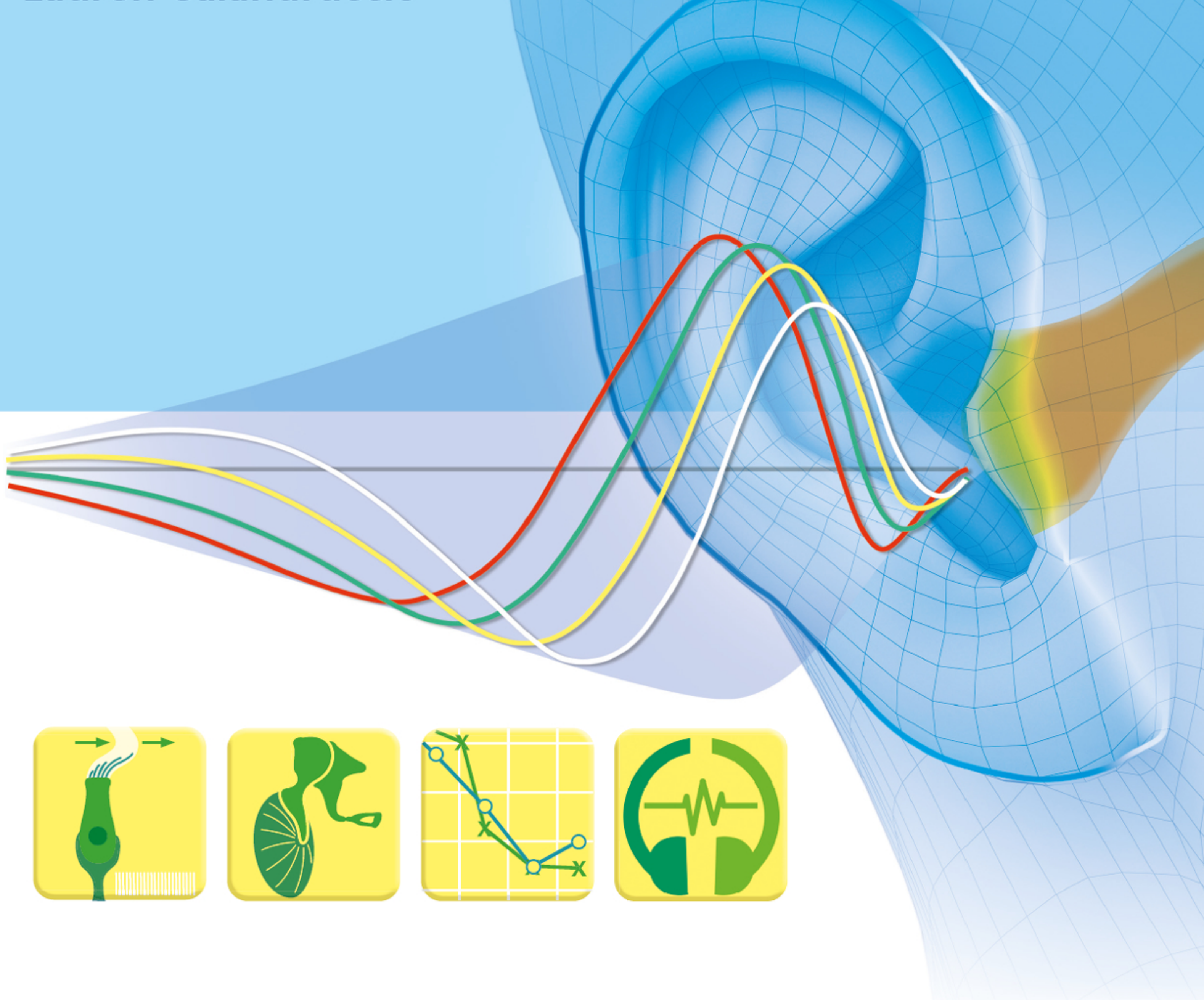
Essentials of Audiology

Fifth Edition

Stanley A. Gelfand
Lauren Calandruccio



Thieme

Essentials of Audiology

5th Edition

Stanley A. Gelfand, PhD, CCC-A/SLP
Professor
Department of Linguistics & Communication Disorders
Queens College of the City University of New York (CUNY)
Flushing, New York;
AuD Program & PhD Program in Speech-Language-Hearing Sciences
CUNY Graduate Center
New York, New York, USA

Lauren Calandruccio, PhD, CCC-A
Louis D. Beaumont University Professor
Department of Psychological Sciences
Case Western Reserve University
Cleveland, Ohio, USA

448 illustrations

Thieme
New York • Stuttgart • Delhi • Rio de Janeiro

Library of Congress Cataloging-in-Publication Data is available with the publisher.

Important note: Medicine is an ever-changing science undergoing continual development. Research and clinical experience are continually expanding our knowledge, in particular our knowledge of proper treatment and drug therapy. Insofar as this book mentions any dosage or application, readers may rest assured that the authors, editors, and publishers have made every effort to ensure that such references are in accordance with **the state of knowledge at the time of production of the book**.

Nevertheless, this does not involve, imply, or express any guarantee or responsibility on the part of the publishers in respect to any dosage instructions and forms of applications stated in the book. **Every user is requested to examine carefully** the manufacturers' leaflets accompanying each drug and to check, if necessary, in consultation with a physician or specialist, whether the dosage schedules mentioned therein or the contraindications stated by the manufacturers differ from the statements made in the present book. Such examination is particularly important with drugs that are either rarely used or have been newly released on the market. Every dosage schedule or every form of application used is entirely at the user's own risk and responsibility. The authors and publishers request every user to report to the publishers any discrepancies or inaccuracies noticed. If errors in this work are found after publication, errata will be posted at www.thieme.com on the product description page.

Some of the product names, patents, and registered designs referred to in this book are in fact registered trademarks or proprietary names even though specific reference to this fact is not always made in the text. Therefore, the appearance of a name without designation as proprietary is not to be construed as a representation by the publisher that it is in the public domain.

Thieme addresses people of all gender identities equally.We encourage our authors to use gender-neutral or gender-equal expressions wherever the context allows.

Thieme Publishers New York
333 Seventh Avenue, 18th Floor
New York, NY 10001, USA
www.thieme.com
+1 800 782 3488, customerservice@thieme.com

Cover design: © Thieme
Cover image source: © Thieme/Martina Berge, using ear © SciePro/stock.adobe.com and headphones © kmsdesen/stock.adobe.com
Typesetting by TNQ Technologies, India

Printed in USA by King Printing Company, Inc. 5 4 3 2 1

ISBN 978-1-68420-398-7

Also available as an e-book:
eISBN (PDF): 978-1-68420-399-4
eISBN (epub): 978-1-63853-106-7

To Janice

In loving memory

Stanley A. Gelfand, PhD, CCC-A/SLP

To my students

Past, present, and future

Lauren Calandruccio, PhD, CCC-A

Contents

Preface ix

1. Acoustics and Sound Measurement 1
2. Anatomy and Physiology of the Auditory System 30
3. Measurement Principles and the Nature of Hearing 70
4. The Audiometer and Test Environment 91
5. Pure Tone Audiometry 107
6. Auditory System and Related Disorders 135
7. Acoustic Immittance Assessment 187
8. Speech Audiometry 220
9. Clinical Masking 255
10. Behavioral Tests for Audiological Diagnosis 279
11. Physiological Methods in Audiology 309
12. Assessment of Infants and Children 341
13. Audiological Screening 365
14. Nonorganic Hearing Loss 391
15. Audiological Management I 411
16. Audiological Management II 430
17. Effects of Noise and Hearing Conservation 488

Appendices 533

Subject Index 551

Author Index 566

Preface

What is audiology? *Audiology* is the clinical profession that deals with hearing and balance disorders. It is also the scientific study of normal and abnormal audition and related areas in the broadest sense.

What is an audiologist? An *audiologist* is a practitioner of audiology as a clinical profession. Audiologists are principally concerned with the identification, evaluation, and management of patients with auditory and balance disorders, as well as with the prevention of hearing impairment. The scope of audiological practice also includes such diverse areas as the evaluation of the vestibular system, noise assessment, and hearing conservation, as well as the physiological monitoring of various neurological functions during surgical procedures. As a result, audiologists possess a broad scope of knowledge and skills, and often have interests in common with a variety of other disciplines such as speech-language pathology, speech and hearing science, education of the deaf and hard-of-hearing, engineering, acoustics, industrial hygiene, musicology, medicine, physiology, psychology, linguistics, and vocational counseling.

Much like many other scholarly professions, pursuing a career in audiology involves a rigorous course of doctoral-level education and training. Most audiologists earn the Doctor of Audiology (AuD) degree, while others with research and academic as well as clinical interests obtain a PhD. Some pursue both. In addition, those qualified to practice audiology must possess professional licenses from the states in which they practice, and are usually certificated by the American Speech-Language-Hearing Association (ASHA) and/or the American Academy of Audiology (AAA).

So who is this book for? Introductory audiology is an essential and fundamental aspect of the education of all students who are interested in the two related professions of speech-language pathology and audiology. This book is primarily intended to serve as a comprehensive introductory text for students who are preparing to enter both of these fields. As such, it tries to address the needs of at least two rather different groups of students. Those planning a career in audiology need a broad overview of the field and a firm understanding of its many basic principles so that they have a solid foundation for the future as doctors of audiology in clinical practice.

The audiological needs of future and practicing speech-language pathologists are just as important, and go well beyond knowing the auditory implications of speech, language, and related disorders, and being able to understand audiological reports. Speech-language pathologists often find themselves working hand-in-hand with their audiological colleagues. They also need to perform certain audiological procedures themselves when these fall within the speech-language pathology scope of practice, especially when screening is involved; and they regularly make interpretations and referrals that are of audiological relevance. Moreover, speech-language pathologists often work with patients with a broad range of hearing and related disorders directly and on an ongoing basis. They frequently must explain the nature and management of auditory disorders to family members, teachers, and other professionals. This is especially true in school settings and long-term care facilities. What is more, cochlear implant and other multidisciplinary programs are enhancing the scope and depth of interactions among speech-language pathologists and audiologists, and are making knowledge and understanding of audiology all the more important for budding speech-language pathologists. With considerations like these in mind, we hope that students who become speech-language pathologists will find this text useful as a reference source long after their audiology courses have been completed. (Of course, we do admit hoping that at least a few speech-language pathology students will be attracted to a career in audiology by what they read here.)

This textbook attempts to provide a comprehensive overview of audiology at the introductory level, including such topics as acoustics, anatomy and physiology, sound perception, auditory disorders and the nature of hearing impairment, methods of measurement, screening, clinical assessment, and clinical management. It is intended to serve as the core text for undergraduate students in speech, language, and hearing, as well as to serve the needs of graduate students who need to learn or review the fundamentals of audiology. It is anticipated that the material will be covered in a one-, two-, or three-term undergraduate sequence, depending on the organization of the communication sciences and disorders curriculum at a given college or university. For example, the first three chapters are often used in the text for an undergraduate hearing science course, while selections from the other chapters might be used in one or two audiology courses.

With these considerations in mind, we have tried to prepare a textbook that is extensive enough for professors to pick and choose material that provides the right depth and scope of coverage for a particular course. For example, text readings can be assigned to cover clinical masking at almost any level from simple to complex by selecting various sections of Chapter 9. It is unlikely that all of that chapter will be assigned in a single undergraduate class.

However, the material is there, if needed, for further study to provide the groundwork for a term paper or independent study report, or for future reference. We also have tried to provide relatively extensive and up-to-date reference lists for similar reasons.

This fifth edition was undertaken to provide the student with an up-to-date coverage of a field that is steadily developing, as well as to take advantage of accumulating experience to improve upon what is included and how it is presented. Many developments and changes have taken place since the prior edition was published. Some of them are in areas of rapidly unfolding development like cochlear implants, hearing aids and related technologies, as well as in electrophysiological assessment. But most of them are the slow, methodical, and often subtle—albeit important—advances that unfold over time in an active clinical science. Other changes reflect the influence of systematic reviews, changes in guidelines, expert position papers, standards and regulations which affect clinical practices, and technical matters. Of course, there are always a few developments that surface the day *after* the textbook is printed—a frustration to the textbook authors, but the kind of thing that makes audiology such an exciting and interesting field.

As with the prior editions, this one was influenced by the input graciously provided by many audiologists involved in clinical practice, research, and teaching and student supervision. In addition, considerable attention was given to the comments and insights of students who were taking or recently completed audiology courses, including those who used the prior edition of this textbook as well as other books. The content and especially the style of the text were substantially influenced by their advice. As a result of their insights, the current edition retains a writing style that has been kept as conversational and informal as possible, and only classroom-proven examples and drawings are included. Similarly, clinical masking, acoustic immittance, and screening have been kept in separate chapters; the material on audiological management continues to be spread over two chapters; and the history of audiology has been omitted.

Many old figures have been updated or replaced and others have been added to keep the material up-to-date. In addition, many older and dated photographs have been replaced to keep the book contemporary and fresh for the reader. More importantly, we have attempted to use figures that reflect and encourage diversity and inclusiveness.

All four prior editions used gender-specific pronouns (*he, she, him, her,* etc.) to maximize clarity for the benefit of the reader, which was well-received although originally undertaken with great trepidation. However, we strongly believe that non-gendered language (e.g., singular *they*) is more appropriate and inclusive, and that it can be incorporated into the text while keeping the material maximally reader-friendly (or at least minimally unfriendly). As a result, we have changed the previously used gender-specific pronouns to gender-neutral wording whenever possible throughout the text.

This book would not exist without the influence of many very special people, with our sincerest apologies to anyone inadvertently omitted.

From Stanley A. Gelfand: I would like to extend my heartfelt thanks to the following very special individuals for their influence, insights, advice, encouragement, assistance, support, devotion, and friendship over the years: Marilyn Agin, Matthew Bakke, Moe Bergman, Arthur Boothroyd, Lauren Calandruccio, Tong Chun, Joseph Danto, Lillian and Sol Gelfand, Judy Gravel, Irving Hochberg, Gertrude and Oscar Katzen, Renee Kaufman, Barbara Kruger, Fred Kruger, Harry Levitt, John Lutolf, Maurice Miller, Edward O'Connor, Sharon O'Connor, Neil Piper, Leslie Ross, Mark Ross, Adrienne Rubenstein, Brian Scanlan, Teresa Schwander, Gerald Studebaker, Shlomo Silman, Carol Silverman, Helen and Harris Topel, Ellen and Glenn Unger, Barbara Weinstein, Mark Weiss, and Joyce West. I am particularly grateful to my current and former colleagues and students in the Department of Linguistics and Communication Disorders at Queens College, the AuD Program and the PhD Program in Speech Language and Hearing Sciences at the City University of New York Graduate Center, and the East Orange VA Medical Center.

From Lauren Calandruccio: To all my amazing mentors, collaborators, and my go-to audiology peeps, thank you! A special thanks to Harvey Abrams, Gin Best, Karen Beverly-Ducker, Kristina Bowdrie, Ann Bradlow, Emily Buss, Laurel Carney, Angela Ciccia, Nicole Corbin, Sumit Dhar, Karen Doherty, Ann Clock Eddins, Andy Faulkner, Erick Gallun, Stan Gelfand, Tucker Gleason, Larry Humes, Lisa Hunter, David Kemp, Wei Wei Lee, Lori Leibold, Nancy Nelson, Gayla Poling, Nia Potier, Stuart Rosen, Sharon Sandridge, Jessica Sullivan, Lee Thompson, Kathy Vander Werff, Peter Wasiuk, and Sharon Williams, and in memory of Judy Gravel and Lu Feng Shi. To Dr. H (Nick Hipskind), thank you for being the first to show me the (cone of) light.

We would also like to express our appreciation to the extraordinary, talented, and dedicated professionals at Thieme Medical Publishers, who have been so helpful, cooperative, and supportive throughout the process of preparing this book.

Finally, our greatest appreciation is expressed to our families.

From Stanley A. Gelfand: To Janice, whose memory will always be a blessing and inspiration; and to my wonderful children, Michael and Hillary, Joshua and Erin, and Jessica and Bobby for their love, encouragement, support, insight,

and unimaginable patience; as well as to my amazing grandchildren, Jacob, Alexander, Jackson, Benjamin, and Joey for the special love and joy that only they can bring.

From Lauren Calandruccio: To Jamie and Aunt Sis, you are the “fun” to my “serious,” thank you for the balance. To my parents, thank you for teaching me how to work hard. That work ethic allowed me to complete this project. To my two Ks, thank you for always cheering me on and for your love and patience (especially when mom has to work). Eric, 20 years and you are still my rock. Your support is the foundation to all my successes. Thank you for everything.

Stanley A. Gelfand, PhD, CCC-A/SLP
Lauren Calandruccio, PhD, CCC-A

1 Acoustics and Sound Measurement

We begin our study of audiology by reviewing the nature of sound because, after all, sound is what we hear. The science of sound is called acoustics, which is a branch of physics, and relies on several basic physical principles. Many useful sources are available for students wishing to pursue the areas covered in this chapter in greater detail (e.g., Peterson & Gross 1972; Hewitt 1974; Kinsler, Frey, Coppens, & Sanders 1982; Sears, Zemansky, & Young 1982; Beranek 1986; Gelfand 2018).

Physical Quantities

The basic physical quantities are mass, time, and length (or distance). All other physical quantities are derived by combining these three basic ones, as well as other derived quantities, in a variety of ways. The principal basic and derived quantities are summarized in **Table 1.1**. These basic quantities are expressed in terms of conventional units that are measurable and repeatable. The unit of **mass (*M*)** is the **kilogram (kg)** or the **gram (g)**; the unit of **length (*L*)** is the **meter (m)** or the **centimeter (cm)**; and the unit of **time (*t*)** is the **second (s)**. Mass is not really synonymous with weight even though we express its magnitude in kilograms. The mass of a body is related to its density, but its weight is related to the force of gravity. If two objects are the same size, the one with greater density will weigh more. However, even though an object's mass would be identical on the earth and the moon, it would weigh less on the moon, where there is less gravity.

When we express mass in kilograms and length in meters, we are using the meter-kilogram-second or **MKS** system. Expressing mass in grams and length in centimeters constitutes the centimeter-gram-second or **cgs** system. These two systems also have different derived quantities. For example, the units of force and work are called newtons and joules in the MKS system and dynes and ergs in the cgs system, respectively. We will emphasize the use of MKS units because this is the internationally accepted standard in the scientific community, known as the **Système International d'Unites (SI)**. Equivalent cgs values will often be given as well because the audiology profession has traditionally worked in cgs units, and the death of old habits is slow and labored. These quantities are summarized with equivalent values in MKS and cgs units in **Table 1.1**. In addition, the correspondence between scientific notation and conventional numbers, and the meanings of prefixes used to describe the sizes of metric units are shown for convenience and ready reference in **Table 1.2** and **Table 1.3**.

Quantities may be scalars or vectors. A **scalar** can be fully described by its magnitude (amount or size), but a **vector** has both direction and magnitude. For example, length is a scalar because an object that is one meter long is always one meter long. However, we are dealing with a vector when we measure the distance between two coins that are one meter apart because their relationship has both magnitude and direction (from point x_1 to point x_2). This quantity is called **displacement (*x*)**. Derived quantities will be vectors if they have one or more components that are vectors; for example, velocity is a vector because it is derived from displacement, and acceleration is a vector because it involves velocity. We distinguish between scalars and vectors because they are handled differently when calculations are being made.

Velocity Everyone knows that "55 miles per hour" refers to the speed of a car that causes it to travel a distance of 55 miles in a one-hour period of time. This is an example of **velocity (*v*)**, which is equal to the amount of displacement (x) that occurs over time (t):

$$v = \frac{x}{t}$$

Displacement is measured in meters and time is measured in seconds (s); thus, velocity is expressed

Table 1.1 Principal physical quantities

Quantity	Formula	MKS (SI) units	cgs units	Comments
Mass (*M*)	M	kilogram (kg)	gram (g)	1 kg = 10^3 g
Time (*t*)	t	second (s)	s	
Area (*A*)	A	m^2	cm^2	1 $m^2 = 10^4$ cm^2
Displacement (*x*)	x	meter (m)	centimeter (cm)	1 m = 10^2 cm
Velocity (*v*)	$v = x/t$	m/s	cm/s	1 m/s = 10^2 cm/s
Acceleration (*a*)	$a = v/t$	m/s^2	cm/s^2	1 $m/s^2 = 10^2$ cm/s^2
Force (*F*)	$F = Ma$	kg · m/s^2	g · cm/s^2	1 N = 10^5 dyne
	$= Mv/t$	newton (N)	dyne	
Pressure (*p*)	$p = F/A$	N/m^2 Pascal (Pa)	dyne/cm^2 microbar (μbar)	2×10^{-5} N/m^2 or 20 μPa (reference value) 2×10^{-4} dyne/cm^2 or μbar (reference value)
Work (*W*)	$W = Fx$	N · m joule (J)	dyne · cm erg	1 J = 10^7 erg
Power (*P*)	$P = W/t$	joule/s	erg/s	1 W = 1 J/s
	$= Fx/t$	watt (W)	watt (W)	1 W = 10^7 erg/s
	$= Fv$			
Intensity (I)	$I = P/A$	W/m^2	W/cm^2	10^{-12} W/m^2 (reference value) 10^{-16} W/cm^2 (reference value)

in meters per second (m/s). Velocity is the vector equivalent of speed because it is based on displacement, which has both magnitude and direction. When we take a trip we usually figure out the distance traveled by making a mental note of the starting odometer reading and then subtracting it from the odometer reading at the destination (e.g., if we start at 10,422 miles and arrive at 10,443 miles, then the distance must have been 10,443 – 10,422 = 21 miles). We do the same thing to calculate the time it took to make the trip (e.g., if we left at 1:30 and arrived at 2:10, then the trip must have taken 2:10 – 1:30 = 40 minutes). Physical calculations involve the same straightforward approach. When an object is displaced, it starts at point x_1 and time t_1 and arrives at point x_2 and time t_2. Its *average velocity* is simply the distance traveled ($x_2 - x_1$) divided by the time it took to make the trip ($t_2 - t_1$):

$$v = \frac{x_2 - x_1}{t_2 - t_1}$$

The term *instantaneous* velocity describes the velocity of a body at a *particular moment* in time. For the math-minded, it refers to the velocity when the displacement and time between one point and the next one approach zero, that is, the derivative of displacement with respect to time:

$$v = \frac{dx}{dt}$$

Acceleration Driving experience has taught us all that a car increases its speed to get onto a highway, slows down when exiting, and also slows down while making a turn. “Speeding up” and “slowing down” mean that the velocity is changing over time. The change of velocity over time is **acceleration (*a*)**. Suppose a body is moving between two points. Its velocity at the first point is v_1, and the time at that point is t_1. Similarly, its velocity at the second point is v_2 and the time at that point is t_2. **Average acceleration** is the difference between the two velocities ($v_2 - v_1$) divided by the time interval by the time interval ($t_2 - t_1$):

$$\boldsymbol{a} = \frac{v_2 - v_1}{t_2 - t_1}$$

Table 1.2 Expressing numbers in standard notation and scientific notation

Standard notation	Scientific notation
0.000001	10^{-6}
0.00001	10^{-5}
0.0001	10^{-4}
0.001	10^{-3}
0.01	10^{-2}
0.1	10^{-1}
1	10^{0}
10	10^{1}
100	10^{2}
1000	10^{3}
10,000	10^{4}
100,000	10^{5}
1,000,000	10^{6}
3600	3.6×10^{3}
0.036	3.6×10^{-2}
0.0002	2×10^{-4}
0.00002	2×10^{-5}

In more general terms, acceleration is written simply as

$$a = \frac{v}{t}$$

Because velocity is the same as displacement divided by time, we can replace v with x/t, so that

$$a = \frac{x/t}{t}$$

which can be simplified to

$$a = \frac{x}{t^2}$$

Consequently, acceleration is expressed in units of *meters per second squared (m/s²)* in the MKS system. When measurements are made in cgs units, acceleration is expressed in centimeters per second squared (cm/s²).

