availability of advanced DSP has allowed after-the-fact improvements in noise reduction and distortion reduction that would have been impossible in an earlier age. It is tantalizing to consider the performance that might have been possible with this medium if such processing had been available and incorporated into the basic architecture of the disc recording system.

18.4.3 Loudspeakers

Loudspeakers and headphones provide complementary transduction to microphones, namely converting electrical audio energy to acoustical energy. Many of the principles and notions regarding microphones can be applied to loudspeakers either directly or by, in a sense, turning the ideas upside down.

In a manner reminiscent of microphones, few loudspeakers convert directly between electrical and acoustical energy. Instead, the electrical energy is converted to mechanical energy, specifically the movement of a diaphragm, often cone-shaped or dome-shaped, which then radiates acoustical energy, sometimes aided by an acoustical coupling horn or lens.

Indeed, many of the same transduction methods that have been used in microphones have at least been tried for loudspeakers including magnetic (also known as *dynamic*), piezoelectric, and condenser.

The genesis of loudspeaker-like transducer actually predates Bell's invention of the telephone by a couple of years when, in 1874, Ernst W. Siemens described a dynamic moving-coil transducer, although he does not seem to have initially done much with it [18.39].

Bell's telephone, patented in 1876, also used a dynamic receiver. With a magnetic armature mechanically coupled to a diaphragm secured at its edges, it was probably closer in spirit to a headphone than a loudspeaker.

There followed a series of refinements that progressively brought the output transducer closer to the modern concept of a loudspeaker. In 1898, Oliver Lodge patented a spacer to maintain the air gaps in the magnetic circuit. Three years later, John Stroh came up with the notion of attaching the cone to the frame via a flexible, corrugated surround that improved the linear travel of the cone. Proper, flexible centering of the rear of the cone in the magnetic gap was provided by the *spider*, developed by Anton Pollak in 1909.

Many of these early loudspeakers used an acoustic horn of some sort to project the loudspeaker output into the room. The horn improves efficiency, partly by improving the acoustic coupling from the loudspeaker to the room, and partly by collimating the sound into a beam directed at the listener. Credit for the first true direct radiator loudspeaker design to include all the elements of what can be regarded as the modern-day loudspeaker is generally accorded to Rice and Kellogg in 1925 [18.40]. Their design took explicit account of the baffle on which the driver was mounted in the overall acoustic design.

There were of course, countless follow-on refinements to the basic electrodynamic driver, a process that continues to present day. Other methods of transduction were also explored. In 1929, E. W. Kellogg filed a patent for an electrostatic loudspeaker, consisting of panels not unlike large condenser microphones. The design has not been as widely accepted as the dynamic driver, in part due to some limitations in maximum output. There have also been piezoelectric drivers, but the limited travel of the piezoelectric element has limited their use to mid-range drivers, tweeters, and earphones.

So, as the Rice/Kellogg baffle indicated, with the basic dynamic loudspeaker configuration established, much of the attention focused on refining the loudspeaker/cabinet combination as a system.

The output ultimately obtained from a raw loudspeaker driver is profoundly affected by the surrounding baffle or cabinetry. In isolation, a loudspeaker driver produces two acoustical outputs: one from the front of the diaphragm or cone, and the second from the back. Unfortunately, the two outputs are out of phase. Were they generated at the same place and time, they would can-

Amazon EX1041 Page 27

cel; the loudspeaker would produce nothing but heat. So part of the goal in designing a complete loudspeaker system is to keep those two outputs from annihilating each other.

At frequencies above the point where the wavelength is comparable to the size of the speaker cone, the radiation will tend to be a directional beam, and the front and back waves will tend to naturally go their own separate ways. At lower frequencies, the front and back radiation will tend to be more omnidirectional, and will therefore tend to cancel. This can be avoided by requiring that the size of the loudspeaker be at least as large as the wavelength of the lowest frequency of interest. For a low-frequency limit of 20 Hz, this would correspond to a diameter of at least 15 m, which at the very least would make for rather unwieldy boom boxes.