Acceleration at a given moment is called **instantaneous acceleration**, and quantitatively oriented readers should note it is equal to the derivative of velocity with respect to time, or

$$a = \frac{dv}{dt}$$

Because velocity is the first derivative of displacement, we find that acceleration is the second derivative of displacement:

$$a = \frac{d^2x}{dt^2}$$

Table 1.3 Examples of prefixes used to express metric units

			Multiply by	
Prefix	**Symbol**	**Definition**	**Standard notation**	**Scientific notation**
micro	μ	millionths	1/1,000,000 *or* 0.000001	$\mathbf{10^{-6}}$
milli	m	thousandths	1/1000 *or* 0.001	$\mathbf{10^{-3}}$
centi	c	hundredths	1/100 *or* 0.01	$\mathbf{10^{-2}}$
deci	d	tenths	1/10 *or* 0.1	$\mathbf{10^{-1}}$
deka	da	tens	10	$\mathbf{10^{1}}$
hecto	h	hundreds	100	$\mathbf{10^{2}}$
kilo	k	thousands	1000	$\mathbf{10^{3}}$
mega	M	millions	1,000,000	$\mathbf{10^{6}}$

Force An object that is sitting still will not move unless some outside influence causes it to do so, and an object that is moving will continue moving at the same speed unless some outside influence does something to change it. This commonsense statement is Newton's first law of motion. It describes the attribute of **inertia**, which is the property of mass to continue doing what it is already doing. The "outside influence" that makes a stationary object move, or causes a moving object to change its speed or direction, is called **force (*F*)**. Notice that force causes the moving object to change velocity or the motionless object to move, which is also a change in velocity (from zero to some amount). Recall that a change of velocity is acceleration. Hence, force is that influence (conceptually a "push" or "pull") that causes a mass to be accelerated. In effect, the amount of "push" or "pull" needed depends on how much mass you want to influence and the amount of acceleration you are trying to produce. In other words, force is equal to the product of mass times acceleration:

$$F = Ma$$

Since acceleration is velocity over time (v/t), we can also specify force in the form

$$F = \frac{Mv}{t}$$

The quantity Mv is called **momentum**, so we may also say that force equals momentum over time.

The amount of force is measured in kg · m/s² because force is equal to the product of mass (measured in kg) and acceleration (measured in m/s²). The *unit of force* is the **newton (N)**, where one newton is the amount of force needed to cause a 1 kg mass to be accelerated by 1 m/s²; hence, 1 N = 1 kg · 1 m/s². (This might seem very technical, but it really simplifies matters; after all, it is easier to say "one newton" than "one kg · m/s².") It would take a 2 N force to cause a 1 kg mass to be accelerated by 2 m/s², or a 2 kg mass to be accelerated by 1 m/s². A 4 N force is needed to accelerate a 2 kg mass by 2 m/s², and a 63 N force is needed to accelerate a 9 kg mass by 7 m/s². In the cgs system, the unit of force is called the **dyne**, which is the force needed to accelerate a 1 g mass by 1 cm/s²; that is, 1 dyne = 1 g · cm/s². It takes 10^5 dynes to equal 1 N.

Many different forces are usually acting upon an object at the same time. Hence, the force we have been referring to so far is actually the **net** or **resultant force**, that is, the "bottom line" effect of all the forces that act upon an object. If a force of 3 N is pushing an object toward the right and a second force of 8 N is also pushing that object toward the right, then the net force would be 3 + 8 = 11 N toward the right. In other words, if two forces push a body in the same direction, then the net force would be the sum of those two forces. Conversely, if a 4 N force pushes an object toward the right at the same time that a 9 N force pushes it toward the left, then the net force is 9 – 4 = 5 N toward the left. Thus, if two forces push an object in opposite directions, then the net force is the difference between the two opposing forces, and it causes the object to accelerate in the direction of the greater force. If two *equal* forces push in *opposite* directions, then the net force is zero. Because the net force is zero it will not cause the motion of the object to change. The situation in which net force is zero is called **equilibrium**. In this case, a moving object will continue moving and an object that is **at rest** (i.e., not moving) will continue to remain still.

Friction When an object is moving in the real world, it tends to slow down and eventually comes to a halt. This happens because anything that is moving in the real world is always in contact with other bodies or mediums. The sliding of one body on the other constitutes a force that opposes the motion, called **resistance** or **friction**.

The opposing force of friction or resistance depends on two parameters. The first factor is that the amount of friction depends on the nature of the materials that are sliding on one another. Simply stated, the amount of friction between two given objects is greater for "rough" materials than for "smooth" or "slick" ones, and is expressed as a quantity called the *coefficient of friction*. The second factor that determines how much friction occurs is easily appreciated by rubbing the palms of your hands back and forth on each other. First rub slowly and then more rapidly. The rubbing will produce heat, which occurs because friction causes some of the mechanical energy to be converted into heat. This notion will be revisited later, but for now we will use the amount of heat as an indicator of the amount of resistance. Your hands become hotter when they are rubbed together more quickly. This illustrates the notion that the amount of friction depends on the *velocity* of motion. In quantitative terms,

$$F_f = Rv$$

where F_f is the force of friction, R is the coefficient of friction between the materials, and v is the velocity of the motion.

Elasticity and restoring force It takes some effort (an outside force) to compress or expand a spring; and the compressed or expanded spring will bounce back to its original shape after it is released. Compressing or expanding the spring is an example of deforming an object. The spring bouncing back to its prior shape is an example of elasticity. More formally, we can say that **elasticity** is the property whereby a deformed object returns to its original form. Notice the distinction between deformation and elasticity. A rubber band and saltwater taffy can both be stretched (deformed), but only the rubber band bounces back. In other words, what makes a rubber band *elastic* is not that it stretches, but rather that it bounces back. The more readily a deformed object returns to its original form, the more elastic (or stiff) it is.

We know from common experiences, such as using simple exercise equipment, that it is relatively easy to *begin* compressing a spring (e.g., a "grip exerciser"), but that it gets progressively harder to *continue* compressing it. Similarly, it is easier to *begin* expanding a spring (e.g., pulling apart the springs on a "chest exerciser") than it is to *continue* expanding it. In other words, the more a spring-like material (an elastic element) is deformed, the more it *opposes* the applied force. The force that opposes the deformation of an elastic or spring-like material is known as the **restoring force**. If we think of deformation in terms of how far the spring has been compressed or expanded from its original position, we could also say that the restoring force increases with displacement. Quantitatively, then, restoring force (F_R) depends on the stiffness (S) of the material and the amount of its displacement as follows:

$$F_R = Sx$$

Pressure Very few people can push a straight pin into a piece of wood, yet almost anyone can push a thumbtack into the same piece of wood. This is possible because a thumbtack is really a simple machine that concentrates the amount of force being exerted over a larger area (the head) down to a very tiny area (the point). In other words, force is affected by the size of the area over which it is applied in a way that constitutes a new quantity. This quantity, which is equal to force divided by area (A), is called **pressure (p)**, so

$$p = \frac{F}{A}$$

Because force is measured in newtons and area is measured in square meters in MKS units, pressure is measured in newtons per square meter (N/m^2).

The unit of pressure is the **pascal (Pa)**, so that 1 Pa = 1 N/m^2. In the cgs system, pressure is measured in dynes per square centimeter ($dynes/cm^2$), occasionally referred to as **microbars (μbars)**.

Work and energy As a physical concept, **work (W)** occurs when the force applied to a body results in its displacement, and the amount of work is equal to the product of the force and the displacement, or

$$W = Fx$$

Because force is measured in newtons and displacement is measured in meters, work itself is quantified in newton-meters (N · m). For example, if a force of 2 N displaces a body by 3 m, then the amount of work is $2 \times 3 = 6$ N. There can only be work if there is displacement. There cannot be work if there is no displacement (i.e., if $x = 0$) because work is the product of force and displacement, and zero times anything is zero. The MKS unit of work is the **joule (J)**. One joule is the amount of work that occurs when one newton of force effects one meter of displacement, or 1 J = 1 N · m. In the cgs system, the unit of work is called the erg, where 1 erg = 1 dyne · cm. One joule corresponds to 10^7 erg.

Energy is usually defined as the capability to do work. The energy of a body at rest is **potential** energy and the energy of an object that is in motion is **kinetic** energy. The total energy of a body is the sum of its potential energy plus its kinetic energy, and work corresponds to the exchange between these two forms of energy. In other words, energy is not consumed when work is accomplished; it is converted from one form to the other. This principle is illustrated by the simple example of a swinging pendulum. The pendulum's potential energy is greatest when it reaches the extreme of its swing, where its motion is momentarily zero. On the other hand, the pendulum's kinetic energy is greatest when it passes through the midpoint of its swing because this is where it is moving the fastest. Between these two extremes, energy is being converted from potential to kinetic as the pendulum speeds up (on each down swing), and from kinetic to potential as the pendulum slows down (on each up swing).

Power The rate at which work is done is called **power (P)**, so that power can be defined as work divided by time,

$$P = \frac{W}{t}$$

The unit of power is called the **watt (W)**. One unit of power corresponds to one unit of work divided by one unit of time. Hence, one watt is equal to one joule divided by one second, or 1 W = 1 J/s. Power is also expressed in watts in the cgs system, where work is measured in ergs. Since 1 J = 10^7 erg, we can also say that 1 W = 10^7 erg/s.

Power can also be expressed in other terms. For example, because $W = Fx$, we can substitute Fx for W in the power formula, to arrive at

$$P = \frac{Fx}{t}$$

We know that $v = x/t$, so we can substitute v for x/t and rewrite this formula as

$$P = Fv$$

In other words, power is also equal to force times velocity.

Intensity Consider a hypothetical demonstration in which one tablespoonful of oil is placed on the surface of a still pond. At that instant the entire amount of oil will occupy the space of a tablespoon. As time passes, the oil spreads out over an expanding area on the surface of the pond, and it therefore also thins out so that much less than all the oil will occupy the space of a tablespoon. The wider the oil spreads the more it thins out, and the proportion of the oil covering any given area gets smaller and smaller, even though the total amount of oil is the same. Clearly, there is a difference between the amount of oil, per se, and the concentration of the oil as it is distributed across (i.e., divided by) the surface area of the pond.

An analogous phenomenon occurs with sound. It is common knowledge that sound radiates outward in every direction from its source, constituting a sphere that gets bigger and bigger with increasing distance from the source, as illustrated by the concentric circles in **Fig. 1.1**. Let us imagine that the sound source is a tiny pulsating object (at the center of the concentric circles in the figure) and that it produces a finite amount of power, analogous to the fixed amount of oil in the prior example. Consequently, the sound power will be divided over the ever-expanding surface as distance increases from the source, analogous to the thinning out of the widening oil slick. This notion is represented in the figure by the thinning of the lines at greater distances from the source. Suppose we measure how much power registers on a certain fixed amount of surface area (e.g., a square inch). As a result, a progressively smaller proportion of the original power falls onto a square inch as the distance from the source increases, represented in the figure by the lighter shading of the same-size ovals at increasing distances from the source.

The examples just described reveal that a new quantity, called **intensity (*I*)**, develops when power is distributed over area. Specifically, intensity is equal to power per unit area, or power divided by area, or

$$I = \frac{P}{A}$$

Because power is measured in watts and area is measured in square meters in the MKS system, intensity is expressed in watts per square meter (W/m^2). Intensity is expressed in watts per square centimeter (W/cm^2) in the cgs system.

Intensity decreases with increasing distance from a sound source according to a rule called the **inverse square law**. It states that the amount of

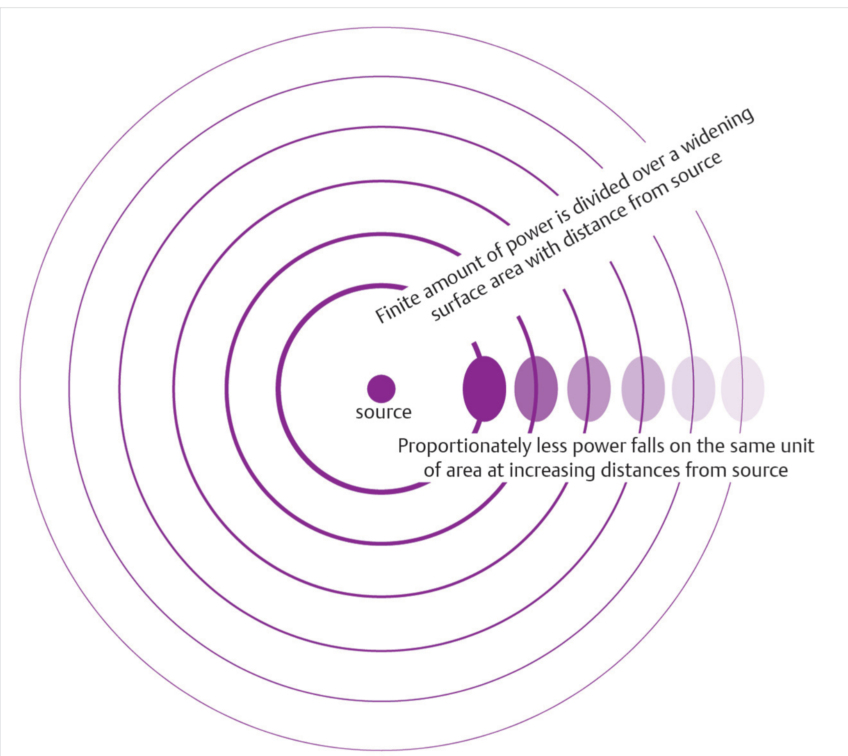


Fig. 1.1 Intensity (power divided by area) decreases with distance from the sound source because a fixed amount of power is spread over an increasing area, represented by the thinning of the lines. Proportionately less power falls on the same unit area (represented by the lighter shading of the ovals) with increasing distance from the source.

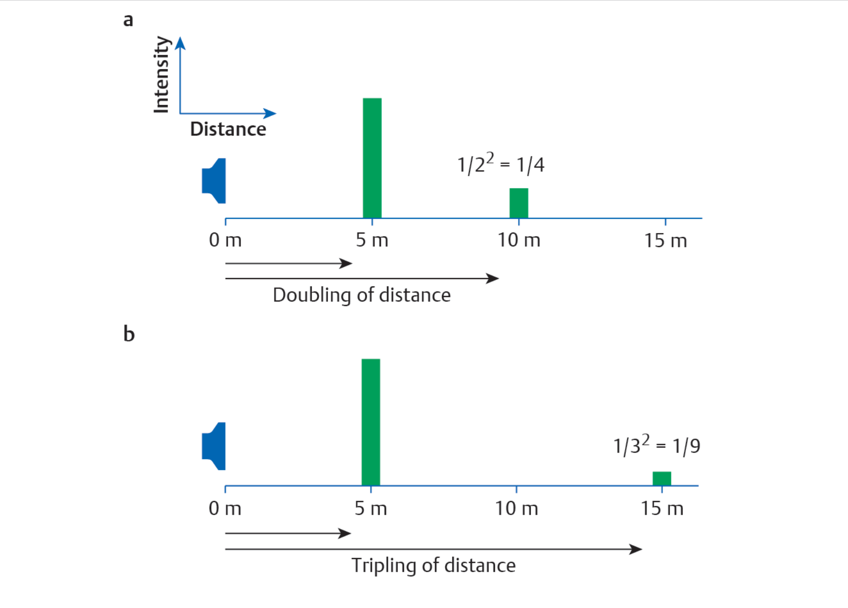


Fig. 1.2 Illustrations of the inverse square law. **(a)** Doubling of distance: The intensity at 10 m away from a loudspeaker is one quarter of its intensity at 5 m because $1/2^2 = 1/4$. **(b)** Tripling of distance: The intensity at 15 m away from the sources is one-ninth of its intensity at 5 m because $1/3^2 = 1/9$.

intensity drops by 1 over the square of the change in distance. Two examples are illustrated in **Fig. 1.2**. **Fig. 1.2a** shows that when the distance from a loudspeaker is doubled from 5 m to 10 m, the amount of intensity at 10 m will be one quarter of what it was at 5 m (because $1/2^2 = 1/4$). Similarly, **Fig. 1.2b** shows that tripling the distance from 5 m to 15 m causes the intensity to fall to one-ninth of its value at the closer point because $1/3^2 = 1/9$.

An important relationship to be aware of is that power is equal to pressure squared,

$$P = p^2$$

and pressure is equal to the square root of power,

$$p = \sqrt{P}$$

In addition, intensity is proportional to pressure squared,

$$I \propto p^2$$

and pressure is proportional to the square root of intensity,

$$p \propto \sqrt{I}$$

This simple relationship makes it easy to convert between sound intensity and sound pressure.

The Nature of Sound

Sound is often defined as a form of vibration that propagates through the air in the form of a wave. **Vibration** is nothing more than the to-and-fro motion (oscillation) of an object. Some examples include a playground swing, a pendulum, the floorboards under a washing machine, a guitar string, a tuning fork prong, and air molecules. The vibration is usually called **sound** when it is transferred from air particle to air particle (we will see how this happens later). The vibration of air particles might have a simple pattern such as the tone produced by a tuning fork, or a very complex pattern such as the din heard in a school cafeteria. Most naturally occurring sounds are very complex, but the easiest way to understand sound is to concentrate on the simplest ones.

Simple Harmonic Motion

A vibrating tuning fork[1] is illustrated in **Fig. 1.3**. The initial force that was applied by striking the tuning fork is represented by the green arrow in frame 1. The progression of the drawings represents the motion of the prongs at selected points in time after the fork has been activated. The two prongs vibrate as mirror images of each other, so that we can describe what is happening in terms of just one prong. The *circular insert* highlights the motion of the right prong. Here the center position is where the prong would be at rest. When the fork is struck the prong is forced inward as shown by arrow **a**. After reaching the leftmost position it bounces back (arrow **b**), accelerating along the way. The rapidly moving prong overshoots the center and continues rightward (arrow **c**). It slows down along the way until it stops for an instant at the extreme right, where it reverses direction again and starts moving toward the left (arrow **d**) at an ever-increasing speed. It overshoots

[1] See Russell (2020) for detailed review of the characteristics and history of tuning forks.

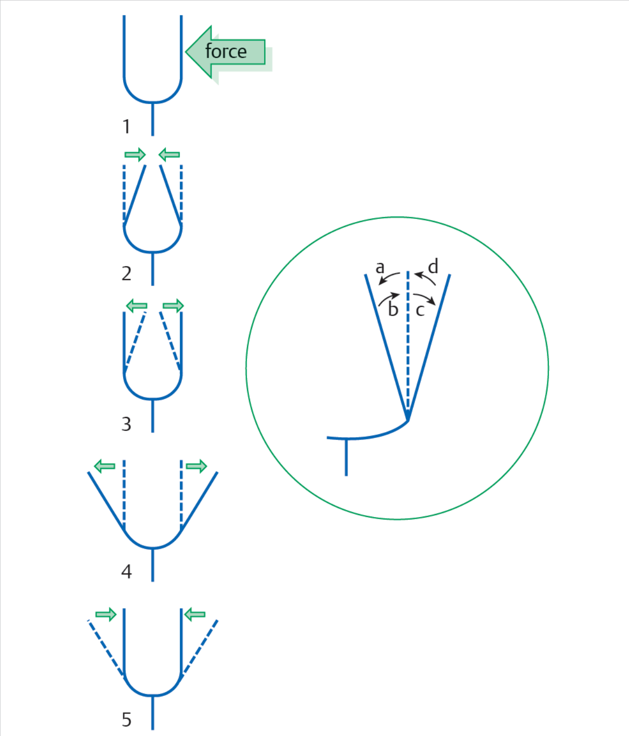


Fig. 1.3 After being struck, a tuning fork vibrates or oscillates with a simple pattern that repeats itself over time. One replication (cycle) of this motion is illustrated going from frames 1 to 5. The arrows in the insert highlight the motion of one of the prongs.

the center again, and as before, the prong now follows arrow **a**, slowing down until it stops momentarily at the extreme left. Here it reverses direction again and repeats the same process over and over again. One complete round trip (or replication) of an oscillating motion is called a **cycle**. The number of cycles that occur in one second is called **frequency**.

This form of motion occurs when a force is applied to an object having the properties of inertia and elasticity. Due to its elasticity, the deformation of the fork caused by the applied force is opposed by a restoring force. In the figure the initial leftward force is opposed by a restoring force in the opposite direction, that is, toward the right. The rightward restoring force increases as the prong is pushed progressively toward the left. As a result, the movement of the prong slows down and eventually stops. Under the influence of its elasticity the prong now reverses direction and starts moving rightward. As the restoring force brings the prong back toward the center, we must also consider its mass. Because the prong has mass, inertia causes it to accelerate as it moves back toward its center resting position. In fact, the prong is moving at its maximum speed as it passes through the resting position. The force of inertia causes the prong to overshoot the center and continue moving rightward. The deformation process begins again once the prong overshoots its resting position. As a result, opposing elastic restoring forces start building up again, now in the leftward direction. Just as before, the increasing (leftward) restoring force eventually overcomes the rightward inertial force, thereby stopping the prong's displacement at the rightmost point, and causing a reversal in the direction of its movement. Hence, the same course of events is repeated again, this time in the leftward direction; then rightward, then leftward, etc., over and over again. This kind of vibration is called **simple harmonic motion** because the oscillations repeat themselves at the same rate over and over again.

We know from experience that the oscillations just described do not continue forever. Instead, they dissipate over time and eventually die out completely. The dying out of vibrations over time is called **damping**, and it occurs due to resistance or friction. Resistance occurs because the vibrating prong is always in contact with the surrounding air. As a result, there will be friction between the oscillating metal and the surrounding air molecules. This friction causes some of the mechanical energy that has been supporting the motion of the tuning fork to be converted into heat. In turn, the energy that has been converted into heat is no longer available to maintain the vibration of the tuning fork. Consequently, the sizes of the oscillations dissipate and eventually die out altogether.

A diagram summarizing the concepts just described is shown in Fig. 1.4. The curve in the figure represents the tuning fork's motion. The amount of displacement of the tuning fork prong around its resting (or center) position is represented by distance above and below the horizontal line. These events are occurring over time, which is represented by horizontal distance (from left to right). The initial displacement of the prong due to the original applied force is represented by the dotted segment of the curve. Inertial forces due to the prong's mass and elastic restoring forces due to the elasticity of the prong are represented by labeled arrows. Damping of the oscillations due to friction is shown by the decline in the displacement of the curve as time goes on. The curve in this diagram is an example of a **waveform**, which is a graph that shows displacement (or another measure of magnitude) as a function of time.

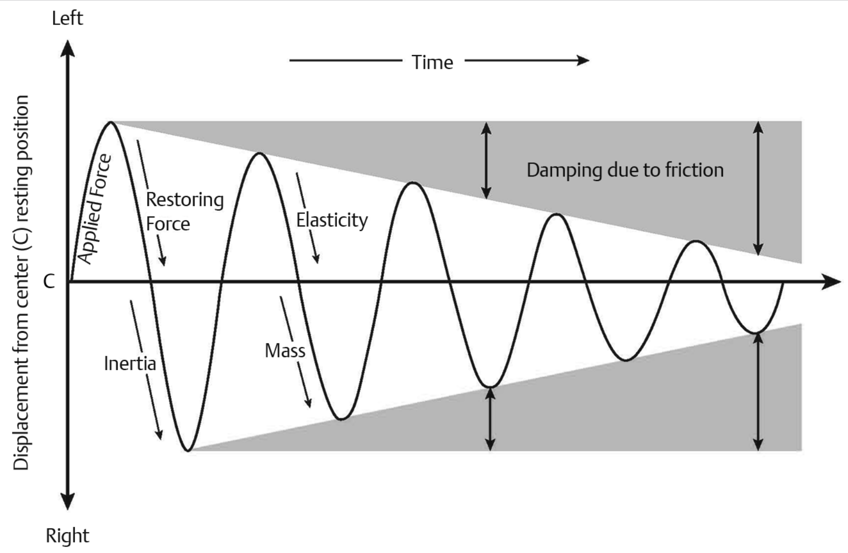


Fig. 1.4 Diagrammatic representation of tuning fork oscillations over time. Vertical displacement represents the amount of the tuning fork prong displacement around its resting position. Distance from left to right represents the progression of time. (Adapted from Gelfand 2018, courtesy of CRC Press.)

Sound Waves

Tuning fork vibrations produce sound because the oscillations of the prongs are transmitted to the surrounding air particles. When the tuning fork prong moves to the right, it displaces air molecules to its right in the same direction. These molecules are thus displaced to the right of their own resting positions. Displacing air molecules toward the right pushes them closer to the air particles to their right. The pressure that exists among air molecules that are not being disturbed by a driving force (like the tuning fork) is known as *ambient or atmospheric pressure*. We can say that the rightward motion of the tuning fork prong exerts a force on the air molecules that pushes them together relative to their undisturbed, resting situation. In other words, forcing the air molecules together causes an increase in air pressure relative to the ambient pressure that existed among the undisturbed molecules. This state of positive air pressure is called **compression**. The amount of compression increases as the prong continues displacing the air molecules rightward. A maximum amount of positive pressure occurs when the prong and air molecules reach their greatest rightward displacement.

The tuning fork prong then reverses direction, overshoots its resting position, and proceeds to its leftmost position. The compressed air molecules reverse direction along with the prong. This occurs because air is an elastic medium, so the particles compressed to the right develop a leftward restoring force. Small as they are, air particles do have mass. Therefore, inertia causes the rebounding air particles to overshoot their resting positions and to continue toward their extreme leftward positions. As the particles move leftward, the amount of compression decreases and is momentarily zero as they pass through their resting positions. As they continue to move to the left of their resting positions, the particles are now becoming increasingly farther from the molecules to their right (compared with when they are in their resting positions). We now say that the air particles are rarefied compared with their resting states, so that the air pressure is now below atmospheric pressure. This state of lower than ambient pressure is called **rarefaction**. When the air particles reach the leftmost position they are maximally rarefied, which means that the pressure is maximally negative. At this point, the restoring force instigates a rightward movement of the air molecules. This movement is enhanced by the push of the tuning fork prongs that have also reversed direction. The air molecules now accelerate rightward (so that the amount of rarefaction decreases), overshoot their resting positions, and continue to the right, and so on. The tuning fork vibrations have now been transmitted to the surrounding particles, which are now also oscillating in simple harmonic motion. Sounds that are associated with simple harmonic motion are called **pure tones**.

Let us consider one of the air molecules that has already been set into harmonic motion by the tuning fork. This air particle now vibrates to-and-fro in the same direction that was originally imposed by the vibrating prong. When this particle moves toward its right it will cause a similar displacement of the particle that is located there. The subsequent leftward motion is also transmitted to the next particle, etc. Thus, the oscillations of one air particle are transmitted to the molecule next to it. The second particle is therefore set into oscillation, which in turn initiates oscillation of the next one, and so forth down the line. In other words, each particle vibrates back and forth around its own resting point, and causes successive molecules to vibrate back and forth around their own resting points, as shown schematically in **Fig. 1.5**. Notice that each molecule vibrates "in place" around its own average position; it is the vibratory pattern that is transmitted through the air.

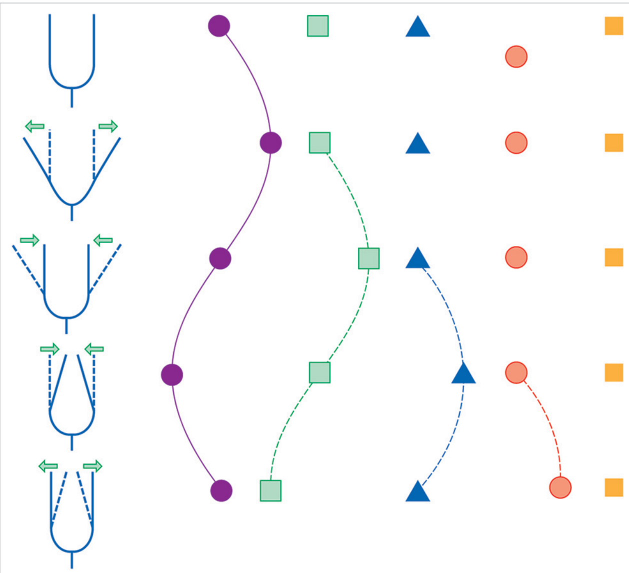


Fig. 1.5 Sound is initiated by transmitting the vibratory pattern of the sound source to nearby air particles, and then the vibratory pattern is passed from particle to particle as a wave. Notice how it is the pattern of vibration that is being transmitted, whereas each particle oscillates around its own average location.

This propagation of vibratory motion from particle to particle constitutes the sound wave. This wave appears as alternating compressions and rarefactions radiating from the sound source in all directions, as already suggested in **Fig. 1.1**. The transmission of particle motion along with the resulting variations in air pressure with distance from the source are represented in **Fig. 1.6**. Most people are more familiar with the kinds of waves that develop on the surface of a pond when a pebble is dropped into the water. These are **transverse waves** because the particles are moving at right angles to the direction that the wave is propagating. That is, the water particles oscillate up and down (vertically) even though the wave moves out horizontally from the spot where the pebble hit the water. This principle can be demonstrated by floating a cork in a pool, and then dropping a pebble in the water to start a wave. The floating cork reflects the motions of the water particles. The wave will move out horizontally, but the floating cork bobs up and down (vertically) at right angles to the wave. In contrast, sound waves are **longitudinal waves** because each air particle oscillates in the same direction in which the wave is propagating (**Fig. 1.6**). Although sound waves are longitudinal, it is more convenient to draw them with a transverse representation, as in **Fig. 1.6**. In such a diagram, the vertical dimension represents some measure of the size of the signal (e.g., displacement, pressure, etc.), and left to right distance represents time (or distance). For example, the waveform in **Fig. 1.6** shows the amount of positive pressure (compression) above the baseline, negative pressure (rarefaction) below the baseline, and distance horizontally going from left to right.

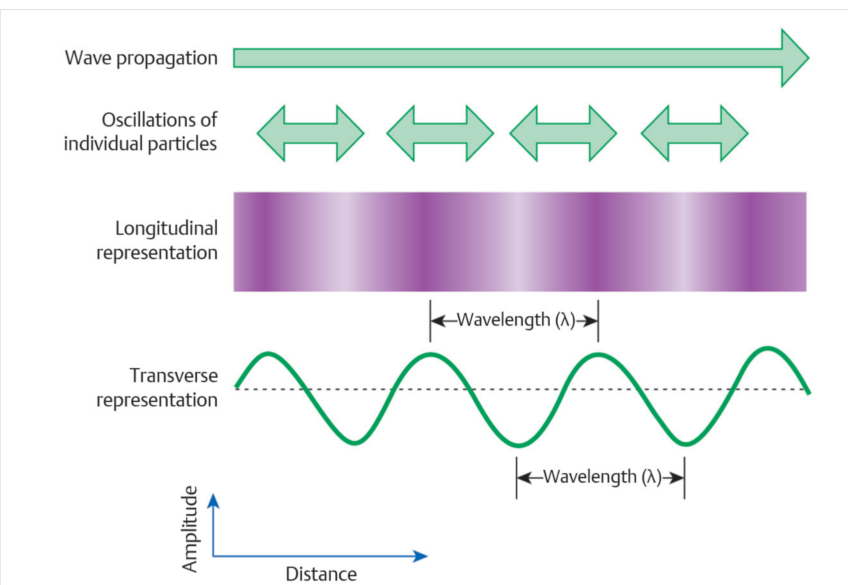


Fig. 1.6 Longitudinal and transverse representations of a sound wave. Wavelength (λ) is the distance covered by one replication (cycle) of a wave, and is most easily visualized as the distance from one peak to the next.

The Sinusoidal Function

Simple harmonic motion is also known as **sinusoidal motion**, and has a waveform that is called a **sinusoidal wave** or a **sine wave**. Let us see why. Fig. 1.7 shows one cycle of a sine wave in the center, surrounded by circles labeled to correspond to points on the wave. Each circle shows a horizontal line corresponding to the horizontal baseline on the sine wave, as well as a radius line (r) that will move around the circle at a fixed speed, much like a clock hand but in a counterclockwise direction.

Point **a** on the waveform in the center of the figure can be viewed as the "starting point" of the cycle. The displacement here is zero because this point is *on* the horizontal line. The radius appears as shown in circle **b** when it reaches 45° of rotation, which corresponds to point **b** on the sine wave. The angle between the radius and the horizontal is called the **phase angle** (θ) and is a handy way to tell location going around the circle and on the sine wave. In other words, we consider one cycle (one "round trip" of oscillation) to be the same as going around a circle one time. Just as a circle has 360°, we also consider one cycle to have 360°. Since $45/360 = {}^{1}/_{8}$, a phase angle (θ) of 45° is the same as one-eighth of the way around a circle or one-eighth of the way into a sine wave. Returning to the circle, the vertical displacement from the horizontal to the point where r intersects the circle is represented by a vertical line labeled d. This vertical line corresponds to the displacement of point **b** on the sine wave, where the displacement of the air particle is represented by the height of the point above the baseline. Notice that we now have a right triangle in the circle, where r is the hypotenuse, θ is an angle, and d is the leg opposite that angle. Recall from high school math that the sine of an angle equals the length of the opposite leg over the length of the hypotenuse. Here, $\sin\theta = d/r$. If we conveniently assume that the length of r is 1, then displacement d becomes the sine of angle θ, which happens to be 0.707. In other words, displacement is determined by the sine of the phase angle, and displacement at any point on the sine wave corresponds to the sine of θ. This is why it is called a sine wave.

The peak labeled **c** on the sine wave corresponds to circle **c**, where the rotating radius has reached the straight-up position. We are now one-fourth of the way into the wave and one-fourth of the way around the circle. Here, $\theta = 90°$ and the displacement is the largest it can be (notice that $d = r$ on the circle). Continuing the counterclockwise rotation of r causes the amount of displacement from the horizontal to decrease, exemplified by point **d** on the sine wave and circle **d**, where θ is 135°. The oscillating air particle has already reversed direction and is now moving back toward the resting position. When it reaches the resting position there is again no displacement, as shown by point **e** on the sine wave and by the fact that r is now superimposed on the horizontal at $\theta = 180°$ in circle **e**. Notice that 180° is one half of the 360° round trip, so we are now halfway around the circle and halfway into the cycle of simple harmonic motion. In addition, displacement is zero at this location (180°).

Continuing the rotation of r places it in the lower left quadrant of circle **f**, corresponding to point **f** on

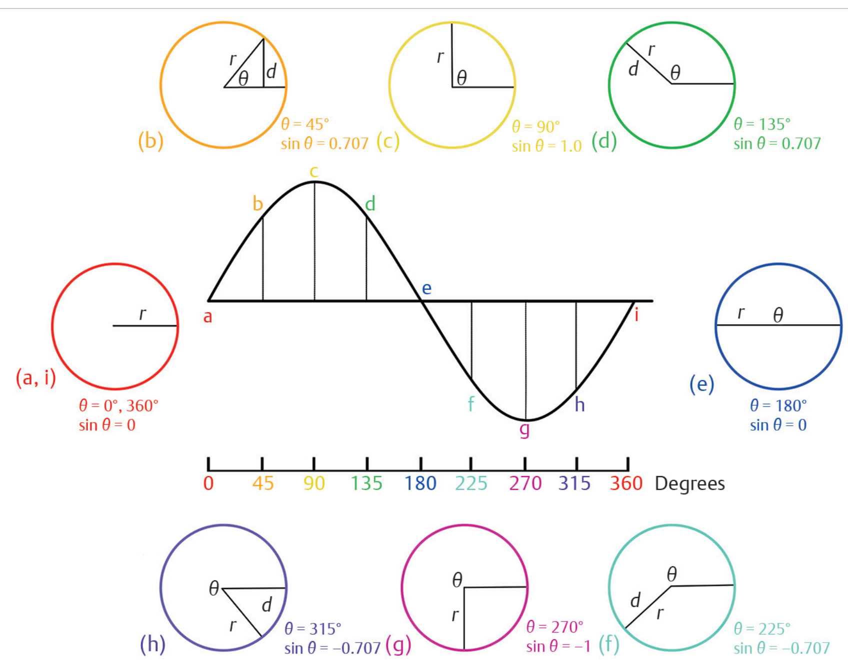


Fig. 1.7 Sinusoidal motion (θ, phase angle; d, displacement). (Adapted from Gelfand 2018, courtesy of CRC Press.)

the wave, where $\theta = 225°$. The oscillating particle has overshot its resting position and the displacement is now increasing in the other direction, so that we are in the rarefaction part of the wave. Hence, displacement is now drawn in the negative direction, indicating rarefaction. The largest negative displacement is reached at point **g** on the wave, where $\theta = 270°$, corresponding to circle **g**, in which *r* points straight down.

The air particle now begins moving in the positive direction again on its way back toward the resting position. At point **h** and circle **h**, the displacement in the negative direction has become smaller as the rotating radius passes through the point where 0 = 315° (point **h** on the wave and circle h). The air particle is again passing through its resting position at point **i**, having completed one round trip or 360° of rotation. Here, displacement is again zero. Having completed exactly one cycle, 360° corresponds to 0°, and circle **i** is the same one previously used as circle **a**.

Recall that *r* rotates around the circle at a fixed speed. Hence, how fast *r* is moving will determine how many degrees are covered in a given amount of time. For example, if one complete cycle of rotation takes 1 second, then 360° is covered in 1 second; 180° is covered in $^1/_2$ second; 90° takes $^1/_4$ second; 270° takes $^3/_4$ second, etc. Hence, the phase angle also reflects the elapsed time from the onset of rotation. This is why the horizontal axis in **Fig. 1.8** can be labeled in terms of phase. As such, the phase of the wave at each of the points indicated in **Fig. 1.7** is 0° at **a**, 45° at **b**, 90° at **c**, 135° at **d**, 180° at **e**, 225° at **f**, 270° at **g**, 315° at **h**, and 360° at **i**, which is also 0° for the next cycle.

Phase is often used to express relationships between two waves that are displaced relative to each other, as in **Fig. 1.8**. Each frame in the figure shows two waves that are identical to each other except that they do not line up exactly along the horizontal (time) axis. The top panel shows two waves that are 45° apart. Here, the wave represented by the thicker line is at 45° at the same time that the other wave (shown by the thinner line) is at 0°. The phase displacement is highlighted by the shaded area and the dotted vertical guideline in the figure. This is analogous to two radii that are always 45° apart as they move around a circle. In other words, these two waves are 45° apart or *out-of-phase*. The second panel shows the two waves displaced from each another by 90°, so that one wave is at 90° when other one is at 0°. Hence, these waves are 90° out-of-phase, analogous to two radii that are always 90° apart as they move around a circle. The third panel shows two waves that are 180° out-of-phase. Here, one wave is at its 90° (positive) peak

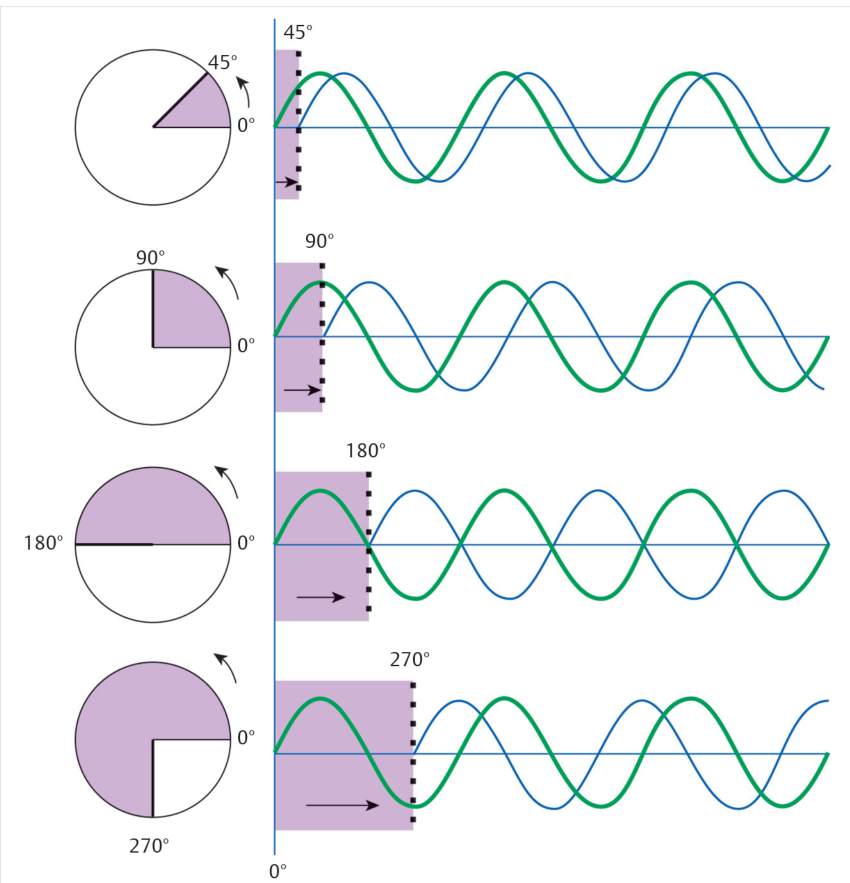


Fig. 1.8 Each panel shows two waves that are identical in every way except they are displaced from one another in terms of phase, highlighted by the shaded areas and the dotted vertical guidelines. Analogous examples of two radii moving around a circle are shown to the left of the waveforms. Top panel: Two waves that are 45° out-of-phase, analogous to two radii that are always 45° apart as they move around a circle. Second panel: Waves that are 90° out-of-phase, analogous to two radii moving around a circle 90° apart. Third panel: Waves that are 180° out-of phase, analogous to two radii that are always 180° apart (pointing in opposite directions) moving around a circle. Bottom panel: Two waves (and analogous radii moving around a circle) that are 270° out-of-phase.

at the same time that the other one is at its 270° (negative) peak, which is analogous to two radii that are always 180° apart as they move around a circle. Notice that these two otherwise identical waves are exact mirror images of each other when they are 180° out-of-phase, just as the two radii are always pointing in opposite directions. The last example in the bottom panel shows the two waves 270° out-of-phase.

Parameters of Sound Waves

We already know that a **cycle** is one complete replication of a vibratory pattern. For example, two cycles are shown for each sine wave in the upper frame of **Fig. 1.9**, and four cycles are shown for each sine wave in the lower frame. Each of the sine waves in this figure is said to be **periodic** because it repeats itself exactly over time. Sine waves are the simplest

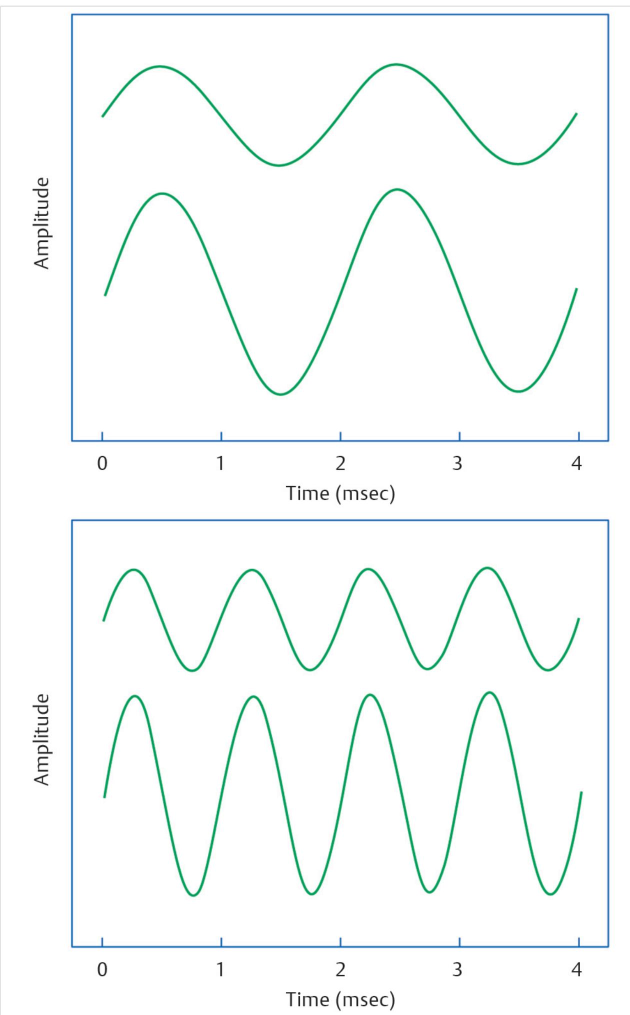


Fig. 1.9 Each frame shows two sine waves that have the same frequency but different amplitudes. Compared with the upper frame, twice as many cycles occur in the same amount of time in the lower frame; thus, the period is half as long and the frequency is twice as high.

kind of periodic wave because simple harmonic motion is the simplest form of vibration. Later we will address complex periodic waves.

The duration of one cycle is called its **period**. The period is expressed in time (t) because it refers to the amount of time that it takes to complete one cycle (i.e., how long it takes for one round trip). For example, a periodic wave that repeats itself every one-hundredth of a second has a period of 1/100 seconds, or $t = 0.01$ seconds. One-hundredth of a second is also 10 thousandths of a second (milliseconds), so we could also say that the period of this wave is 10 milliseconds.

Similarly, a wave that repeats itself every one-thousandth of a second has a period of 1 millisecond or 0.001 seconds; and the period would be 2 milliseconds or 0.002 seconds if the duration of one cycle is two-thousandths of a second.

The number of times a waveform repeats itself in one second is its frequency (f), or the number of **cycles per second (cps)**. We could say that frequency is the number of cycles that can fit into one second. Frequency is expressed in units called **hertz (Hz)**, which means the same thing as cycles per second. For example, a wave that is repeated 100 times per second has a frequency of 100 Hz; the frequency of a wave that has 500 cycles per second is 500 Hz; and a 1000 Hz wave has 1000 cycles in one second.

If frequency is the number of cycles that occur each second, and period is how much time it takes to complete one cycle, then frequency and period must be related in a very straightforward way. Let us consider the three examples that were just used to illustrate the relationship of period and frequency:

- A period of 0.01 seconds corresponds to a frequency (f) of 100 Hz.
- A period of 0.002 seconds corresponds to a frequency of 500 Hz.
- A period of 0.001 seconds corresponds to a frequency of 1000 Hz.

Now, notice the following relationships among these numbers:

- $1/100 = 0.01$ and $1/0.01 = 100$.
- $1/500 = 0.002$ and $1/0.002 = 500$.
- $1/1000 = 0.001$ and $1/0.001 = 1000$.

In each case, the period corresponds to 1 over the frequency, and the frequency corresponds to 1 over the period. In formal terms, frequency equals the reciprocal of period, and period equals the reciprocal of frequency,

$$f = \frac{1}{t}$$

Each wave in the upper frame of **Fig. 1.9** contains two cycles in 4 milliseconds, and each wave in the lower frame contains four cycles in the 4 milliseconds. If two cycles in the upper frame last 4 milliseconds, then the duration of one cycle is 2 milliseconds. Hence, the period of each wave in the upper frame is 2 milliseconds ($t = 0.002$ s), and the frequency is 1/0.002, or 500 Hz. Similarly, if four cycles last 4 milliseconds in the lower frame, then one cycle has a period of 1 millisecond ($t = 0.001$ **s**), and the frequency is 1/0.001, or 1000 Hz.

Fig. 1.9 also illustrates differences in the amplitude between waves. **Amplitude** denotes size or magnitude, such as the amount of displacement, power, pressure, etc. The larger the amplitude at some point along the horizontal (time) axis, the greater its distance from zero on the vertical axis. With respect to the figure, each frame shows one wave that has a smaller amplitude and an otherwise identical wave that has a larger amplitude.