Happily, it is not necessary to rely on the loudspeaker itself to avoid low-frequency cancellation. Instead, a smaller driver can be mounted on a baffle, and the baffle can be used to prevent interaction of the front and back radiation. Indeed, a common concept in loudspeaker system theory is that of the *infinite baffle*, which precludes cancellation at any frequency, at the cost of being physically impossible. Even assuming a willingness to forego response below 20 Hz, a baffle of at least 15 m would still be required, which is not much more practical than an unadorned 15 m woofer. This leads to the notion of putting the loudspeaker in a box, or loudspeaker cabinet. (A speaker mounted in a car door or trunk lid is operating in an approximate infinite baffle, although the car interior can be considered a semi-sealed cabinet.)

The simplest arrangement is to make the box completely sealed, so that the rear radiation is trapped within the box, and cannot emerge to cancel the front radiation. Of course, this requires making the box quite rigid, else the rear speaker radiation will simply be conducted through the walls of the box. One problem with sealed-box systems is that the springiness of the air in the cabinet will decrease the natural compliance of the loudspeaker, so that a driver with a respectably low free-air resonance may have a much higher resonance mounted in a sealed box, with correspondingly reduced low-frequency response.

This issue was addressed in 1954 by Edgar Villchur, in a manner which in hindsight seems remarkably straightforward, namely increasing the compliance of the speaker to obtain the desired performance when it was mounted in a sealed box. This approach, referred to as an acoustic suspension design, made the cabinet an inherent part of the design, and is still in use today.

The alternative to the sealed-box loudspeaker enclosure is, logically enough, the vented box, in which at least a portion of the back radiation of the driver is allowed to emerge from the cabinet, often delayed or phase-shifted to better align it with the front radiation. One of the early examples of this approach was the bass reflex design by Albert Thuras of Bell Laboratories, in 1930. Other vented designs have used elaborate internal tubing, internal ductwork (such as the Klipsch folded horns), and or provided mass loading of the exit port via a diaphragm (the *passive radiator*). The goal is to improve efficiency, partially by relieving some of the back pressure behind the driver, but mostly by adding some of the back radiation to the front radiation in a nondestructive fashion, to achieve some increase in output, typically on the order of 3 dB.

More recently, hybrid enclosures have been developed, sometimes called bandpass enclosures, in which a vented or sealed box is used behind the driver and a vented box is used in front of the driver, to provide improved low-frequency efficiency over a somewhat narrow range of frequencies [18.41]. This design is popular in some smaller low-frequency speakers.

Regardless of its other characteristics, a bare loudspeaker cabinet is likely to impart a variety of spurious resonances to the response from sound waves bouncing helter-skelter around the inside of the cabinet. For this reason, loudspeaker cabinets are often lined or stuffed with sound-absorbing material. This alters the overall response as well as damping the resonances, and must be accounted for in the overall design.

Much of the modern modeling of loudspeaker systems is based on the work of Thiele and Small in the 1970s, who established the basic parameters governing the performance of loudspeaker systems, and developed accurate mathematical models incorporating those parameters, allowing rapid and precise mathematical prototyping [18.42].

Still, the implications of achievable, cost-effective drivers and practical cabinet size is that there is precious little that can be done to get usable bass response down to 20 Hz from a small loudspeaker system. It is possible to apply low-level equalization to the signal upstream of the loudspeaker, but this can increase the cost and the power requirements, and will not be effective unless the driver can handle increased signal level without serious distortion. The closed-box design is considered one of the better choices for the use of equalization, as its response roll-off below primary resonance is only 12 dB per octave, which is gentler than most vented systems.

Amazon EX1041 Page 28

Most of the issues relating to speaker cabinetry pertain to low frequencies. Midrange and high-frequency drivers may also have their backs sealed, but the wavelengths are shorter, and the required resonant frequencies are a lot easier to attain.

The notion of using multiple drivers in a loudspeaker system seems to have originated in 1931 with a two-way system designed by Frederick of Bell Laboratories [18.43]. A three-way system followed within two years. It is simply much easier to design a system to have flat frequency response with multiple drivers, optimized to handle a subset of the audio band, than to try to accomplish the task with a single driver.