As illustrated in **Fig. 1.10**, the **peak-to-peak amplitude** of a wave is the total vertical distance between its negative and positive peaks, and **peak amplitude** is the distance from the baseline to one peak. However, neither of these values reflects the overall, ongoing size of the wave because the amplitude is constantly changing. At any instant an oscillating particle may be at its most positive or most negative displacement from the resting position, or anywhere between these two extremes, including the resting position itself, where the displacement is zero. The magnitude of a sound at a given instant **(instantaneous amplitude)** is applicable only for that moment, and will be different at the next moment. Yet we are usually interested in a kind of "overall average" amplitude that reveals the magnitude of a sound wave throughout its cycles. A simple average of the positive and negative instantaneous amplitudes will not work because it will always be equal to zero. A different kind of overall measure is therefore used, called the **root-mean-square (RMS)** amplitude. Even though measurement instruments provide us with RMS amplitudes automatically, we can understand RMS by briefly reviewing the steps that would be used to calculate it manually: (1) All of the positive and negative values on the wave are squared, so that all values are positive (or zero for values on the resting position itself). (2) A mean (average) is calculated for the squared values. (3) This average of the squared values is then rescaled back to the "right size" by taking its square root. This is the RMS value. The RMS amplitude is numerically equal to 70.7% of (or 0.707 times) the peak amplitude (or 0.354 times the peak-to-peak amplitude). Even though these values technically apply only to sinusoids, for practical purposes RMS values are used with all kinds of waveforms.

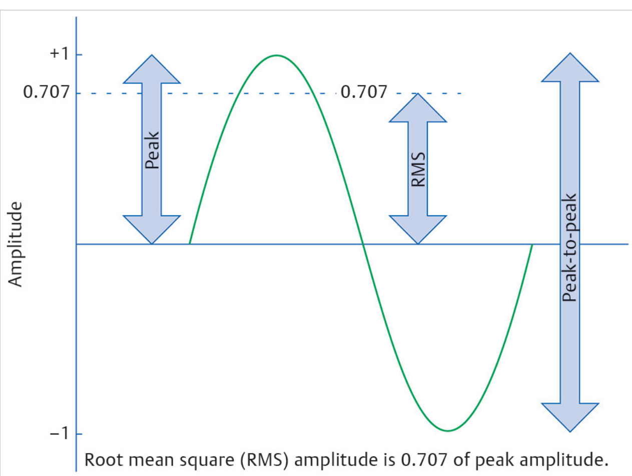


Fig. 1.10 Peak, root-mean-square (RMS), and peak-to-peak amplitude.

Referring back to **Fig. 1.6**, we see that the distance covered by one cycle of a propagating wave is called its **wavelength** (λ). We have all seen water waves, which literally appear as alternating crests and troughs on the surface of the water. Using this common experience as an example, wavelength is simply the distance between the crest of one wave and the crest of the next one. For sound, wavelength is the distance between one compression peak and the next one, or one rarefaction peak and the next one, that is, the distance between any two successive peaks in **Fig. 1.6**. It is just as correct to use any other point, as long as we measure the distance between the same point on two successive replications of the wave. The formula for wavelength is

$$\lambda = \frac{c}{f}$$

where f is the frequency of the sound and c is the speed of sound (~ 344 m/s in air). This formula indicates that wavelength is inversely proportional to frequency. Similarly, frequency is inversely proportional to wavelength:

$$f = \frac{c}{\lambda}$$

These formulas show that wavelength and frequency are inversely proportional to each other. In other words, low frequencies have long wavelengths and high frequencies have short wavelengths.

Complex Waves

When two or more pure tones are combined, the result is called a complex wave. Complex waves may contain any number of frequencies from as few as two up to an infinite number of them. **Complex**

periodic waves have waveforms that repeat themselves over time. If the waveform does not repeat itself over time, then it is an **aperiodic wave**.

Combining Sinusoids

The manner in which waveforms combine into more complex waveforms involves algebraically adding the amplitudes of the two waves at every point along the horizontal (time) axis. Consider two sine waves that are to be added. Imagine that they are drawn one above the other on a piece of graph paper so that the gridlines can be used to identify similar moments in time (horizontally) for the two waves, and their amplitudes can be determined by simply counting boxes vertically. The following exercise is done at every point along the horizontal time axis: (1) Determine the amplitude of each wave at that point by counting the boxes in the positive and/or negative direction. (2) Add these two amplitudes *algebraically* (e.g., + 2 plus + 2 is + 4; –3 plus –3 is –6; and + 4 plus –1 is + 3, etc.). (3) Plot the algebraic sum just obtained on the graph paper at the same point along the horizontal time axis. After doing this for many points, drawing a smooth line through these points will reveal the combined wave.

Several examples of combining two sinusoids are illustrated in **Fig. 1.11**. This figure shows what occurs when two sinusoids being combined have exactly the same frequencies and amplitudes. The two sinusoids being combined in **Fig. 1.11a** are in phase with each other. Here, the combined wave looks like the two identical components, but has an amplitude twice as large. This case is often called complete **reinforcement** for obvious reasons. The addition of two otherwise identical waves that are 180° out-of-phase is illustrated in **Fig. 1.11b**. In this case, the first wave is equal and opposite to the second wave at every moment in time, so that algebraic addition causes the resulting amplitude to be zero at every point along the horizontal (time) axis. The result is complete **cancellation**.

When the sinusoids being combined are identical but have a phase relationship that is any value other than 0° (in-phase) or 180° (opposite phase), then the appearance of the resulting wave will depend on how the two components happen to line up in time. **Fig. 1.11c** shows what happens when the two otherwise identical sinusoids are 90° out-of-phase. The result is a sinusoid with the same frequency as the two (similar) original waves but that differs in phase and amplitude.

Complex Periodic Waves

The principles used to combine any number of similar or dissimilar waves are basically the same as

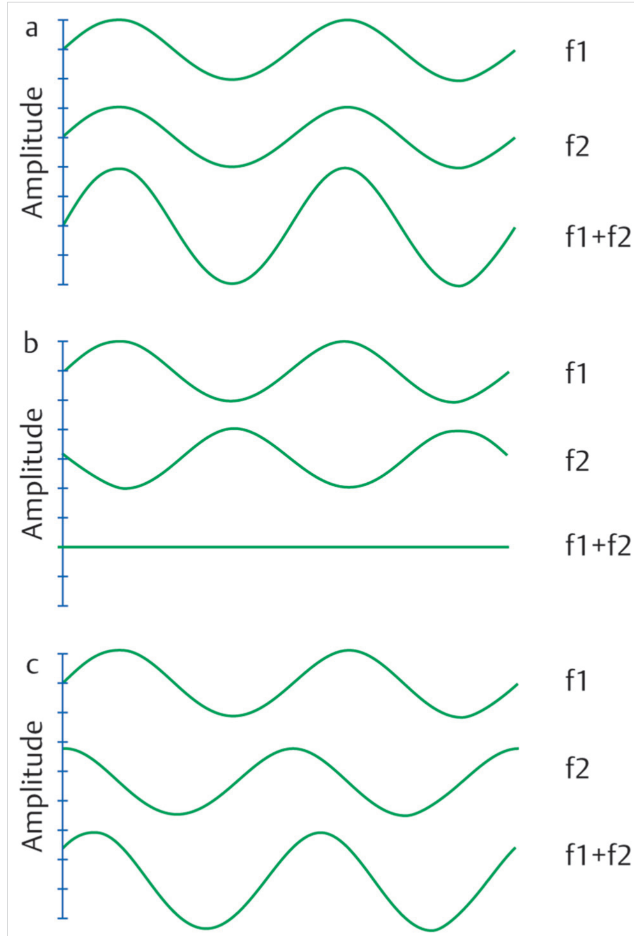


Fig. 1.11 Combining sinusoids that have the same frequency and amplitude when they are **(a)** in-phase (showing complete reinforcement); **(b)** 180° out-of-phase (showing cancellation); and **(c)** 90° out-of-phase.

those just described for two similar waves: Their amplitudes are algebraically summed on a point-by-point basis along the horizontal (time) axis, regardless of their individual frequencies and amplitudes or their phase relationships. However, combining unequal frequencies will not produce a sinusoidal result. Instead, the combined waveform depends on the nature of the particular sounds being combined. For example, consider the three different sine waves labeled *f*1, *f*2, and *f*3 in **Fig. 1.12**. Wave *f*1 has a frequency of 1000 Hz, *f*2 is 2000 Hz, and *f*3 is 3000 Hz. The lower waveforms show various combinations of these sinusoids. The combined waves (*f*1 + *f*2, *f*1 + *f*3, and *f*1 + *f*2 + *f*3) are no longer sinusoids, but they are periodic because they repeat themselves at regular intervals over time. In other words, they are all complex periodic waves.

Notice that the periods of the complex periodic waves in **Fig. 1.12** are the same as the period of *f*1, which is the lowest frequency component for each of them. The lowest frequency component of a

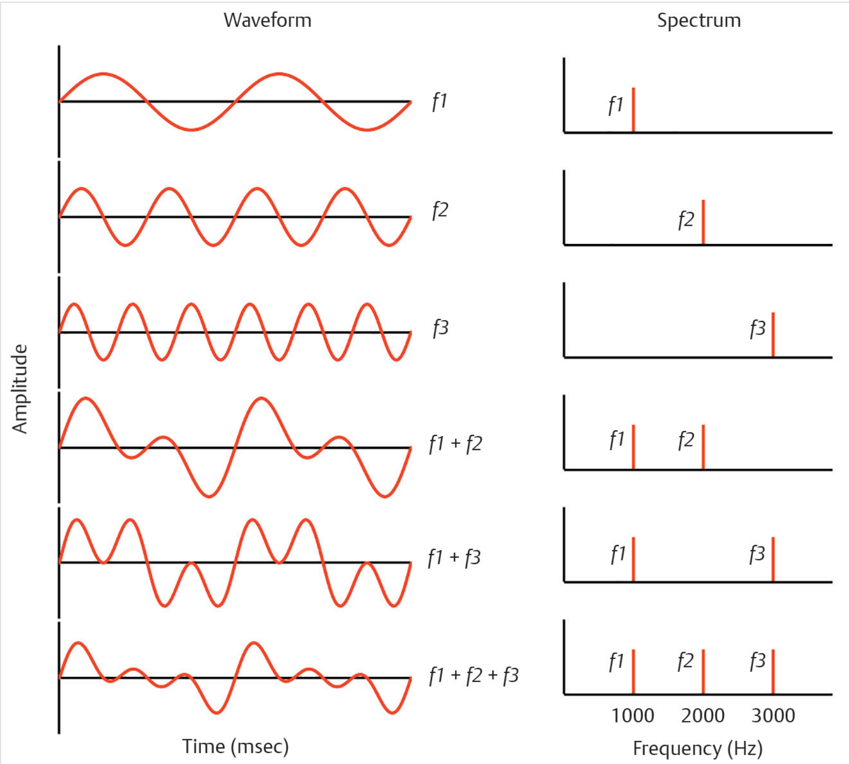


Fig. 1.12 Waveforms *(left)* and corresponding spectra *(right)* for three harmonically related sine waves (pure tones) of frequencies *f*1, *f*2, and *f*3; and complex periodic waves resulting from the in-phase addition of *f*1 + *f*2, *f*1 + *f*3, and *f*1 + *f*2 + *f*3. Notice that the fundamental frequency is *f*1 for all of the complex waves. Also notice that each pure tone spectrum has one vertical line, while the spectrum of each complex periodic sound has a separate vertical line for each of its components.

complex periodic wave is called its **fundamental frequency**. The fundamental frequency of each of the complex periodic waves in the figure is 1000 Hz because *f*1 is the lowest component in each of them. The period (or the time needed for one complete replication) of a complex periodic wave is the same as the period of its fundamental frequency. **Harmonics** are whole number or integral multiples of the fundamental frequency. In other words, the fundamental is the largest whole number common denominator of a wave's harmonics, and the harmonics are integral multiples of the fundamental frequency. In fact, the fundamental is also a harmonic because it is equal to 1 times itself. In the case of wave *f*1 + *f*2 + *f*3, 1000 Hz is the fundamental (first harmonic), 2000 Hz is the second harmonic, and 3000 Hz is the third harmonic.

Another example of combining sinusoids into a complex periodic wave is given in **Fig. 1.13.** Here, the sine waves being added are odd harmonics of the fundamental (1000 Hz, 3000 Hz, 5000 Hz, etc.), and their amplitudes get smaller with increasing frequency. The resulting complex periodic wave becomes squared off as the number of odd harmonics is increased, and is called a **square wave** for this reason.

Waveforms show how amplitude changes with time. However, the frequency of a pure tone (sine wave) is not directly provided by its waveform, and the frequencies in a complex sound cannot be determined by examining its waveform. In fact, the same frequencies can result in dramatically different-looking complex waveforms if their phase relationships are changed. Hence, another kind of graph is needed when we want to know what frequencies are present. This kind of graph is a **spectrum**, which shows amplitude on the *y*-axis as a function of frequency along the *x*-axis. Several examples are given in **Fig. 1.12** and **Fig. 1.13**. The frequency of the pure tone is given by the location of a vertical line along the horizontal (frequency) axis, and the amplitude of the tone is represented by the height of the line. According to Fourier's **theorem**, complex sounds can be mathematically dissected into their constituent pure tone components. The process of doing so is called **Fourier analysis**, which results in the information needed to plot the spectrum of a complex sound. The spectrum of a complex periodic sound has as many vertical lines as there are component frequencies. The locations of the lines show their frequencies, and their heights show their amplitudes, as illustrated in **Fig. 1.13**.

Aperiodic Waves

Aperiodic sounds are made up of components that are not harmonically related and have waveforms that do not repeat themselves over time, which is why they are called aperiodic. The extreme cases of aperiodic sounds are transients and random noise.

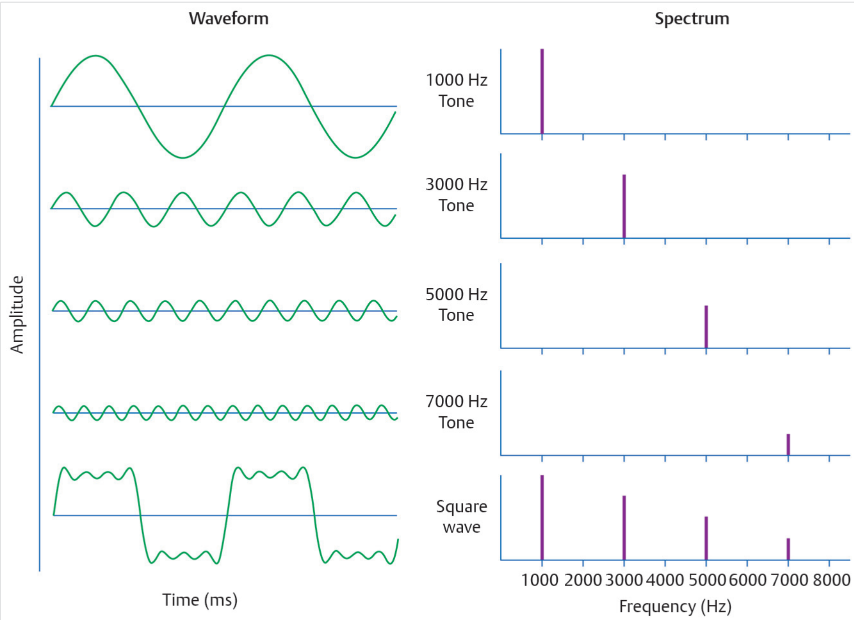


Fig. 1.13 Waveforms (*left*) and corresponding spectra (*right*) of odd harmonics combined to form a square wave (*bottom*). Notice that the spectrum of a pure tone has one vertical line, whereas the spectrum of a complex periodic sound has a separate vertical line for each of its frequency components.

A **transient** is an abrupt sound that is extremely brief in duration. It is aperiodic by definition because its waveform is not repeated (**Fig. 1.14a**). **Random noise** has a completely random waveform (**Fig. 1.14b**) so that it contains all possible frequencies at the same average amplitude over the long run. Random noise is also called **white noise** in analogy to white light because all possible frequencies are represented.

The spectrum of white noise is depicted in **Fig. 1.15a**. Individual vertical lines are not drawn because there would be an infinite number of them. It is more convenient to draw a continuous line over their tops and leave out the vertical lines themselves. This kind of spectrum is used for most aperiodic sounds and is called a **continuous spectrum**. *Random noise* has a *flat* continuous spectrum because the amplitudes are the same, on average, for all frequencies. An ideal *transient* also has *a flat spectrum*.

Most aperiodic sounds do not have flat spectra because they have more amplitude concentrated in one frequency range or the other. This notion is demonstrated by a simple experiment: Tap or scratch several different objects. The resulting noises will sound different from one another because they have energy concentrations in different frequency ranges. This is exactly what is represented on the spectrum. For example, different continuous spectra might show concentrations of aperiodic sounds with greater amounts of amplitude at higher frequencies (**Fig. 1.15b**), at lower frequencies (**Fig. 1.15c**), or within a particular band (range) of frequencies (**Fig. 1.15d**).

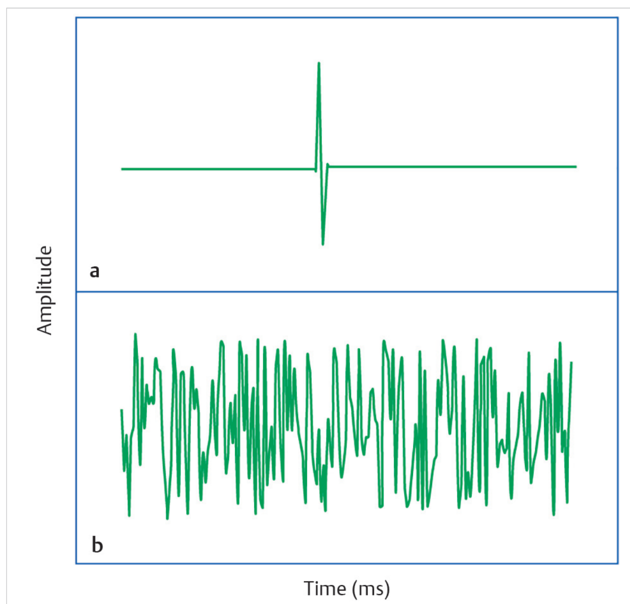


Fig. 1.14 Artist's conceptions of the waveforms of **(a)** a transient and **(b)** random or white noise.

Standing Waves and Resonance

The frequency(ies) at which a body or medium vibrates most readily is (are) called its **natural** or **resonant frequency(ies)**. Differences in resonant frequency ranges enable different devices or other objects to act as filters by transmitting energy more readily for certain frequency ranges than for others. Examples of the frequency ranges that are transmitted by high-, low-, and band-pass filters are illustrated in **Fig. 1.15**.

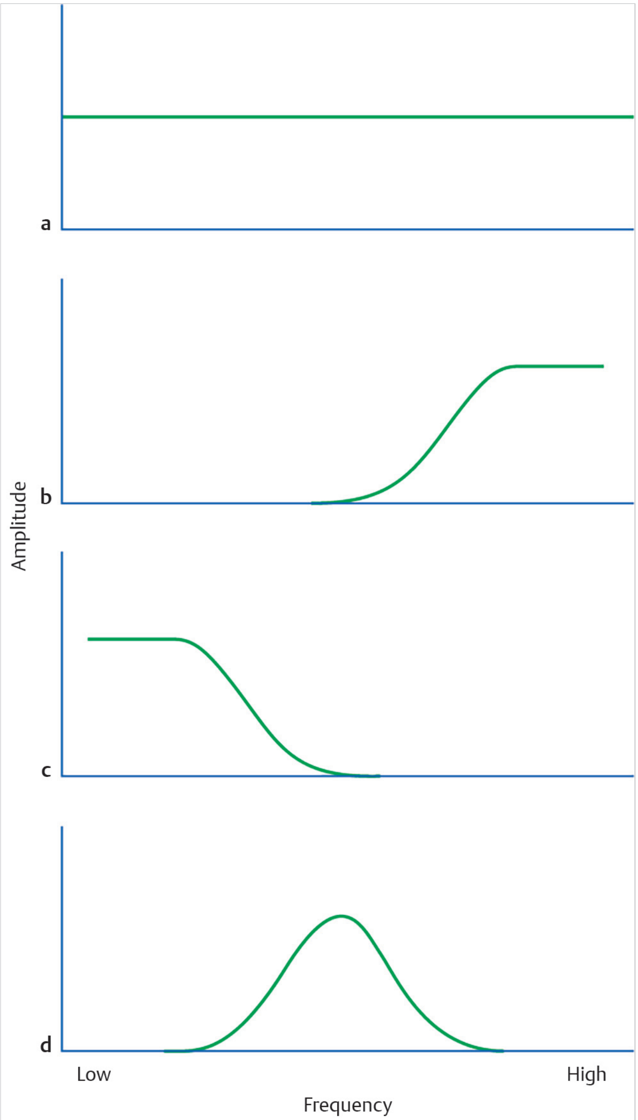


Fig. 1.15 Idealized continuous spectra showing **(a)** equal amplitude at all frequencies for a transient or white noise; **(b)** greater amplitude in the higher frequencies (also a high-pass filter); **(c)** greater amplitude in the lower frequencies (also a low-pass filter); and (**d**) amplitude concentrated within a certain band of frequencies (also a band-pass filter).

Vibrating strings Consider what happens when you pluck a guitar string. The waves initiated by the pluck move outward toward the two tied ends of the string. The waves are then reflected back and they propagate in opposite directions. The result is a set of waves that are moving toward each other, a situation that is sustained by continuing reflections from the two ends. Being reflections of one another, all of these waves will have the same frequency, and they will, of course, be propagating at the same velocity. Recall that waves interact with one another so that their instantaneous displacements add algebraically. As a result, the net displacement of the string at any point along its length will be due to the way in which the superimposed waves interact. The resulting combined wave appears as a pattern that is standing still even though it is derived from the interaction of waves, which themselves are propagating. Consequently, the points of maximum displacement (peaks of the combined wave pattern) and no displacement (baseline crossings of the combined wave pattern) will occur at fixed locations along the string. This pattern is called a **standing wave**.

Places along the string where there is zero displacement in the standing wave pattern are called **nodes**, and the locations where maximum displacement occurs are called **antinodes**. The string is tied down at its two ends so that it cannot move at these locations. In other words, the displacement must always be zero at the two ends of the string. This means that the standing wave pattern must have nodes that occur at the two ends of the string. Just as no displacement occurs at each end of the string because it is most constrained at these points, the greatest displacement occurs in the middle of the string, where it is least constrained and therefore has the most freedom of motion. In other words, the standing wave pattern will have a node at each end of the string and an antinode at the center of the string. This pattern is illustrated in **Fig. 1.16a**. Notice that the antinode occurs halfway between the nodes just as peaks (at 90° and 270°) alternate with zero displacements (at 0° and 180°) in a cycle of a sine wave.

The standing wave pattern that has a node at each end and an antinode in the center is not the only one that can occur on a given string; rather, it is just the *longest* one. This longest standing wave pattern is called the **first mode** of vibration. This standing wave pattern goes from no displacement to a peak and back to no displacement, which is analogous to going from 0° to 180° on a wave cycle. In other words, the pattern comprises exactly half of a cycle. Because we are dealing with displacement as a function of *distance* along the string (rather than over time), the parameter of the cycle with which we are dealing here is its wavelength (λ). In other words, the length of the longest standing wave pattern is the length of the whole string, and this length corresponds to one half of a wavelength ($\lambda/2$). Of course, if we know $\lambda/2$, then we can easily figure out λ. Now, a given wavelength is associated with a particular frequency because $f = c/\lambda$ (recall that c is the speed of sound). Consequently, the first mode of vibration is equal to half the wavelength ($\lambda/2$) of some frequency, which will, in turn, be its frequency of vibration. This will be the lowest **resonant frequency** of the string, which is its *fundamental frequency.*

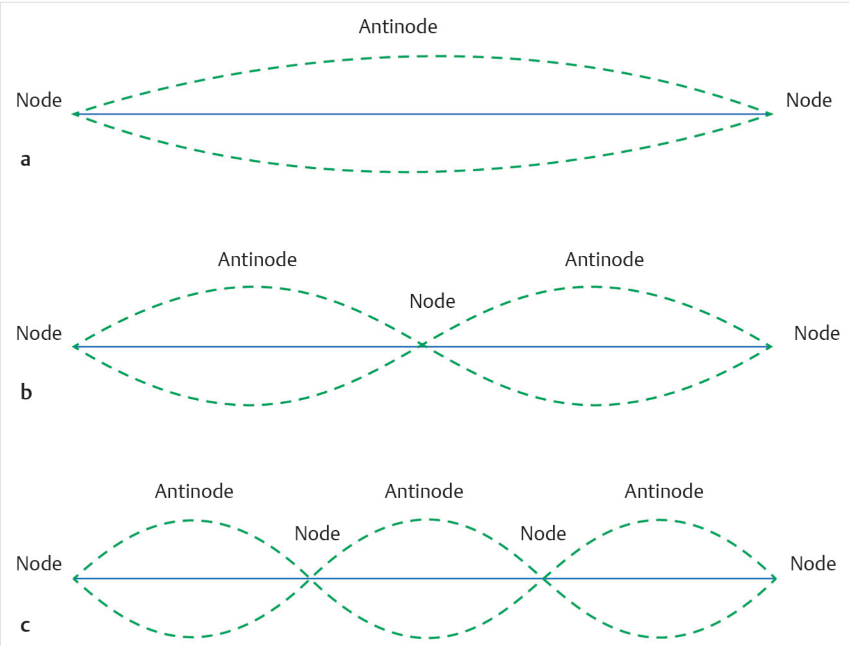


Fig. 1.16 Standing wave patterns corresponding to the **(a)** first, **(b)** second, and **(c)** third modes of a vibrating string.

Finding this frequency is a matter of substituting what we know into the formula $f = c/\lambda$. We know the length of the string (L), and we also know that $L = \lambda/2$. Therefore, $\lambda = 2L$. Substituting $2L$ for λ, the string's lowest resonant frequency is found with the formula $f = c/2L$. In reality, the speed of sound (c) is different for a vibrating string than it is for air. For the benefit of the math-minded student, the value of c for a string equals the square root of the ratio of its tension (T) to its mass (M), so that the real formula for the string's resonant frequency F_0 is

$$F_0 = \frac{1}{2L} \cdot \sqrt{\frac{T}{M}}$$

Other standing waves can also develop, provided they meet the requirement that there must be a node at each end of the string. For this criterion to be met, the string must be divided into exact halves, thirds, fourths, etc., as illustrated in **Fig. 1.16**. These standing wave patterns are called the second, third, fourth, etc., modes of vibration. Because the segments of the second mode are exactly half of the length of first mode, they produce a frequency that is exactly twice the fundamental. If we call the fundamental the first harmonic, then the second mode produces the second harmonic. Similarly, the segments of the third mode are exactly one-third the length of the first mode, so that they produce a third harmonic that is exactly three times the fundamental. The same principles apply to the fourth mode and harmonic, and beyond.

Vibrations in tubes The column of air inside a tube can be set into vibration by various means, such as blowing across the tube's open end. If this is done with different-size tubes, we would find that (1) shorter tubes are associated with higher pitches than longer ones and (2) the same tube produces a higher pitch when it is open at both ends compared with when it is open at one end and closed at the other.

When a column of air is vibrating in a tube that is open at both ends, the least amount of particle displacement occurs in the center of the tube, where the pressure is greatest. The greatest amount of displacement occurs at the two open ends, where the pressure is lowest. Hence, there will be a standing wave that has a displacement node in the middle of the tube and antinodes at the two ends, as illustrated in **Fig. 1.17a**. This standing wave pattern involves one half of a cycle in the sense that going from one end of the tube to the other end involves going from a displacement peak to a zero crossing to another peak. This trip would cover 180° (half of a cycle) on a sine wave, and thus a distance corresponding to half of a wavelength. Because this longest standing wave involves half of a wavelength ($\lambda/2$), the tube's lowest resonant (fundamental) frequency must have a wavelength that is twice the length of the tube (where $\lambda = 2L$). For this reason, tubes open at both ends are **half-wavelength resonators**. In other words, the lowest resonant frequency of a tube open at both ends is determined by the familiar formula $f = c/2L$. We could also say that the longest standing wave pattern is the first mode of vibration for the tube and that it is related to its fundamental frequency (lowest harmonic). As for the vibrating string, each successive higher mode

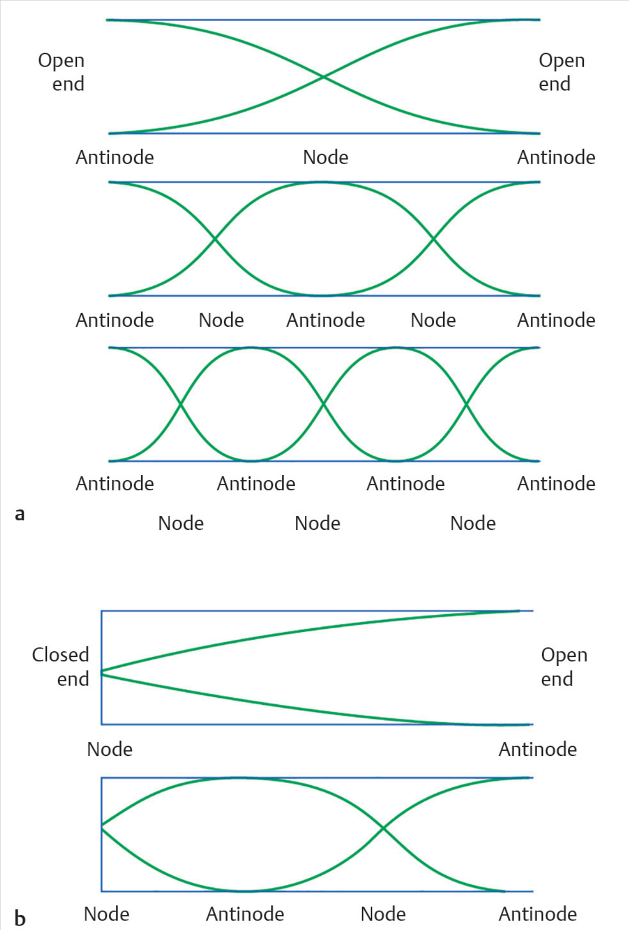


Fig. 1.17 Standing waves patterns in **(a)** a tube open at both ends (a half-wavelength resonator) and **(b)** a tube open at one end and closed at the other end (a quarter-wavelength resonator).

corresponds to exact halves, thirds, etc., of the tube length, as illustrated in **Fig. 1.17a**. In turn, these modes produce harmonics that are exact multiples of the fundamental frequency. Harmonics will occur at each multiple of the fundamental frequency for a tube open at both ends.

Air particles vibrating in a tube that is closed at one end and open at the other end are restricted most at the closed end. As a result, their displacement will be least at the closed end, where the pressure is the greatest. Thus, in terms of displacement, there must be a node at the closed end and an antinode at the open end, as illustrated in **Fig. 1.17b**. This pattern is analogous to the distance from a zero crossing to a peak, which corresponds to one quarter of a cycle (0° to 90°), and a distance of one quarter of a wavelength ($\lambda/4$). Because the length of the tube corresponds to $\lambda/4$, its lowest resonant frequency has a wavelength that is four times the length of the tube ($4L$). Hence, $f = c/4L$. For this reason, a tube that is open at one end and closed at the other end is a **quarter-wavelength resonator**. Because a node can occur at only one end, these tubes have only odd modes of vibration and produce only odd harmonics of the fundamental frequency (e.g., $f1$, $f3$, $f5$, $f7$, etc.), as illustrated in the figure.

Immittance

Immittance is the general term used to describe how well energy flows through a system. The *opposition* to the flow of energy is called **impedance (*Z*)**. The inverse of impedance is called **admittance (*Y*)**, which is the *ease* with which energy flows through a system.

The concept of impedance may be understood in terms of the following example. (Although this example only considers mass, we will see that immittance actually involves several components.) Imagine two metal blocks weighing different amounts. Suppose you repetitively push and pull the lighter block back and forth across a smooth table top with a certain amount of effort. This is a mechanical system in which a sinusoidally alternating force (the pushing and pulling) is being applied to a mass (the block). The effort with which you are pushing (and pulling) the block is the amount of applied force, and the velocity of the block reflects how well energy flows through this system to effect motion. A particular block will move at a certain velocity given the amount of effort you are using to push (and pull) it. If the *same* amount of *effort* was used to push and pull the *heavier* block, then it would move slower than the first one. In other words, the heavier block (greater mass) moves with less velocity than the lighter block (smaller mass) in response to the same amount of applied force. We can say that the flow of energy is opposed more by the heavier block than by the lighter one. For this reason, the heavier block (greater mass) has more impedance and less admittance than the lighter block (smaller mass).

This example shows that impedance and admittance are viewed in terms of the relationship between an applied force and the resulting amount of velocity. In effect, higher impedance means that more force must be applied to result in a given amount of velocity, and lower impedance means that less force is needed to result in a given amount of velocity. For the mathematically oriented, we might say that impedance (Z) is the ratio of force to velocity:

$$Z = \frac{F}{v}$$

The amount of impedance is expressed in **ohms**. The larger the number of ohms, the greater the opposition to the flow of energy.

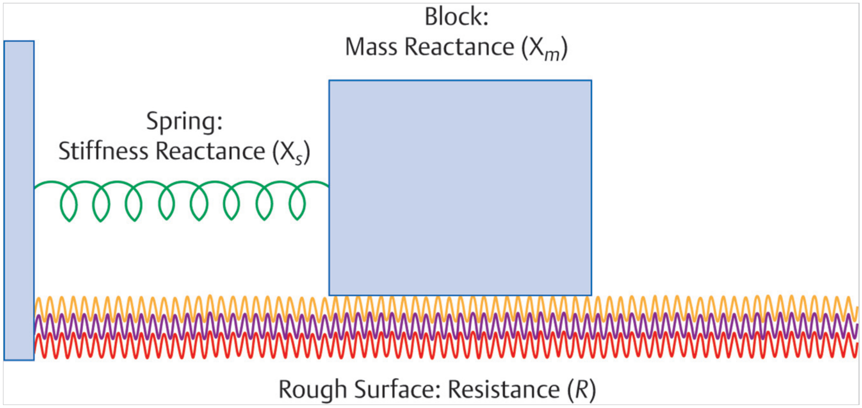


Fig. 1.18 The components of impedance are (1) mass reactance (X_m), represented by the block; (2) stiffness reactance (X_s), represented by the spring; and (3) resistance (R), represented by the rough surface under the block.

Admittance (Y) is the reciprocal of impedance:

$$Y = \frac{1}{Z}$$

and is therefore equal to the ratio of velocity to force:

$$Y = \frac{v}{F}$$

As we might expect, the unit of admittance is the inverse of the ohm, and is therefore called the **mho**. The more mhos, the greater the ease with which energy flows. The admittance values that we are concerned with in audiology are very small, and are thus expressed in **millimhos (mmhos)**.

Impedance involves the complex interaction of three familiar physical components: mass, stiffness, and friction. In **Fig. 1.18** mass is represented by the block, stiffness (or compliance) by the spring, and friction by the rough surface under the block. Let's briefly consider each of these components. Friction dissipates some of the energy being introduced into the system by converting it into heat. This component of impedance is called resistance (R). The effect of resistance occurs in-phase with the applied force (**Fig. 1.19**). Some amount of friction is always present. Opposition to the flow of energy due to mass is called **mass (positive) reactance (X_m)** and is related to inertia. The opposition due to the stiffness of a system is called **stiffness (negative) reactance (X_s)**, and is related to the restoring force that develops when an elastic element (e.g., a spring) is displaced.

Mass and stiffness act to oppose the applied force because these components are out-of-phase with it (**Fig. 1.19**). They oppose the flow of energy by storing it in these out-of-phase components before effecting motion. First, consider the mass all by itself. At the same point in time (labeled 1) the applied force is maximal (in the upward direction) and the velocity of the block (mass) is zero (crossing the horizontal line in the positive direction). One quarter of

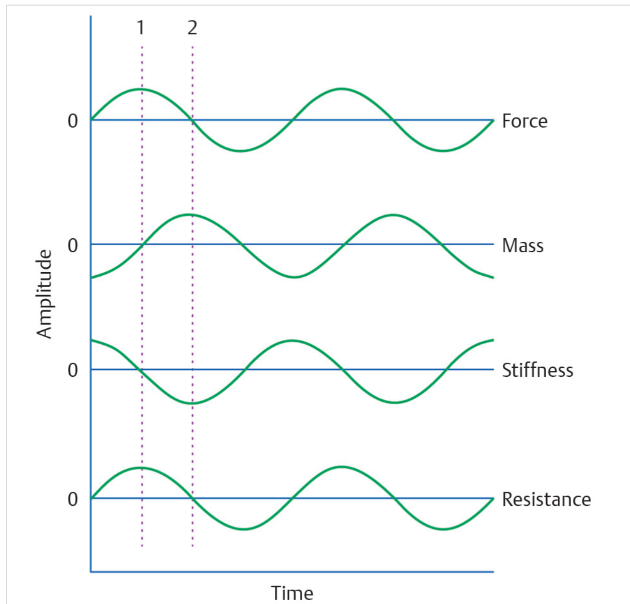


Fig. 1.19 Relationship between a sinusoidally applied force (top) and the velocities associated with mass, stiffness, and resistance. The dotted lines labeled 1 and 2 show two moments in time. Resistance is in-phase with the applied force (F). Mass and stiffness are 90° out-of-phase with force and 180° out-of-phase with each other.

a cycle later, at the time labeled 2, the applied force is zero and the velocity of the mass is now maximal (in the upward direction). Hence, a sinusoidally applied force acting on a mass and the resulting velocity of the mass are a quarter-cycle (90°) out-of-phase. To appreciate this relationship, hold a weight and shake repetitively back and forth from right to left. You will feel that you must exert the most effort (maximal force) at the extreme right and left points of the swing, where the direction changes. Notice that the weight is momentarily still (i.e., its velocity is zero) at the extreme right and left points because this is where it changes direction. On the other hand, the weight will be moving the fastest (maximum velocity) as it passes the midpoint of the right-to-left swing, which is also where you will be using the least effort (zero force).

Now consider the stiffness all by itself at the same two times in Fig. 1.19. At time 1, when the applied force is maximal (upward), the velocity of the spring (stiffness) is zero (crossing the horizontal line in the negative direction). One quarter-cycle later, at time 2, the applied force is zero, and the velocity of the spring is now maximal (downward). Hence, a sinusoidally applied force acting on a spring and the resulting velocity of the spring (stiffness) are a quarter-cycle (90°) out-of-phase. This occurs in the opposite direction of what we observed for the mass (whose motion is associated with inertia) because the motion of the spring (stiffness) is associated with restoring force. You can appreciate this relationship of the stiffness component by alternately compressing and expanding a spring. You must push or pull the hardest (maximum applied force) at the moment when the spring is maximally expanded (or compressed), which is also when the spring is not moving (zero velocity) because it is about to change direction. Similarly, you will exert no effort (zero applied force) and the spring will be moving the fastest (maximum velocity) as it moves back through its "normal" position (where it is neither compressed nor expanded).

Notice that the mass and stiffness reactances are 180° out-of-phase with each other (Fig. 1.19). This means that the effects of mass reactance and stiffness reactance oppose each other. As a result, the **net reactance (X_{net})** is the difference between them, so that

$$X_{net} = X_s - X_m$$

when stiffness reactance is larger, or

$$X_{net} = X_m - X_s$$

when mass reactance is larger. For example, if X_s is 850 ohms and X_m is 140 ohms, then X_{net} will be 850 – 140 = 710 ohms of stiffness reactance. If X_m is 1000 ohms and X_s is 885 ohms, then X_{net} will be 1000 – 885 = 115 ohms of mass reactance.

The overall impedance is obtained by combining the resistance and the net reactance. This cannot be done by simple addition because the resistance and reactance components are out-of-phase. (Recall here the difference between scalars and vectors mentioned at the beginning of the chapter.) The relationships in Fig. 1.20 show how impedance is derived from resistance and reactance. The size of the resistance component is plotted along the *x*-axis. Reactance is plotted on the *y*-axis, with mass (positive) reactance represented upward and stiffness (negative) reactance downward. The net reactance here is plotted downward because X_s is greater

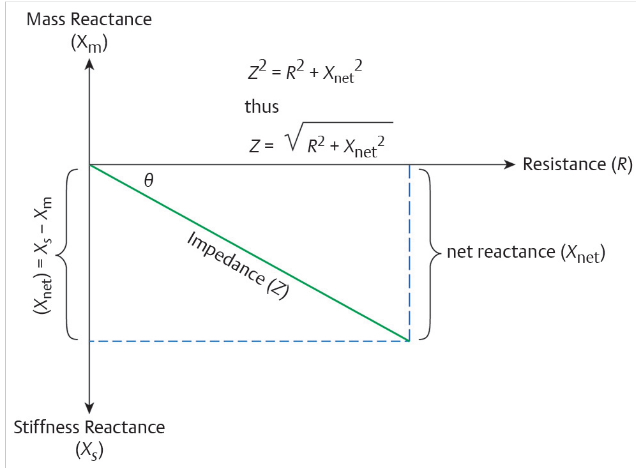


Fig. 1.20 Impedance (Z) is the complex interaction of resistance (R) and the net reactance (X_{net}, which is equal to $X_s - X_m$). Notice the impedance value is determined by the vector addition of the resistance and reactance. The angle (9) between the horizontal leg of the triangle (resistance) and its hypotenuse (impedance) is called the phase angle.

than X_m, so that X_{net} is negative. Notice that R and X_{net} form two legs of a right triangle, and that Z is the hypotenuse. Hence, we find Z by the familiar Pythagorean theorem ($a^2 + b^2 = c^2$), which becomes $Z^2 = R^2 + X_{net}^2$. Removing the squares gives us the formula for calculating impedance from the resistance and reactance components:

$$Z = \sqrt{R^2 + X_{net}^2}$$

Resistance tends to be essentially the same at all frequencies. However, reactance depends on frequency (f) in the following way: (1) mass reactance is proportional to frequency,

$$X_m = 2\pi f M$$

where M is mass; and (2) stiffness reactance is inversely proportional to frequency,

$$X_s = \frac{S}{2\pi f}$$

where S is stiffness. In other words, X_m gets larger as frequency goes up, and X_s gets larger as frequency goes down. Because of these frequency relationships, impedance also depends on frequency:

$$Z = \sqrt{R^2 + \left(\frac{S}{2\pi f} + 2\pi f M\right)^2}$$

In addition, there will be a frequency where X_m and X_s are equal, and thus cancel. This is the resonant frequency, where the only component that is opposing the flow of energy is resistance.

Admittance is the reciprocal of impedance,

$$Y = \frac{1}{Z}$$

and the components of admittance are the reciprocals of resistance and reactance: **conductance (*G*)** is the reciprocal of resistance,

$$G = \frac{1}{R}$$

stiffness (compliance) susceptance (*B_s*) is the reciprocal of stiffness reactance,

$$B_s = \frac{1}{X_s}$$

and mass **susceptance (*B_m*)** is the reciprocal of mass reactance:

$$B_m = \frac{1}{X_m}$$

Stiffness susceptance is proportional to frequency (B_s increases as frequency goes up), and mass susceptance is inversely proportional to frequency (B_m decreases as frequency goes up). Net susceptance (B_{net}) is the difference between B_s and B_m. The formula for admittance is

$$Y = \sqrt{G^2 + B_{\text{net}}^2}$$

where B_{net} is (B_s – B_m) when B_s is bigger and (B_m – B_s) when B_m is larger.

Up to this point we have discussed immittance in mechanical terms. **Acoustic immittance** is the term used for the analogous concepts when dealing with sound. The opposition to the flow of sound energy is called **acoustic impedance (*Z_a*)**, and its reciprocal is **acoustic admittance (*Y_a*)**. Thus,

$$Z_a = \frac{1}{Y_a}$$

and

$$Y_a = \frac{1}{Z_a}$$

When dealing with acoustic immittance, we use **sound pressure (*p*)** in place of force, and velocity is replaced with the velocity of sound flow, called **volume velocity (*U*)**. Thus, acoustic impedance is simply the ratio of sound pressure to volume velocity,

$$Z_a = \frac{p}{U}$$

and acoustic admittance is the ratio of volume velocity to sound pressure,

$$Y_a = \frac{U}{p}$$

The components of acoustic immittance are based on the acoustic analogies of friction, mass, and stiffness (compliance). Friction develops between air molecules and a mesh screen, which is thus used to model **acoustic resistance (*R_a*)**. **Mass (positive) acoustic reactance (+*X_a*)** is represented by a slug of air in an open tube. Here an applied sound pressure will displace the slug of air as a unit, so that its inertia comes into play. A column of air inside a tube open at one end and closed at the other end represents **stiffness (negative) acoustic reactance (–*X_a*)** because sound pressure compresses the air column like a spring. The formulas and relationships for acoustic immittance are the same as those previously given, except that the analogous acoustic values are used. For example, acoustic impedance is equal to

$$Z_a = \sqrt{R_a^2 + X_a^2}$$

where X_a is the net difference between stiffness acoustic reactance (–X_a) and mass acoustic reactance (+X_a). Similarly, the formula for acoustic admittance is

$$Y_a = \sqrt{G_a^2 + B_a^2}$$

where B_a is the net difference between stiffness acoustic susceptance (+B_a) and mass acoustic susceptance (–B_a).

Expressing Values in Decibels

It is extremely cumbersome to express sound magnitudes in terms of their actual intensities or pressures for several reasons. To do so would involve working in units of watts/m^2 (or watts/cm^2) and newtons/m^2 (or dynes/cm^2). In addition, the range of sound magnitudes with which we are concerned in audiology is enormous; the loudest sound that can be tolerated has a pressure that is roughly 10 million times larger than the softest sound that can be heard. Even if we wanted to work with such an immense range of cumbersome values on a linear scale, we would find that it is hard to deal with them in a way that has relevance to the way we hear. As a result, these absolute physical values are converted into a simpler and more convenient form called **decibels (dB)** to make them palatable and meaningful.

The decibel takes advantage of ratios and logarithms. Ratios are used so that physical magnitudes can be stated in relation to a reference value that has meaning to us. It makes sense to use the *softest sound that can be heard by normal people* as our **reference value**. This reference value has an *intensity* of

$$10^{-12} \text{ W/m}^2$$

in MKS units, which corresponds to

$$10^{-16}\ \mathrm{W/cm^2}$$

in the cgs system. The same softest audible sound can also be quantified in terms of its *sound pressure*. This reference pressure is

$$2 \times 10^{-5}\ \mathrm{N/m^2}$$

or

$$20\ \mu\mathrm{Pa}$$

in the MKS system.[2] In cgs units this reference pressure is

$$2 \times 10^{-4}\ \mathrm{dyne/cm^2}$$

(The student will find that 2×10^{-4} dyne/cm^2 is also written as 0.0002 dynes/cm^2, 2×10^{-4} µbar, or 0.0002 µbar, especially in older literature.) The appropriate reference value (intensity or pressure, MKS or cgs) becomes the denominator of our ratio, and the intensity (or pressure) of the sound that is actually being measured or described becomes the numerator. As a result, instead of describing a sound that has an intensity of 10^{-10} W/m^2, we place this value into a ratio so we can express it in terms of how it compares to our reference value (which is 10^{-12} W/m^2). Hence, this ratio would be

$$\frac{10^{-10}\mathrm{W/m^2}}{10^{-12}\mathrm{W/m^2}}$$

This ratio reduces to simply 10^2, or 100.

Regardless of what this ratio turns out to be, it is replaced with its common logarithm because equal ratios correspond to different distances on a linear scale, but equal ratios correspond to equal distances on a logarithmic scale. In other words, the linear distance between two numbers with the same ratio relationship (e.g., 2 to 1) is small for small numbers and large for large numbers, but the logarithm of that ratio is always the same. For example, all of the following pairs involve 2/1 ratios. Even though the linear distance between the numbers in the pairs gets wider as the absolute sizes of the numbers get larger, the logarithm of all of the ratios stays the same (2/1 = 2, and log 2 is always 0.3; **Table 1.4**).