Part of the motivation for using multiple drivers is that the loudspeaker, like the microphone, is an inherently spatial entity. Sound radiates from a loudspeaker in all directions, depending on angle and frequency. In an anechoic environment, with a listener situated on the main axis of the loudspeaker, the response on that axis is well defined, and the only one that really matters. In a real room, however, the multidirectional acoustical output of the loudspeaker will reflect off walls and other surfaces, generally producing a complicated multiplicity of arrivals at the listener's ears. There is no easy way to characterize this process, or of the resulting perception, although in the 1970s, Roy Allison made significant strides in characterizing the loudspeaker–room interaction at low frequencies [18.44]. However, it certainly establishes the importance of the loudspeaker spatial response at all angles, and suggests the desirability of a loudspeaker having a consistent radiated frequency response (radiation pattern) at any angle.

This is another area in which multiple drivers can be beneficial. Again, the output of a driver will begin to get directional at a frequency corresponding to a wavelength about twice the diameter of the driver. Assuming consistently wide dispersion is desired, a 30 cm woofer will be useful up to about 500 Hz, a 7.5 cm midrange continuing to 2 kHz and a 2.5 cm tweeter maintaining wide dispersion to at least 6 kHz. These figures may be conservative, in part because controlled decoupling of the speaker cone can result in a progressively smaller radiation area at successively higher frequencies, decreasing the effective diameter and broadening the resulting dispersion. Other approaches to maintaining a desired radiation pattern across frequency include the use of acoustic lenses and phased arrays of multiple drivers within a given frequency range. It is rare for a speaker to have truly constant directionality across its entire pass-band, in part from that fact that most are at least somewhat directional at mid and high frequencies, and, because of the long wavelengths involved, almost unavoidably omnidirectional at low frequencies. (Such long wavelengths may be difficult to localize, enabling the use of satellite systems with only one woofer to support two or more channels.)

The use of two or more frequency-selective drivers generally imposes the need for a crossover network to apportion the signal on the basis of frequency. Not only does this avoid wasting power, but it can be absolutely necessary to avoid overdriving or burning-out a driver. Some care is generally required in the transition frequency region to maintain flat response, and although there are well-regarded canonical topologies, such as the Linkwitz–Riley crossover developed in 1976 [18.45], real-world deviations from ideal driver behavior may engender the use of recursive approximation techniques to optimize system response fully.

The combination of crossover networks, drivers, and room loading will often result in a loudspeaker having an electrical input impedance characteristic that is strongly frequency dependent, and which poses a challenge for a power amplifier.

As is no doubt evident, a loudspeaker system and its associated radiation pattern is a fiendishly complicated device that does not readily lend itself to characterization by traditional audio metrics or specifications. In consideration of such specifications, it is difficult to make definitive statements that fully characterize the sound of a speaker.

The bandwidth of most speakers may approach the ear's 20–20 kHz bandwidth, but few speakers make it all the way, particularly at low frequencies, and bandwidths on the order of, say, 70–14 kHz are far more common, at least in consumer equipment.

The deviation of the frequency response is difficult to talk about meaningfully, but it seems desirable for speakers to be within 1 dB of flat response across their useful bandwidth, at least on axis, and better speakers can approach or achieve this, although at low frequencies the room exerts enough effect as to likely require tuning or experimentation with speaker position to maintain a flat response. The response at other angles should probably be smooth and well controlled, but the exact optimal specification has not been established. Perhaps the ideal speaker would adjust its radiation pattern to match the instruments being reproduced. (But then what is the preferred radiation pattern of a purely electronic instrument, such as a synthesizer?) Although it is necessary to measure the response of a loudspeaker in a room in order to accurately gauge the low-frequency response, a direct narrow-band or sine-wave measure-

Amazon EX1041 Page 29

ment of the system response in a room at middle and high frequencies is likely to yield a useless, mad scribble of a response, owing to the complex interactions of the echoes at the point of measurement. Fortunately, the perceived response should be a lot smoother than what is measured, in part because the ear differentiates sound arrivals on the basis of direction. Measurement techniques that attempt to address this include the use of directional microphones, anechoic and semi-anechoic measurement, and time-gated response measurement.