The general decibel formula is expressed in terms of *power* as follows:

$$PL = 10\log\frac{\mathrm{P}}{\mathrm{P_0}}$$

Here, *PL* stands for **power level** (in dB), *P* is the power of the sound being measured, and P_0 is the reference power to which the former is being compared. The word *level* is added to distinguish the raw physical quantity (power) from the corresponding decibel value (which is a logarithmic ratio about the power). Similarly, intensity expressed in decibels is called **intensity level (IL)** and sound pressure expressed in decibels is called **sound pressure level (SPL)**.

Table 1.4 All pairs of numbers that have the same ratio between them (e.g., 2:1) are separated by the same logarithmic distance even though their linear distances are different

Pairs of numbers with 2:1 ratios	Distances between the absolute numbers get wider	Logarithms of all 2:1 ratios are the same
2/1	1	**0.3**
8/4	4	**0.3**
20/10	10	**0.3**
100/50	50	**0.3**
200/100	100	**0.3**
2000/1000	1000	**0.3**

Most sound measurements are expressed in terms of intensity or sound pressure, with the latter being the most common. The formula for decibels of **intensity level** is

$$IL = 10\log\frac{I}{I_0}$$

where *IL* is intensity level in dB, *I* is the intensity of the sound in question (in W/m^2), and I_0 is the reference intensity (10^{-12} W/m^2). If the value of *I* is 10^{-10} W/m^2, then

$$IL = 10\log\frac{10^{-10}\ \mathrm{W/m^2}}{10^{-12}\ \mathrm{W/m^2}}$$

$$= 10\log\frac{10^{-10}}{10^{-12}}\ (\text{notice that W/m}^2\text{ cancels out})$$

$$= 10\log 10^{(-10)-(-12)}$$

$$= 10\log 10^2$$

$$= 10 \times 2$$

$$= 20\ \mathrm{dB\ re:}\ 10^{-12}\mathrm{W/m^2}$$

$$= 20\ \mathrm{dB\ IL}$$

Consequently, an absolute intensity of 10^{-10} W/m^2 has an intensity level of 20 dB re: 10^{-12} W/m^2, or 20 dB IL. The phrase "re: 10^{-12} W/m^2" is added

[2] One pascal (Pa) = 1 N/m^2. Thus, 10^{-6} N/m^2 = 1 micropascal (µPa), 10^{-5} N/m^2 = 10 µPa, and 2×10^{-5} N/m^2 = 20 µPa.

because the decibel is a dimensionless quantity that has real meaning only when we know the reference value, that is, the denominator of the ratio.

The formula for decibels of **sound pressure level (dB SPL)** is obtained by replacing all of the intensity values with the corresponding values of *pressure squared* (because $I \propto p^2$):

$$SPL = 10\log\frac{p^2}{p_0^2}$$

Here, p is the measured sound pressure (in N/m^2) and p_0 is the reference sound pressure (2×10^{-5} N/m^2, or 20 µPa). This form of the formula is cumbersome because of the squared values, which can be removed by applying the following steps:

$$SPL = 10\log\frac{p^2}{p_0^2}$$

$$SPL = 10\log\left(\frac{p}{p}\right)^2 \left(\text{because } \frac{x^2}{y^2} = \left(\frac{x}{y}\right)^2\right)$$

$$SPL = 10 \times 2\log\left(\frac{p}{p_0}\right) \quad (\text{because } x^2 = 2\log x)$$

$$SPL = 20\log\left(\frac{p}{p_0}\right)$$

Therefore, the commonly used simplified formula for decibels of SPL is

$$SPL = 20\log\frac{p}{p_0}$$

where the multiplier is 20 instead of 10 as a result of removing the squares from the unsimplified version of the formula.

Let us go through the exercise of converting the absolute sound pressure of a sound into dB SPL. We will assume that the sound being measured has a pressure of 2×10^{-4} N/m^2. Recall that the reference pressure is 2×10^{-5} N/m^2. The steps are as follows:

$$SPL = 20\ \log\frac{2 \times 10^{-4}\ N/m^2}{2 \times 10^{-5}\ N/m^2}$$

$$= 20\log\frac{10^{-4}}{10^{-5}} (\text{notice that } N/m^2 \text{ cancels out})$$

$$= 20\log 10^{(-4)-(-5)}$$

$$= 20\log 10^1$$

$$= 20 \times 1 = 20$$

$$= 20\ \text{dB re}:\ 2 \times 10^{-5} N/m (\text{or } 20\ \mu Pa)$$

$$= 20\ \text{dB}\ SPL$$

Hence, a sound pressure of 2×10^{-4} N/m^2 corresponds to a sound pressure level of 20 dB re: 2×10^{-5} N/m^2 (or 20 µPa), or 20 dB SPL.

What is the decibel value of the reference itself? In other words, what would happen if the intensity (or pressure) being measured is equal to the reference intensity (or pressure)? In terms of intensity, the answer is found by using the reference value (10^{-12} w/m^2) as both the numerator (I) and denominator (I_0) in the dB formula, so that

$$IL = 10\ \log\frac{10^{-12}\ W/m^2}{10^{-12}\ W/m^2}$$

$$= 10\log 1 (\text{anything divided by itself equals } 1)$$

$$= 10 \times 0$$

$$= 0\ \text{dB re:}\ 10^{-12} W/m^2$$

$$= 0\ \text{dB IL}$$

Consequently, the intensity *level* of the reference intensity is 0 dB IL. Similarly, 0 dB SPL means that the measured sound pressure corresponds to that of the reference sound:

$$SPL = 20\ \log\frac{2 \times 10^{-5}\ N/m^2}{2 \times 10^{-5}\ N/m^2}$$

$$= 20\log 1 (\text{anything divided by itself equals } 1)$$

$$= 20 \times 0$$

$$= 0\ \text{dB re}:\ 2 \times 10^{-5} N/m^2\ (\text{or } 20\mu Pa)$$

$$= 0\ \text{dB SPL}$$

Notice that 0 dB IL or 0 dB SPL means that the sound being measured is equal to the reference value; it does not mean "no sound." It follows that negative decibel values indicate that the magnitude of the sound is lower than the reference; for example, −10 dB means that the sound in question is 10 dB below the reference value.

Sound Measurement

The magnitude of a sound is usually measured with a device called a **sound level meter (SLM)**. This device has a high-quality microphone that picks up the sound and converts it into an electrical signal that is analyzed by an electronic circuit, and then displays the magnitude of the sound on a meter in decibels of sound pressure level (dB SPL). An example of an SLM is shown in **Fig. 1.21**. SLMs are used to calibrate or establish the accuracy of audiometers and other instruments used to test hearing, as well as to measure noise

Fig. 1.21 An example of a digital sound level meter with a measuring microphone attached.

Fig. 1.22 An example of a piston phone acoustical calibrator **(a)**, an analog barometer **(b)**, and a digital barometer **(c)**. (Photographs courtesy of GRAS Sound and Vibration.)

levels for such varied purposes as determining whether a room is quiet enough for performing hearing tests or identifying potentially hazardous noise exposures. SLMs or equivalent circuits that perform the same function are also found as components of other devices, such as hearing aid test systems.

The characteristics of SLMs are specified in the ANSI S1.4–2014 (R2019) standard. The accuracy of the measurements produced by an SLM is established using a compatible **acoustical calibrator**, which is a device that produces a known precise signal that is directed into the SLM microphone. Fig. 1.22 shows an example of one type of acoustical calibrator, known as a **piston phone** because of the way it works. A barometer is shown to the right of the piston phone, which is needed because these measurements are affected by barometric pressure. For example, if the calibrator produces a signal that is exactly 114 dB SPL, then the SLM is expected to read this amount when it is connected to the calibrator (within certain tolerances allowed in the ANSI standard). If the meter reading deviates from the actual value of 114 dB SPL, then its controls are adjusted to reset the meter to the right value, or it might be necessary to have the SLM repaired and recalibrated by the manufacturer or an instrumentation service company.

The microphone of an SLM picks up all sounds that are present at all frequencies within its operating range. However, the user might want to emphasize or de-emphasize certain frequency ranges (we'll talk about why later). The SLM has settings to make these adjustments or corrections, which are called **weightings**. In other words, the SLM has weightings that can treat the original sound levels as though they were stronger or weaker at certain frequencies. For example, imagine that the *actual* level picked up by the microphone at, say, 700 Hz is 49 dB. If the SLM has a *weighting* of −11 dB at that frequency, then it will be *treated as though* it was 49 − 11 = 38 dB. If the weighting is + 2 dB, then it would be treated as though it was 49 + 2 = 51 dB. One of these weightings is actually *no weighting at all*, or *unweighted*, so that the sound levels at all frequencies are treated exactly as they were picked up by the microphone. In this case, we could say that the *weighting* is *zero decibels*. This setting on the SLM is therefore called *Z* for *zero*, as in *zero adjustments* or a *zero weighting*. This setting is therefore the ***Z*-weighting** (sometimes called the *Z scale*). So, when the SLM is set to the *Z*-weighting it will display the overall SPL based on all of the sounds picked up by the microphone without making any adjustments or corrections. In this case, the

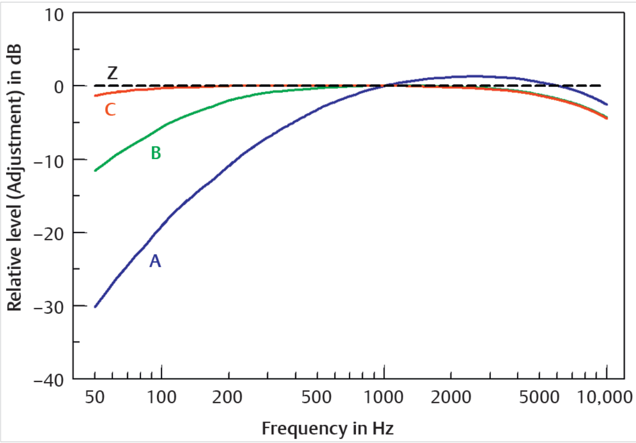


Fig. 1.23 Frequency response curves for the *A*-, *B*-, *C*-, and *Z*-weightings (see text).

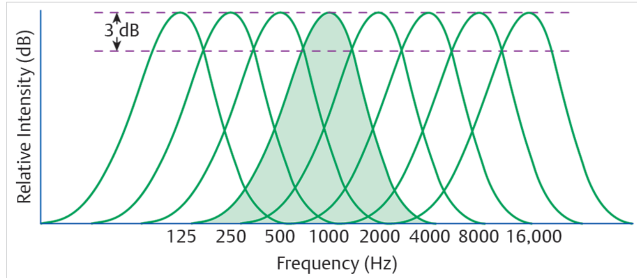


Fig. 1.24 A series of octave-bands representing those typically used in octave-band analysis. Notice the bands overlap at their 3 dB down points. The 1000 Hz octave-band is highlighted for clarity.

sound level is expressed in **dB SPL** or **dBZ**. (Some older SLMs called this the "linear" setting.)

SLMs also have other weightings that do change the emphasis given to certain parts of the spectrum, and these are used much more often than Z. In fact, the ANSI standard for SLMs actually considers the Z-weighting optional. We all know that turning up the *bass* on a radio, TV, or music system makes the low pitches more pronounced, whereas turning the bass down makes the lows less noticeable. In other words, the bass control determines whether the low frequencies will be emphasized or de-emphasized. The *treble* control does the same thing for the high frequencies. SLM **weightings** do essentially the same thing, mainly by de-emphasizing the low frequencies.

Fig. 1.23 shows several of the weightings that may be found on SLMs, of which **A** and **C** are the most common. The y-axis is relative level in decibels, and shows how many decibels are subtracted or added by the weighting at a given frequency. A horizontal line at 0 dB refers to how the sound would be without the weighting network. In other words, 0 dB here means "unchanged" or "0 dB of change," and corresponds to the Z-weighting as described above. Almost all of the adjustments are negative, and show how much the sound level is de-emphasized at each frequency.

The **A-weighting** considerably de-emphasizes the low frequencies, as shown by its curve, which gets progressively more negative as frequency decreases below 1000 Hz. For example, the curve shows that the A-weighting network de-emphasizes sounds by ~ 4 dB at 500 Hz, 11 dB at 200 Hz, 19 dB at 100 Hz, and 30 dB at 50 Hz. This is analogous to turning the bass all the way down on a stereo system. Sound level measurements using the A-weighting are expressed as **dBA**. The A-weighting is commonly used when it is desirable to exclude the effects of the lower frequencies, and are especially useful in noise level measurements.

The **B-weighting** network also de-emphasizes the lower frequencies, but not as much as the A-weighting. For example, the amount of reduction is only ~ 6 dB at 100 Hz. Sound level measurements using the B-weighting are expressed as **dBB**. However, the B-weighting is rarely used and no longer included in the ANSI standard for SLMs.

The **C-weighting** is barely different from a flat, unweighted (*Z*) response. Sound level measurements using the C-weighting are expressed as **dBC**. The C-weighting is commonly used as a proxy for the completely flat (*Z*) response, and useful in noise level measurements.

In addition to frequency weightings, SLMs often have **filters** that allow it to "look at" a certain range of frequencies instead of all of them. **Octave-band filters** separate the overall frequency range into narrower ranges, each of which is one octave wide, as illustrated in Fig. 1.24. For example, the range from 355 to 710 Hz is an octave-band because $710 = 2 \times 355$, and the bandwidth from 2800 to 5600 Hz is also an octave-band because $5600 = 2 \times 2800$. An octave-band is named according to its **center frequency**, but keep in mind that the center is defined as the *geometric mean* of the upper and lower cutoffs rather than the arithmetic midpoint between them. Hence, the 500 Hz octave-band goes from approximately 355 to 710 Hz, and the 4000 Hz octave-band includes 2800 to 5600 Hz. The center frequencies and the upper and lower cutoff frequencies of the octave-bands typically used in acoustical measurements are listed in Table 1.5.

Measuring a noise on an octave-band by octave-band basis is called **octave-band analysis** and makes it possible to learn about the spectrum of a sound instead of just its overall level. An even finer level of analysis can be achieved by using **third-octave-band filters**, in which case each filter is one-third of an octave wide. For example, the 500 Hz third-octave filter includes the frequencies between approximately 450 and 560 Hz, and the 4000 Hz third-octave-band goes from 3550 to

Table 1.5 Examples of octave-band center frequencies, and lower and upper cutoff frequencies

Center frequency (Hz)	Lower cutoff (Hz)	Upper cutoff (Hz)
31.5	22.4	45
63	45	90
125	90	180
250	180	355
500	355	710
1000	710	1400
2000	1400	2800
4000	2800	5600
8000	5600	11,200
16,000	11,200	22,630

4500 Hz. Octave-band and third-octave-band filters are useful when we want to concentrate on the sound level in a narrow frequency range without contamination from other frequencies. For example, it is usually better to measure the level of a 1000 Hz tone while using a filter centered around 1000 Hz than to do the same thing with the all-inclusive Z setting because the filter excludes other frequencies that would contaminate the results.

We can combine **octave-band levels (OBLs) or third-octave-band levels ($^1/_3$-OBLs)** to arrive at the overall level of a sound. There are two ways to combine OBLs into overall SPL. The simpler approach involves adding the OBLs in successive pairs using the rules for adding decibels shown in **Table 1.6**. It is very easy to use this table. First, find the difference in decibels between the two sounds being combined. For example, if one sound is 80 dB and the other is 76 dB, then the difference between them is 4 dB. Then find the increment that corresponds to the difference. According to the table, the increment for a 4 dB difference is 1.4 dB. Now, just add this increment to the larger of the original two sounds. The larger value in our example is 80 dB, so we add the 1.4 dB increment to 80 dB (80 + 1.4 = 81.4). Hence, combining 80 and 76 dB results in a total of 81.4 dB. To combine octave-bands into an overall level, simply arrange their OBLs from largest to smallest, and combine pairs successively using the increments in the table. A complete example is shown in **Appendix A**.

Table 1.6 Combining decibels: find the difference in decibels between the two sounds, and then add the corresponding decibel increment to the larger original decibel value

Difference in dB between original sounds	Increment in dB (add to larger original sound)
0	3.0
1	2.6
2	2.2
3	1.8
4	1.4
5	1.2
6	1.0
7	0.8
8	0.6
9	0.5
10	0.4
11	0.35
12	0.3
13	0.25
14	0.2
15	0.15
16	0.1

The more precise method for combining octave-band levels into an overall SPL is to use the following formula for logarithmic addition:

$$L = 10\log\sum_{i=1}^{n} 10^{Li/10}$$

In this formula, L is the overall (combined) level in dB SPL, n is the number of bands being combined, i is the ith band, and L_i is the OBL of the ith band. An example showing how this formula is used may be found in **Appendix A**.

The same methods can be used to combine octave-band levels into an A-weighted sound level (dBA), except that a correction factor is applied to each OBL. This correction factor is the amount by which the A-weighting de-emphasizes the level of the sounds within each octave-band. **Table 1.7** shows the corrections (dBA weightings) that can be used to convert unweighted octave-band levels into

Table 1.7 Corrections (dBA weightings) to convert unweighted octave-band levels into A-weighted octave-band levels

Octave-band center	
Frequency (Hz)	dBA weighting
31.5	–39.4
63	–26.2
125	–16.1
250	–8.6
500	–3.2
1000	0
2000	+ 1.2
4000	+ 1.0
8000	–1.1

A-weighted octave-band levels. For example, dBA de-emphasizes the 125 Hz octave-band by 16.1 dB. Thus, if the 125 Hz OBL is 60 dB, we correct it to its dBA value by subtracting: 60 – 16.1 = 43.9 dBA. A full example is shown in **Appendix A**. The formula for more precisely converting OBLs into dBA is as follows:

$$L_A = 10\log \sum_{i=1}^{n} 10^{(L_i+k_i)/10}$$

The symbols here are the same as in the previous formula except L_A is now the overall (combined) level in dBA, and k_i is the correction factor that must be applied to the OBL of the ith band to convert it into its equivalent value in dBA (which is the reason for the term $(L_i + k_i)$ in the equation). **Appendix A** shows an example of how this formula is used.

Study Questions

1. Define and specify the units of measurement for the following terms: displacement, velocity, acceleration, force, work, and power.
2. Explain pressure and intensity, state the reference values for each of them, and explain why we have these reference values.
3. Explain what happens to the intensity of a sound with increasing distance from the sound source.
4. Define simple harmonic motion and explain how its characteristics are depicted by a sine wave.
5. Define the terms *cycle*, *period*, *frequency*, and *wavelength*, and explain how they are related to each other.
6. Define friction and explain why it causes damping.
7. Define complex periodic and aperiodic waves, and describe how their characteristics are shown on the waveform and spectrum.
8. What are resonant frequencies and how are they related to standing waves?
9. Define impedance and describe how it is related to mass and stiffness.
10. What are the formulas and reference values for intensity level and sound pressure level in decibels? Explain how the magnitude of a sound is expressed in decibels.

References

American National Standards Institute (ANSI). ANSI S1.4–2014 (R2019) Part 1/IEC 61672–1–2013. American National Standard Electroacoustics–Sound Level Meters–Part 1: Specifications (A Nationally Adopted International Standard). New York, NY: ANSI; 2014

Beranek LL. Acoustics. New York, NY: American Institute of Physics; 1986

Gelfand SA. Hearing: An Introduction to Psychological and Physiological Acoustics. 6th ed. Boca Raton, FL: CRC Press; 2018

Hewitt P. Conceptual Physics. Boston, MA: Little, Brown; 1974

Kinsler LE, Frey AR, Coppens AB, Sanders JB. Fundamentals of Acoustics, 3rd ed. New York, NY: Wiley; 1982

Peterson APG, Gross EE. Handbook of Noise Measurement. 7th ed. Concord, MA: General Radio; 1972

Russell DA. The tuning fork: an amazing acoustics apparatus. Acoust Today. 2020; 16(2):48–55

Sears FW, Zemansky MW, Young HD. University Physics. 6th ed. New York, NY: Addison Wesley; 1982

2 Anatomy and Physiology of the Auditory System

General Overview

Hearing and its disorders are intimately intertwined with the anatomy and physiology of the auditory system, which is composed of the ear and its associated neurological pathways. The auditory system is fascinating, but learning about it for the first time means that we must face many new terms, relationships, and concepts. For this reason it is best to begin with a general bird's-eye view of how the ear is set up and how a sound is converted from vibrations in the air to a signal that can be interpreted by the brain. A set of self-explanatory drawings illustrating commonly used anatomical orientations and directions is provided in **Fig. 2.1** for ready reference.

The major parts of the ear are shown in **Fig. 2.3**. One cannot help but notice that the externally visible auricle, or pinna, and the ear canal (external auditory meatus) ending at the eardrum (tympanic membrane) make up only a small part of the overall auditory system. This system is divided into several main sections: The *outer ear* includes the pinna and ear canal. The air-filled cavity behind the eardrum is called the *middle ear,* also known as the tympanic cavity. **Fig. 2.2** shows how the structures of the hearing system are oriented within the head. Notice that the middle ear connects to the pharynx by the Eustachian tube. Medial to the middle ear is the *inner ear.* Three tiny bones (malleus, incus, and stapes), known as the ossicular chain, act as a bridge from the eardrum to the oval window, which is the entrance to the inner ear (**Fig. 2.3**).

The inner ear contains the sensory organs of hearing and balance. Our main interest is with the structures and functions of the hearing mechanism. Structurally, the inner ear is composed of the vestibule, which lies on the medial side of the oval window; the snail-shaped cochlea anteriorly; and the three semicircular canals posteriorly. The entire system may be envisioned as a complex configuration of fluid-filled tunnels or ducts in the temporal bone, which is descriptively called the labyrinth. The labyrinth, which courses through the temporal bone, contains a continuous membranous duct within it, so that the overall system is arranged as a duct within a duct. The outer duct contains one kind of fluid (perilymph) and the inner duct contains another kind of fluid (endolymph). The part of the inner ear concerned with hearing is the cochlea. It contains the organ of Corti, which in turn has hair cells that are the actual sensory receptors for hearing. The balance (vestibular) system is composed of the semicircular canals and two structures contained within the vestibule, called the utricle and saccule.

The sensory receptor cells are in contact with nerve cells (neurons) that make up the eighth cranial (vestibuloacoustic) nerve, which connects the peripheral ear to the central nervous system. The auditory branch of the eighth nerve is often called the auditory or cochlear nerve, and the vestibular branches are frequently referred to as the vestibular nerve. The eighth nerve leaves the inner ear through an opening on the medial side of the temporal bone called the internal auditory meatus (canal), and then enters the brainstem. Here, the auditory portions of the nerve go to the cochlear nuclei and the vestibular parts of the nerve go to the vestibular nuclei.

The hearing process involves the following series of events. Sounds entering the ear set the tympanic membrane into vibration. These vibrations are conveyed by the ossicular chain to the oval window. Here, the vibratory motion of the ossicles is transmitted to the fluids of the cochlea, which in turn stimulate the sensory receptor (hair) cells of the organ of Corti. When the hair cells respond, they activate the neurons of the auditory nerve.