Timing precision is another specification that is difficult to pin down with loudspeakers, again partly due to the effects of room echoes. In one respect, most speakers do pretty well in first-arrival timing, since the ear's temporal pre-masking resolution is on the order of 1–2 ms, corresponding to about 30–60 cm of path length, and speaker systems with all drivers mounted on a common front panel will generally be well within this tolerance for inter-driver path-length differences. Still, audible time disparities may result from primary resonances, particularly at low frequencies, and high-order crossovers.

In the area of distortion, loudspeakers also do pretty well, all things considered. At normal listening levels, a properly designed loudspeaker exhibits fairly low nonlinearity, with a fairly gentle limiting characteristic, leading to low IM distortion and fairly low THD, mostly third order, which is likely to be masked by most signals. Most properly operating loudspeakers do not produce drastic amounts of distortion until they reach the limits of their travel. To be sure, some loudspeakers can sound audibly harsh, but this is not necessarily a byproduct of distortion, it may be more related to an uneven response or radiation pattern. One bugaboo of many drivers is cone breakup, wherein the cone no longer acts as a piston. Although this can look unsightly in slow motion, it will not necessarily produce distortion unless there is nonlinearity present, and most speaker cones, being passive structures of paper, plastic, or metal in and of themselves, are not especially nonlinear. While such breakup is likely to affect the response and/or the radiation pattern of the speaker, it may not produce much distortion per se.

Although the basic dynamic loudspeaker has now been in use for over 70 years, and has been greatly refined in that time, it is still the subject of active research, in both design and evaluation. One of the holy grails yet to be achieved is, like the diaphragm-free microphone, the cone-less loudspeaker that more directly converts electrical energy to acoustical energy. Aside from eliminating some troubling mechanical elements and their associated resonances, this might improve the efficiency, which is currently typically around 1% for consumer loudspeakers.

Another long-sought goal is the projection loudspeaker, which can make a sound appear to emanate from a specified remote location. The have been some encouraging results in the use of ultrasonics [18.46] and large-scale speaker arrays [18.47] in this regard.

18.4.4 Amplifiers

Of all the common audio analog devices, electronic amplifiers of all types probably have the easiest time meeting the preferred specifications of a good-quality audio system. This in no small part derives from the fact that they generally contain no mechanical processing stages, and simply have to push a bunch of very lightweight electrons around. Properly designed, an amplifier should have little difficulty achieving a dynamic range approaching 120 dB, nor is it likely to deviate significantly from the preferred specifications for bandwidth, response deviation, phase/timing response, or distortion.

This is despite the fact that most amplification devices, mainly vacuum tubes, transistors, and field-effect transistors (FETs), are not themselves inherently very linear. In the years following de Forest's invention of the triode, considerable and rapid strides were made in not only the refinement of the devices themselves, but in circuit designs which optimized the linearity and response of the complete amplifiers. A major milestone was the invention in 1928 of the negative-feedback amplifier by *Black* of Bell Laboratories [18.9]. This arrangement compares the output signal with the input signal, and to the extent that they may tend to differ, generates an instantaneous correction signal. Successful application of negative feedback carries with it specific requirements on the performance of the raw amplifier, such as maximum phase shift, and some of these requirements were still being uncovered years after the initial development of negative feedback.

Several standard types of amplifiers are commonly found in audio use, each with specific requirements and design challenges, including preamplifiers, line amplifiers, power amplifiers, and radio frequency (RF) and intermediate frequency (IF) amplifiers for use in radio. Preamplifiers, for example, are required to handle the generally very small signal from transducers such as microphones, phonograph cartridges, tape heads, and optical photocells. Linearity is usually not a major problem, because of the small signals involved, but low noise and impedance matching to the transducer are often sig-

Amazon EX1041 Page 30

22.1 Dynamics of the Linear Single-Degree-of-Freedom (1-DOF) Oscillator

In the low-frequency range, the linear motion of an electrodynamical loudspeaker, for example, can be described by a single-degree-of-freedom oscillator. This is also the case for any complex structure considered as a rigid body for which the equivalent stiffness is represented by a unique spring, and where all the dissipation mechanisms are approximated by a unique dashpot or *fluid damping* constant; one can think of the rigid-body mode of vibration of an automobile. Also the 1-DOF oscillator has fundamental mathematical properties of great interest in the field of vibrations, since linear vibrations of discrete and continuous structures can be represented by a set of oscillators, as will be shown in the following sections.

The motion of a standard 1-DOF oscillator is governed by a second-order differential equation involving an excitation force $F(t)$ and the response of the structure. Using a formulation with the velocity $v(t)$, we get

$$F = M\frac{\mathrm{d}v}{\mathrm{d}t} + Rv + K\int v\,\mathrm{d}t\,. \tag{22.1}$$

This equation accounts for the motion of a mass M subjected to a restoring force due to a linear spring of stiffness K and a dashpot R which represents a linear viscous fluid damping (Fig. 22.1).

Equation (22.1) can be written alternatively, using the mass displacement $\xi(t)$ and the usual reduced parameters:

$$F = M\left(\frac{\mathrm{d}^2\xi}{\mathrm{d}t^2} + 2\zeta_0\omega_0\frac{\mathrm{d}\xi}{\mathrm{d}t} + \omega_0^2\xi\right)\,, \tag{22.2}$$

where $\omega_0 = \sqrt{M/K}$ is the eigenfrequency of the lossless system and $\zeta_0 = R/2M\omega_0$ is the nondimensional damping coefficient.

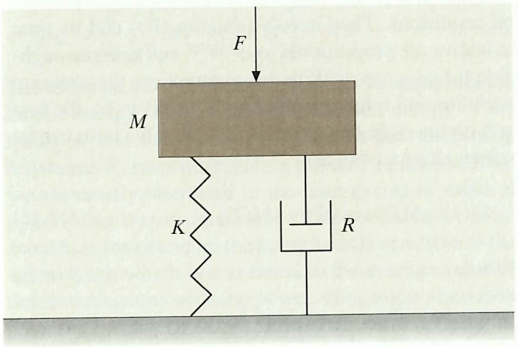


Fig. 22.1 Single-DOF oscillator

22.1.1 General Solution

One strategy for solving a linear equation such as (22.2) is to make use of the Laplace transform [22.1]. Using this transform, (22.2) becomes

$$s^2\mathcal{X}(s) - s\xi(0) - \dot{\xi}(0) + 2\zeta_0\omega_0\left[s\mathcal{X}(s) - \xi(0)\right] + \omega_0^2\mathcal{X}(s) = \frac{\mathcal{F}(s)}{M}\,, \tag{22.3}$$

where $\mathcal{X}(s)$ and $\mathcal{F}(s)$ are the Laplace transforms of $\xi(t)$ and $F(t)$, respectively. Using the tables and properties of the Laplace transforms, we obtain the general solution

$$\xi(t) = \frac{1}{M}\int_0^t F(\theta)G(t-\theta)\,\mathrm{d}\theta + \left[2\zeta_0\omega_0\xi(0) + \dot{\xi}(0)\right]G(t) + \xi(0)\dot{G}(t)\,, \tag{22.4}$$

where $G(t)$ can take different forms, depending on the value of the damping coefficient ζ_0:

- if $\zeta_0 < 1$, which corresponds to the so-called *underdamped* case, then

$$G(t) = \mathrm{e}^{-\zeta_0\omega_0 t}\frac{\sin\omega_0\sqrt{1-\zeta_0^2}\,t}{\omega_0\sqrt{1-\zeta_0^2}}\,; \tag{22.5}$$

- if $\zeta_0 = 1$, which corresponds to the *critical* case, we obtain

$$G(t) = t\,\mathrm{e}^{-\omega_0 t}\,; \tag{22.6}$$

- finally, if $\zeta_0 > 1$, which corresponds to the *overdamped* case, we get

$$G(t) = \mathrm{e}^{-\zeta_0\omega_0 t}\frac{\sinh\omega_0\sqrt{\zeta_0^2-1}\,t}{\omega_0\sqrt{\zeta_0^2-1}}\,. \tag{22.7}$$

22.1.2 Free Vibrations

Free vibrations are characterized by the absence of an external force. With $F(t) = 0$ in (22.4), we get

$$\xi(t) = \left[2\zeta_0\omega_0\xi(0) + \dot{\xi}(0)\right]G(t) + \xi(0)\dot{G}(t)\,, \tag{22.8}$$

Amazon EX1041 Page 31