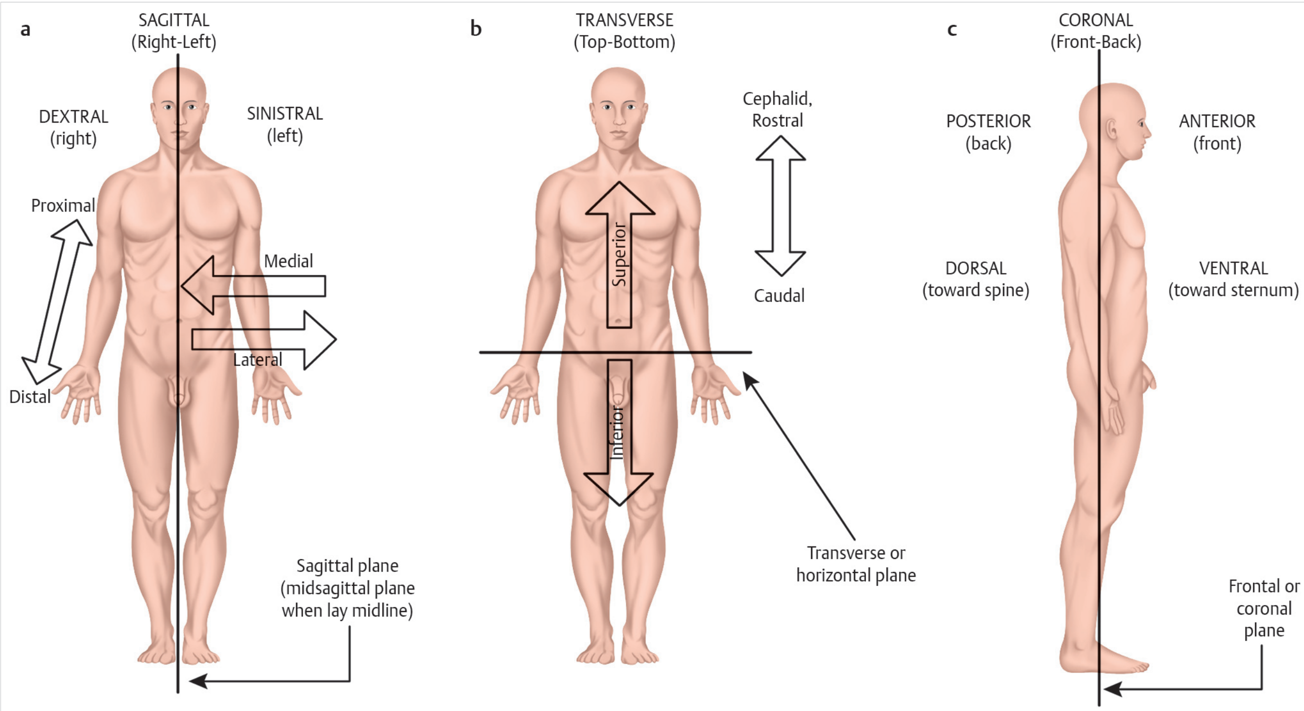


Fig. 2.1 **(a–c)** Commonly encountered anatomical planes, orientations, and directions.

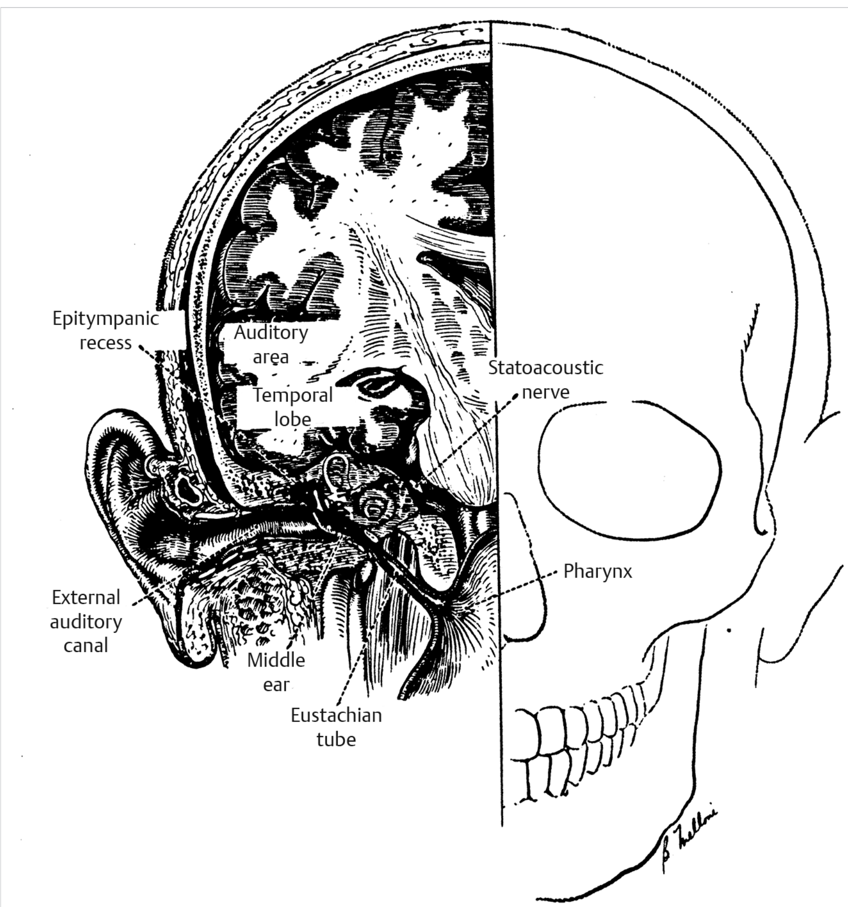


Fig. 2.2 The auditory system in relation to the brain and skull. (Courtesy of Abbott Laboratories.)

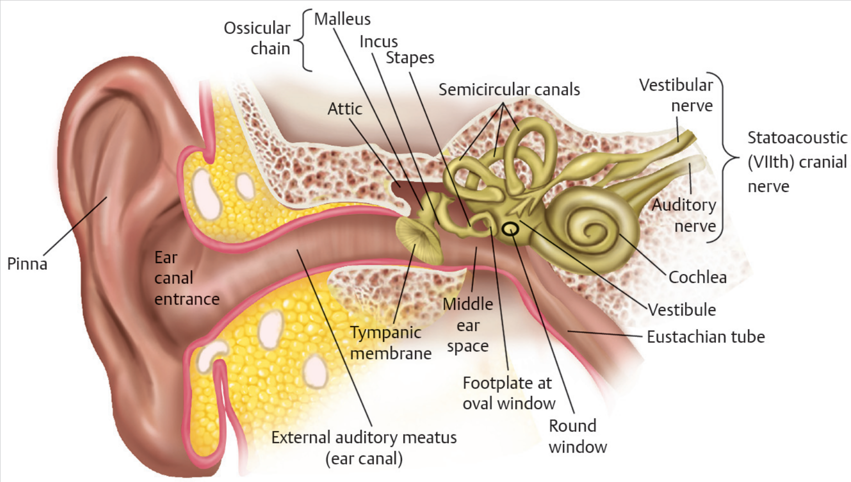


Fig. 2.3 The major parts of the peripheral ear.

The signal is now in the form of a neural code that can be processed by the nervous system.

The outer ear and middle ear are collectively called the **conductive system** because their most apparent function is to bring (conduct) the sound signal from the air to the inner ear. The cochlea and eighth cranial nerve compose the **sensorineural system**, so named because it involves the physiological response to the stimulus, activation of the associated nerve cells, and the encoding of the sensory response into a neural signal. The aspect of the central nervous system that deals with this neurally encoded message is generally called the **central auditory nervous system**.

Temporal Bone

To be meaningful, a study of the ear must begin with a study of the **temporal bone**. Most of the structures that make up the ear are contained within the temporal bone (**Fig. 2.2**). In fact, the walls of these structures and all of the bony aspects of the ear, except for the ossicles, are actually parts of the temporal bone itself. Recall from your anatomy class that the skeleton of the head is composed of 8 cranial bones and 14 facial bones. The right and left temporal bones compose the inferior lateral aspects of the cranium. Beginning posteriorly and moving clockwise, the temporal bone articulates with the occipital bone behind, the parietal bone behind and above, the sphenoid and zygomatic bones to the front, and the mandible anteriorly below. All of these connections, except for the articulation with the mandible, are firmly united, seam-like fibrous junctions called sutures. The articulation with the mandible is via the highly mobile **temporomandibular joint**.

Lateral and medial views of the temporal bone are shown in **Fig. 2.4**. The lateral surface of the bone faces the outside of the head and the medial surface faces the inside of the head. The temporal bone is composed of five sections, including the mastoid, petrous, squamous, and tympanic parts, and the styloid process.

The **squamous part** is a very thin, fan-shaped portion on the lateral aspect of the bone. It articulates with the parietal bone posteriorly and superiorly, and with the sphenoid bone anteriorly. The prominent zygomatic process runs anteriorly to join with the zygomatic bone, forming the zygomatic arch on the medial side of the temporal bone. Just below the base of the zygomatic process is a depression called the **mandibular fossa**, which accepts the condyle of the mandible to form the temporomandibular joint just anterior to the ear canal.

The **petrous part** is pyramid-shaped and medially oriented so that it forms part of the base of the cranium. This extremely hard bone contains the inner ear and the internal auditory meatus through which the eighth cranial nerve travels on its way to the brainstem, so that much of the discussion pertaining to the inner ear is also a discussion of this part of the temporal bone.

The **mastoid part** composes the posterior portion of the temporal bone. It extends posteriorly from the petrous part, below and behind the squamous part. The mastoid articulates with the occipital bone posteriorly and with the parietal bone superiorly. It has an inferiorly oriented, cone-shaped projection below the skull base called **mastoid process**. The mastoid contains an intricate system of interconnecting air cells that vary widely in size, shape, and number. These are connected with an anterosuperior cavity called the **tympanic antrum**, which is located just behind the middle ear cavity. An opening called the **aditus ad antrum** connects the antrum with the attic or upper part of the middle ear cavity. The roof of the antrum (and the middle ear) is composed of a thin bony plate called the **tegmen**

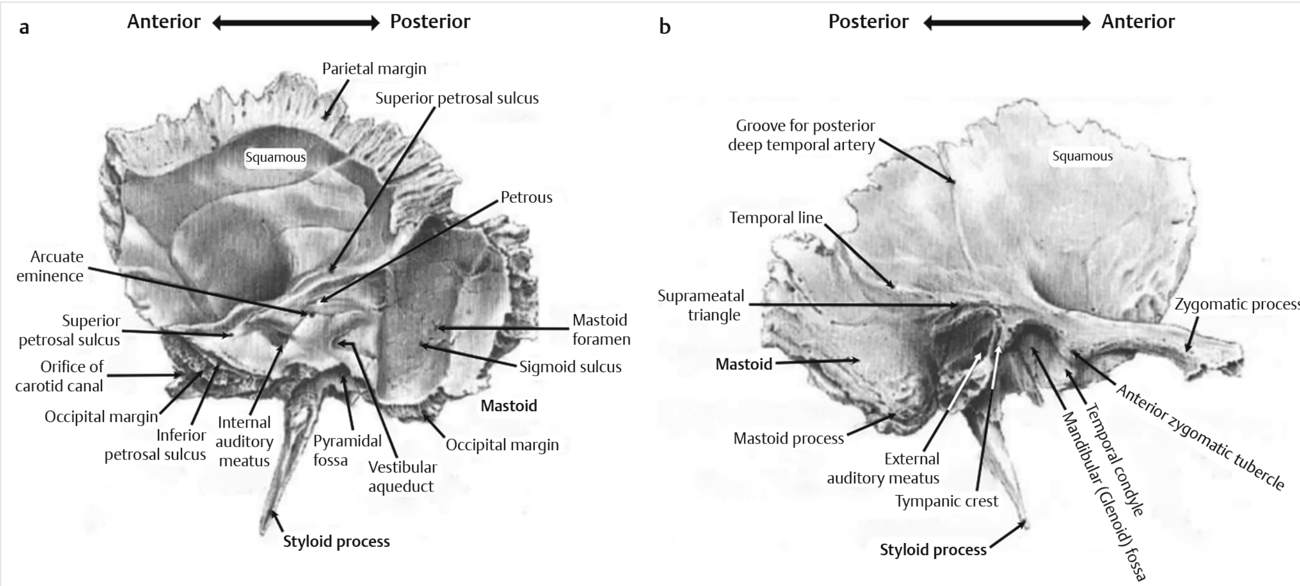


Fig. 2.4 **(a)** Medial and **(b)** lateral views of the right temporal bone. (Adapted from Proctor 1989, with permission.)

tympani, which separates them from the part of the brain cavity known as the middle cranial fossa. Its medial wall separates it from the lateral semicircular canal of the inner ear. The middle ear, antrum, and air cells compose a continuous, air-filled system. Hence, it is not hard to imagine how an untreated middle ear infection can spread to the mastoid air cell system and beyond.

The tympanic part is inferior to the squamous and petrous parts and anterior to the mastoid. The tympanic part forms the inferior and anterior walls of the ear canal, as well as part of its posterior wall.

The **styloid process** is an anteroinferior pillar-like projection from the base of the temporal bone that varies widely in size. It does not contribute to the auditory structures but is of interest to us as the origin of several muscles involved in the speech mechanism.

■ Outer and Middle Ear

The **outer ear** is composed of the pinna and the ear canal, ending at the eardrum. The **tympanic membrane** is generally considered to be part of the **middle ear** system, which includes the middle ear cavity and its contents, and "ends" where the ossicles transmit the signal to the inner ear fluids at the oval window.

Pinna

The externally visible aspect of the ear is an odd-shaped appendage called the **pinna** or **auricle**. The internal structure of the pinna is composed principally of elastic cartilage (except for the earlobe).

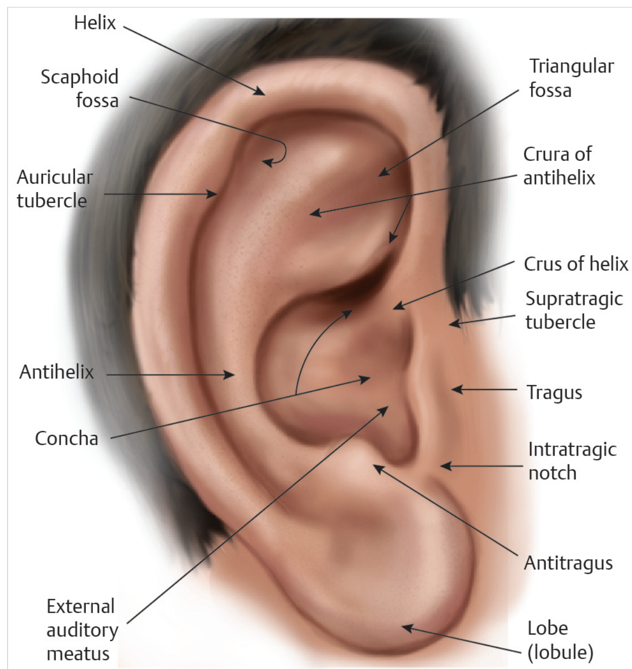


Fig. 2.5 Major landmarks on the pinna.

It also contains some undifferentiated intrinsic muscle tissue, as well as several extrinsic muscles, although these are vestigial structures in humans. However, these muscles are not vestigial in many lower animals that are able to orient their pinnae with respect to the location of a sound source.

The major landmarks of the pinna are highlighted in **Fig. 2.5**. Notice that the pinna is not symmetrical. For example, and most obviously, it has a flap-like extension that angles away from the skull in the backward direction, so that the pinna overlaps the side of the head posteriorly, superiorly, and inferiorly, but not anteriorly. It also has an intricate arrangement of ridges, bumps, and furrows. The entrance to the ear canal is at the bottom of a large, cup-shaped depression called the concha, and is partially covered by a posteriorly directed projection or ridge called the tragus. The ridged rim along most of the perimeter of the pinna is the **helix**. Starting at the very top of the pinna, the helix courses anteriorly and downward, making a hook-like turn in the posterior direction to form a shelf above the concha, which is the **crus of the helix**. Again, beginning at the top of the pinna, the helix courses downward along the posterior perimeter, reaching the **earlobe**, or **lobule**, at the bottom. The **scaphoid fossa** is the furrow just anterior to the helix as it courses down the posterior rim of the pinna. The ridge anterior to the scaphoid fossa is called the **antihelix**. The antihelix runs parallel with the helix and splits superiorly into the two **crura of the antihelix**. The splitting of the antihelix into two crura creates a triangle-shaped depression called the **triangular fossa**. If we follow the antihelix downward, it widens at the bottom to form the **antitragus**, an upward pointing mound that is located posteroinferior to the tragus and superior to the earlobe. The space or angle between the tragus and the antitragus is the **intertragic notch (or incisure)**.

External Auditory Meatus

The ear canal is more formally called the **external auditory meatus (canal)**. On average, it is ~ 9 mm high by 6.5 mm wide, and is roughly 2.5 cm to 3.5 cm long. The ear canal is not quite a straight tube, but has two curves forming a slightly S-like pathway.

These curves usually make it difficult to get an unobstructed view of the eardrum, so that it is usually necessary to straighten the canal before looking inside with an otoscope. An otoscope is the familiar instrument used for examining the ear (**Fig. 2.6**). The ear canal is straightened by gently pulling up and back on the pinna, as illustrated in **Fig. 2.7**.

Fig. 2.6 Example of otoscope used for examining the ear, with reusable and disposable speculum attached. The speculum is the cone-shaped attachment that is inserted into the ear canal. (Photograph courtesy of HEINE Optotechnik.)

The external auditory meatus is lined with tight-fitting skin. However, the outer third of the canal is different from the inner two-thirds in several ways. The underlying material is cartilage in the outer third and bone for the remainder of its length. The bony portion of the canal is derived from (1) the tympanic part of the temporal bone, which forms the floor and anterior wall, as well as the inferoposterior wall; (2) the squamous part, making up the roof and part of the posterior wall; and (3) the condyle of mandible, which contributes to the inferoanterior wall at the temporomandibular joint. The cartilaginous portion contains hairs as well as a plentiful distribution of **sebaceous (oil)** and **ceruminous (wax) glands**, although this is not the case for the bony portion of the canal. Sebaceous glands are also present in the concha. Their secretions serve lubricating and antimicrobial functions, and also help to keep the canal free of debris and even some foreign bodies and insects.

Tympanic Membrane

The external auditory meatus ends at the **tympanic membrane** or **eardrum**, which is tilted at an angle of ~ 55° to the canal. The tympanic membrane is firmly attached to the **tympanic sulcus**, a groove in the bony canal wall, by a ring of fibrocartilaginous connective tissue called the **tympanic annulus** (or **annular ligament**). The ring has a deficiency at the top due to a tiny interruption in the tympanic sulcus known as the **notch of Rivinus**. The eardrum is a smooth, translucent, and sometimes almost transparent membrane with an average thickness of only ~ 0.074 mm. It is slightly taller (~ 0.9 to 1.0 cm) than it is wide (~ 0.8 to 0.9 cm), and it is concave rather than flat. The peak of the cone-like inward displacement is called the **umbo**. Structurally, the eardrum is often described as having three layers, although more correctly there are four of them. The most lateral layer of the tympanic membrane is continuous with the skin of the ear canal, and the most medial layer is continuous with the mucous membrane of the middle ear. Sandwiched between them are two fibrous layers. One of them is composed of radial fibers reminiscent of the spokes of a wheel, and the other layer is made of essentially concentric circular fibers.

The tympanic membrane is connected to the malleus, which is the first of the three middle ear bones. Specifically, a long, lateral process of the malleus called the **manubrium** attaches almost vertically to the eardrum, with its tip at the umbo and continuing upward toward the position of 1 o'clock in the right ear and 11 o'clock in the left ear. This attachment of the manubrium of the malleus forms the **malleal prominence**. Ligamentous bands called the **anterior** and **posterior malleal folds** run from both sides of the malleal prominence to the notch of Rivinus, forming a triangular area between them on the eardrum. The largest part of the tympanic membrane lies *outside* or below the malleal folds, and is called the **pars tensa** ("tense part") because it contains all four layers described above. The superior area of the eardrum between the malleal folds *is missing* the two fibrous layers, and is called the **pars flaccida** ("flaccid area") for this reason. It is also known as **Shrapnell's membrane**.

Fig. 2.7 The ear canal is straightened to facilitate otoscopic inspection by gently pulling up and back on the pinna. (Photograph courtesy of HEINE Optotechnik.)

Fig. 2.8 shows some of the major landmarks that are identified when looking at the eardrum through an otoscope. In addition to the landmarks already mentioned, notice that the light from the otoscope is reflected back in a characteristic way, called the **cone of light** or **light reflex**. This reflection is seen as a bright area on the anteroinferior surface of the eardrum, radiating from the tip of the manubrium to the 5 o'clock position in the right ear and the 7 o'clock position in the left ear. It is often possible to identify one or more middle ear landmarks, especially when viewing relatively transparent eardrums.

Middle Ear

The cavity in the temporal bone behind the tympanic membrane is called the **middle ear**, **tympanum**, or **tympanic cavity**. The posterosuperior portion of the middle ear space is usually viewed as an "attic room" above the main tympanic cavity, and is called the **epitympanic recess** or the **attic**. This space

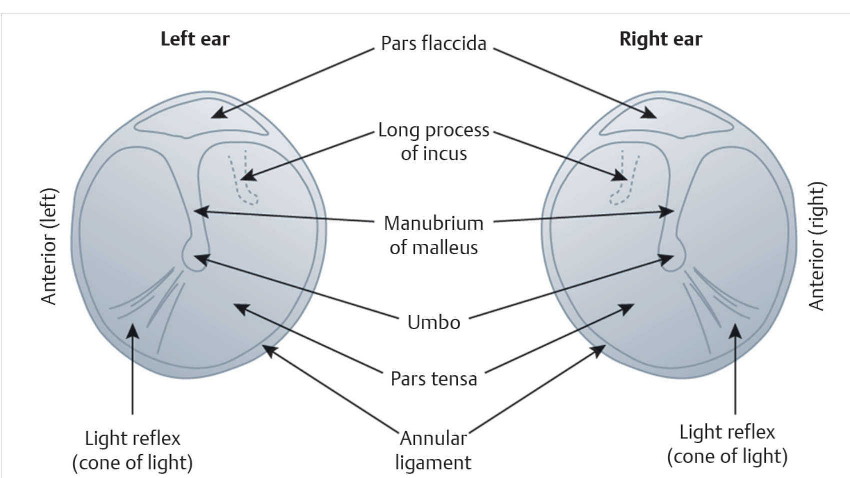


Fig. 2.8 Major otoscopic landmarks of the tympanic membrane.

accommodates the more massive portions of the two larger ossicles, the incus and the malleus. Using a certain amount of artistic license, **Fig. 2.9** depicts the middle ear space as a box-shaped room. Keep in mind that the box is only an analogy used to describe *relative* directions and relationships. In addition, **Fig. 2.10** shows a more realistic drawing of the tympanic cavity and its contents looking backward toward its posterior wall. The tympanic membrane forms the lateral wall, and is shown folded downward to reveal the inside of the "room" in the figure. The middle ear really has irregularly shaped curved surfaces as in the other figures, *not* flat walls with right-angled corners. The floor of the tympanic cavity separates it from the **jugular bulb** below. The ceiling is the **tegmen tympani**, which is the thin bony plate that separates the tympanic cavity from the brain cavity above. Low down on the anterior wall (~ 3 mm up from the floor) is the opening of the **Eustachian tube** (sometimes called the **auditory tube**). The **internal carotid artery canal** is located on the other side of (i.e., anterior to) the anterior wall, just below the Eustachian tube. Just above the Eustachian tube is the **tensor tympani semicanal**, which contains the **tensor tympani muscle**. The tensor tympani semicanal and Eustachian tube are separated by a bony shelf or septum. There is a curved bony projection on the anterior/medial wall that points into the middle ear space, called the **cochleariform process**. The tendon of the tensor tympani muscle bends around the cochleariform process and proceeds in the lateral direction on its way to the malleus (**Fig. 2.10**).

The prominent bulge on the medial wall is the **promontory** of the basal turn of the cochlea. The **oval window** of the cochlea (with its attachment to the stapes) is located posterosuperior to the promontory, and the **round window** of the cochlea is posteroinferior to it. The **facial nerve canal prominence** is situated superior to the oval window.

The posterior wall separates the tympanic cavity from the mastoid. The **aditus ad antrum** is an opening located superiorly on the rear wall, and provides communication between the epitympanic recess of

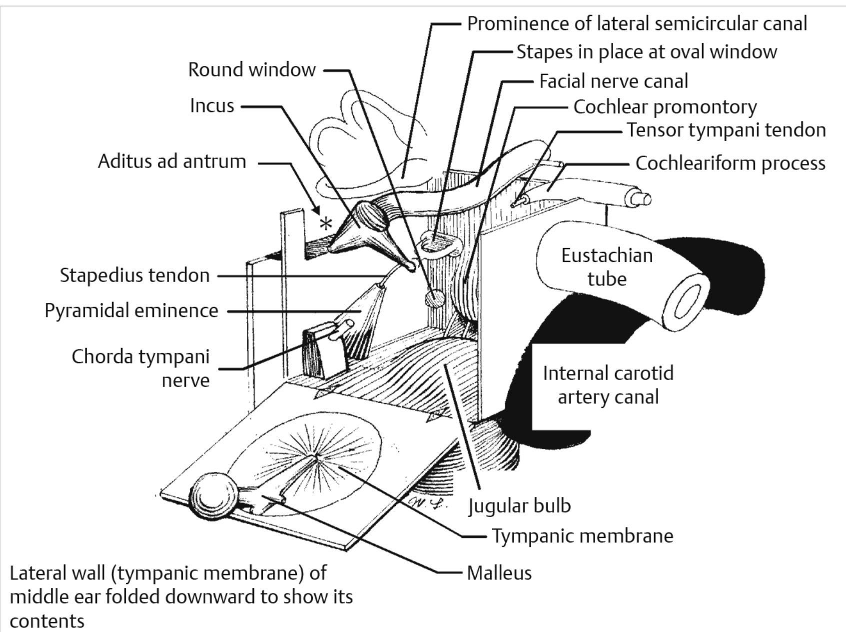


Fig. 2.9 An artist's conceptualization of the middle ear cavity as a room, with the lateral wall (including the eardrum and attached malleus) folded down to reveal the inside. The stapes is shown in place in the oval window. (Adapted from Proctor 1989, with permission.)

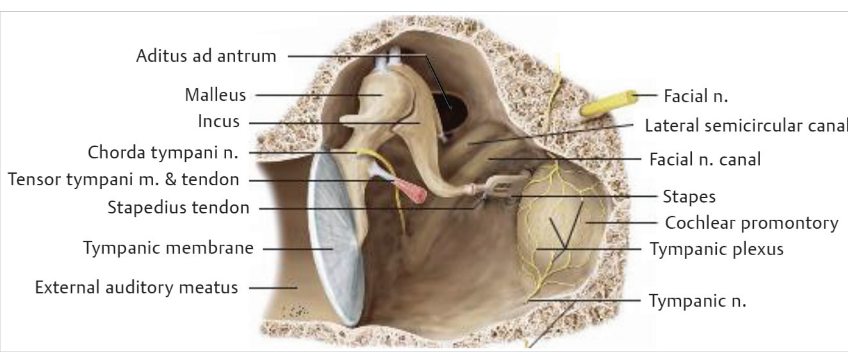


Fig. 2.10 Artist's representation of the middle ear looking backward toward the posterior wall (m., muscle; n., nerve). (Adapted from Schuenke, Schulte, Schumacher, Ross, Lamperti, & Voll 2010; Fig. 9.3b, p. 144, with permission.)

the middle ear cavity and the antrum of the mastoid air cell system. The **pyramidal eminence** or **pyramid** is a prominence on the posterior wall that contains the body of the **stapedius muscle**. The **stapedial tendon** exits from the apex of the pyramid and proceeds to the stapes. The **fossa incudis** is a recess on the posterior wall that accommodates the short process of the incus. The **chorda tympani nerve** is a branch of the **facial (seventh cranial) nerve** that enters the middle ear from an opening laterally at the juncture of the posterior and lateral walls, runs just under the neck of the malleus, and leaves the middle ear cavity via the opening of the **anterior chordal canal (of Huguier)** that is anterior to the tympanic sulcus.

Ossicular Chain

Three tiny bones known as the ossicles or the **ossicular chain** transmit the sound-induced vibrations of the tympanic membrane to the cochlea via the oval window (**Fig. 2.11**). The ossicles are the smallest bones in the body, and include the **malleus**, **incus**, and **stapes**. Instead of being attached to other bones, the ossicular chain is suspended within the middle ear cavity by ligaments and tendons, as well as its attachments to the tympanic membrane and the oval window. The malleus ("hammer" or "mace") is ~ 8 to 9 mm long and weighs ~ 25 mg. Its **manubrium** (handle) is firmly embedded between the fibrous and mucous membrane layers of the eardrum, forming the lateral attachment of the ossicular chain. The neck of the malleus is a narrowing between its manubrium and head. Its lateral process produces a bulge on the eardrum that is often visible otoscopically. There is also an **anterior process** near the junction of the neck and manubrium. The head of the malleus connects with the body of the incus by a diarthrodial or double-saddle joint, the **incudomallear articulation**, such that these two bones move as a unit.

The incus, which is ~ 7 mm long and weighs roughly 30 mg, is commonly called the "anvil" but looks more like a tooth with two roots. The **short process** is posteriorly oriented and is accommodated by the fossa incudis on the back wall of the middle ear. The **long process** descends from the body of the incus, parallel with the manubrium of the malleus, and then hooks medially to end at a rounded nodule called the **lenticular process**. In turn, the lenticular process articulates with the head of the stapes via a true ball-and-socket or enarthrodial joint called the **incudostapedial joint**.

The stapes bears a close resemblance to a "stirrup," which is its common name. Its head is connected via the neck to two strut-like processes called the **anterior** and **posterior crura**, which lead down to the oval shaped **footplate**. The opening between the crura and footplate is called the **obturator foramen**. The stapes weighs only 3 to 4 mg. It is ~ 3.5 mm long, and the footplate has an area of roughly 3.2 mm^2. The footplate is attached to the oval window by the **annular ligament**, forming the medial attachment of the ossicular chain.

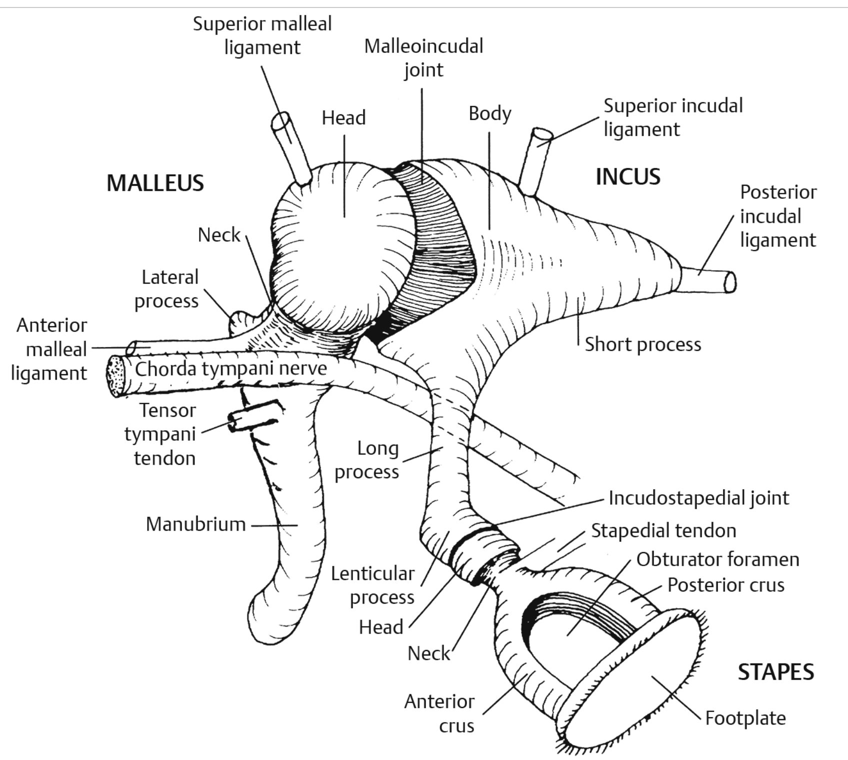


Fig. 2.11 The ossicular chain in place within the middle ear. (Adapted from Tos 1995, with permission.)

In addition to its lateral attachment to the tympanic membrane and its medial attachment at the oval window via the annular ligament, the ossicular chain is also supported by several ligaments and the tendons of the two middle ear muscles. The **superior malleal ligament** runs from the roof (tegmen tympani) of the attic down to the head of the malleus. The **anterior malleal ligament** goes from the anterior tympanic wall to the anterior process of the malleus. The **lateral malleal ligament** extends from the bony margin of the notch of Rivinus to the neck of the malleus. **The posterior incudal ligament** (actually a fold of mucous membrane rather than a ligament) runs from the fossa incudis to the short process of the incus.

Middle Ear Muscles

The **tensor tympani muscle** is innervated by the **trigeminal (fifth cranial) nerve**. It is housed within the tensor tympani semicanal in the anterior middle ear wall superior to the Eustachian tube. This muscle is ~ 25 mm long and arises from the cartilage of the Eustachian tube, the walls of its semicanal, and the segment of sphenoid bone adjacent to the canal. The **tensor tympani tendon** bends around the cochleariform process and proceeds in the lateral direction to insert on the malleus at the top of the manubrium near the neck. Contraction of the tensor tympani pulls the malleus in the anteromedial direction, thereby stiffening the ossicular chain.

The **stapedius** muscle is the smallest skeletal muscle in the body, with an average length of only 6.3 mm. It is contained within the pyramidal eminence of the posterior wall of the tympanic cavity, and is innervated by the **facial (seventh cranial) nerve**. The **stapedius tendon** exists via the apex of the pyramidal eminence and runs anteriorly to insert on the posterior aspect of the neck of the stapes. Contraction of the stapedius pulls the stapes posteriorly. Even though the middle ear muscles pull in more or less opposite directions, they both exert forces that are perpendicular to the normal motion of the ossicles, and their contraction has the effect of stiffening the ossicular chain, thereby reducing the amount of energy that is delivered to the inner ear.

The **acoustic reflex** refers to the reflexive middle ear muscle contraction that occurs in response to high levels of sound stimulation (Gelfand 1984, 2018). In humans this is at least principally a stapedius reflex, while the tensor tympani contracts as part of a startle reaction to very intense sounds, and also in response to certain kinds of nonacoustic stimulation such as an air jet directed at the eye. The acoustic reflex involves both muscles in some animals.

The fundamental aspects of the **acoustic (stapedius) reflex arc** are as follows (Borg 1973; Wiley & Block 1984). The sensory (afferent) pathway proceeds via the auditory nerve to the ventral cochlear nucleus, and then to the superior olivary complex on *both sides* of the brainstem (with the crossover to the opposite side by way of the trapezoid body). The motor (or efferent) pathway is followed bilaterally, and runs from the motor nuclei of the **facial (seventh cranial) nerve** on both sides, and then via the facial nerves to the stapedius muscles. The motor pathway for tensor tympani activation goes from the **trigeminal (fifth cranial) nerve** nuclei to the tensor tympani muscles via the trigeminal nerves. Because contraction of the stapedius and tensor tympani muscles stiffens the middle ear system, the transmission of low frequencies is reduced (Simmons 1959; Møller 1965; Rabinowitz 1976). This change has been observed as a decrease in hearing sensitivity or loudness for low-frequency sounds (Smith 1943; Reger 1960; Morgan & Dirks 1975), although these effects are not always found (Morgan, Dirks, & Kamm 1978).

The purpose of the acoustic reflex is not really known, although several theories have been proposed. The *protection theory* suggests the acoustic reflex protects the inner ear from potentially damaging sound levels. However, it is unlikely that this is its main purpose because the reflex has a delay (latency) that would make it an ineffective protection against sudden sounds, and the contraction adapts over time, limiting its protection against ongoing sounds. The *fixation theory* holds that the middle ear muscles maintain the appropriate positioning and rigidity of ossicles, and the *accommodation theory* states that the muscles modify the characteristics of the conductive system so that the absorption of sound energy is maximized. Simmons (1964) suggested the acoustic reflex improves the audibility of environmental sounds by attenuating internal sounds. This enhancement of signal-to-noise ratio would improve the survival rates for both the fox who is hunting for dinner and the rabbit who is trying to avoid being the main course. Similarly, Borg (1976) proposed that one of the purposes of the reflex is to improve the listener's dynamic range by attenuating low-frequency sounds. Discussions of these and other reflex theories may be found in several sources (Jepsen 1963; Simmons 1964; Borg, Counter, & Rosier 1984).

Eustachian Tube

The **Eustachian (auditory) tube** provides for the aeration and drainage of the middle ear system, and makes it possible for air pressure to be the same on both sides of the eardrum. It runs from the anterior

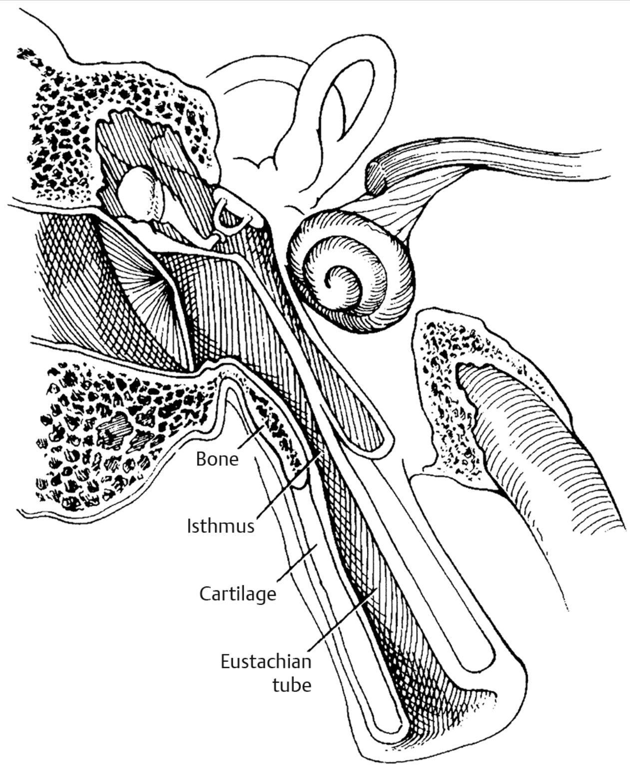


Fig. 2.12 The Eustachian tube in relation to the ear. Note that the bony portion of the tube meets the cartilaginous portion at the isthmus. (Adapted from Hughes 1985, with permission.)

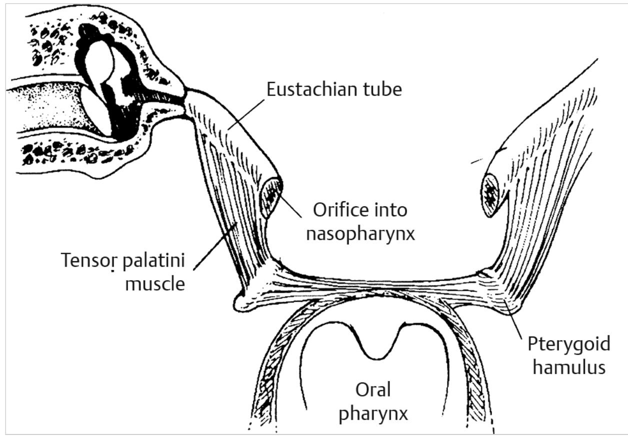


Fig. 2.13 Relationship of the Eustachian tube to the tensor palatini muscle, highlighting the parts of the muscle that arise from the tubal cartilage, hook around the pterygoid hamulus of the sphenoid bone, and insert into the palate. (Adapted from Hughes 1985, with permission.)

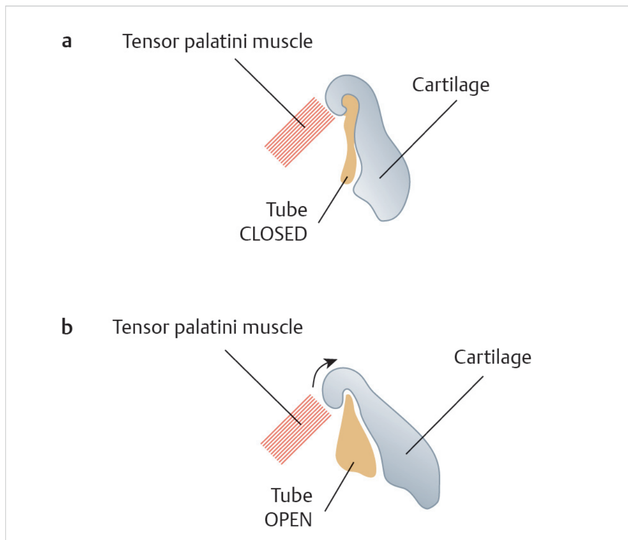


Fig. 2.14 **(a)** The hook-shaped arrangement of the Eustachian tube cartilage keeps the tube in the normally closed state. **(b)** The tube is opened when the cartilage hook is uncurled by action of the tensor palatini muscle.

middle ear wall to the posterior wall of the nasopharynx behind the inferior nasal turbinate, as illustrated in **Fig. 2.12** and **Fig. 2.13**. In adults the Eustachian tube follows a 3.5 to 3.8 cm long course that is directed inferiorly, medially, and anteriorly, tilting downward at an angle of ~ 45°. However, it is important to be aware that the Eustachian tube is almost horizontal in infants and young children (see **Fig. 6.4** in Chapter 6). The first third of the tube beginning at the middle ear is surrounded by bone and the remainder is surrounded by an incomplete ring of elastic cartilage. The meeting point of the **bony** and **cartilaginous portions** is called the **isthmus**. The lumen of the Eustachian tube is narrowest at the isthmus, where it is only 1 to 2 mm across compared with 3 to 6 mm in the remainder of the tube. A prominence on the pharyngeal wall, the **torus tubarius**, is formed by the cartilage of the Eustachian tube and other tissues (e.g., the salpingopalatine, salpingopharyngeus, and tensor palatini muscles).

The cartilaginous portion of the Eustachian tube is arranged as shown in **Fig. 2.14**. Notice that the cartilage hooks around the tube from above, forming the incomplete ring alluded to above. A portion of the tensor palatini muscle is attached to the hooked segment of the cartilage (**Fig. 2.13** and **Fig. 2.14**). At rest, the cartilage keeps the Eustachian tube closed (**Fig. 2.14a**). The lumen of the tube is opened when the cartilage is uncurled due to the pull exerted by the tensor palatini muscle (**Fig. 2.14b**). This occurs during swallowing, yawning, and other actions that cause the tensor palatini muscle to contract. Negative pressure develops in the middle ear (compared with the outside, atmospheric pressure) when this mechanism fails to open the Eustachian tube frequently and effectively. We have all experienced this phenomenon as fullness in the ears that is (hopefully) alleviated when the tube "pops open" due to a swallow, yawn, or some other maneuver. Swelling and/or blockage by mucus due to colds or allergy, and obstruction of the pharyngeal orifice by hypertrophic (enlarged) adenoids, are just a few of the causes of a nonpatent Eustachian tube. The resulting negative pressure within the closed-off middle ear space often precipitates the development of clinically significant middle ear disease.

Functioning of the Conductive Mechanism

Sound entering the outer ear is picked up by the tympanic membrane. The vibrations of the eardrum are transmitted to the ossicular chain, which vibrates essentially in the right-left plane. This vibration is represented as a rocking motion of the stapes footplate in the oval window, as shown in **Fig. 2.15**. The resulting inward and outward displacements of the oval window are thus transmitted to the cochlear fluid. However, instead of just serving as an inert passageway that allows sound to travel to the cochlea from the surrounding air, the conductive mechanism actually modifies sound in several ways that have a direct bearing on how and what we hear.

Head-Related Transfer Function

Sounds reaching the eardrum are affected by the acoustics of the external auditory meatus, which operates as a quarter-wavelength resonator because it is essentially a tube with one open end and one closed end (at the tympanic membrane). Sounds entering the ear will be enhanced if they are close to the resonant frequency range, resulting in a boost in the sound pressure level (SPL) reaching the eardrum, called the **ear canal resonance effect**. To determine this boost in level (or gain) at the eardrum, the sound presented from a loudspeaker is monitored by a microphone outside the patient's ear and also by a special kind of microphone that monitors the sound level inside the subject's ear canal, very close to the eardrum. The difference between these two measurements is the amount of gain provided by the ear canal resonance, shown as a function of frequency in **Fig. 2.16**. On this kind of graph, 0 dB means "unchanged," that is, that the SPL at the eardrum is the same as it is outside the person's ear. Positive values show the amounts of gain provided by the ear canal resonances (negative values mean that the level is lower at the eardrum than outside the ear). We see clearly that the resonance characteristics of the ear canal provide a sound level boost of as much as 15 to 20 dB in the frequency range from roughly 2000 to 5000 Hz. In addition, the middle ear has a resonant region between ~ 800 Hz and ~ 5000 to 6000 Hz (Zwislocki 1975). The resonant responses of the conductive system affect our hearing sensitivity for sounds at different frequencies, as discussed in Chapter 3.

The graph in **Fig. 2.16** is technically called the **head-related transfer function (HRTF)** because it shows how the sound reaching the eardrum is affected by the direction of the sound source relative to the head. In other words, the HRTF shows how the spectrum is changed by the acoustical path from the loudspeaker to the eardrum. Thus, it is sometimes

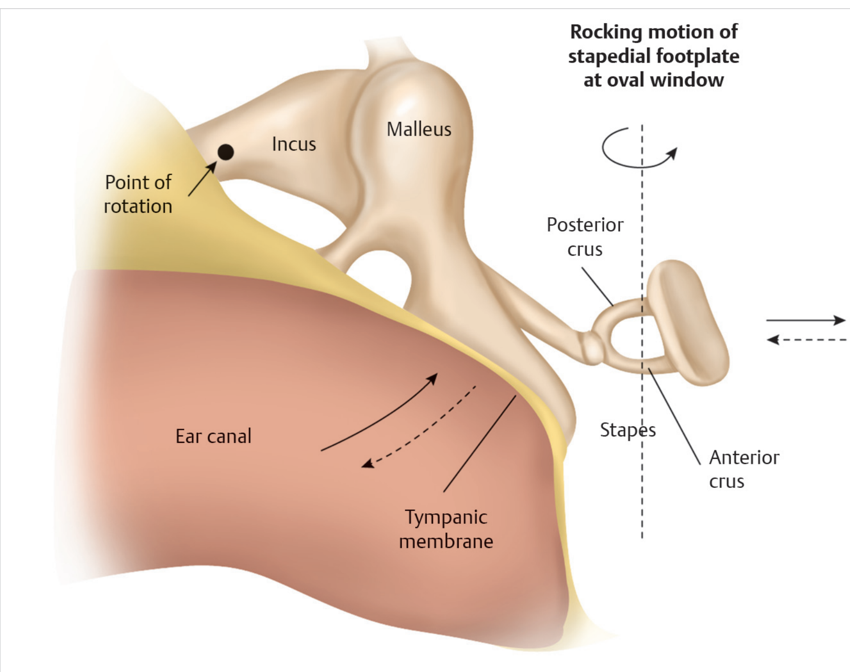


Fig. 2.15 The middle ear transmits the signal to the cochlea via a rocking motion of the stapedial footplate at the oval window. (After Bekesy 1941.)

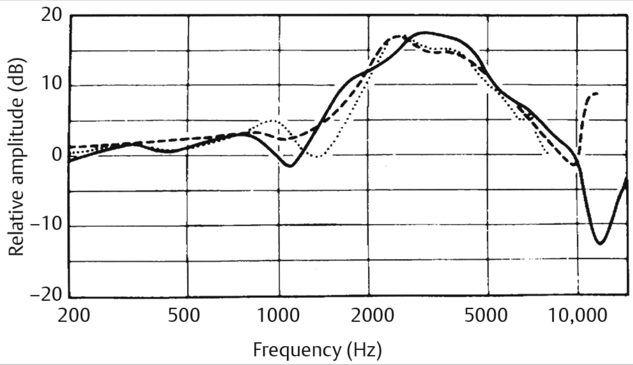


Fig. 2.16 Average head-related transfer functions (sound level at the eardrum compared with outside of the ear) for sounds presented from a loudspeaker directly in front of the subjects. (*Dotted line,* Wiener & Ross 1946; *dashed line,* Shaw 1974; *solid line,* Mehrgardt & Mellert 1977.) (From Mehrgardt & Mellert 1977, with permission of the *Journal of the Acoustical Society of America.)*

also called the **sound field to eardrum transfer function**. The HRTF actually reflects the accumulated effect of all factors that influence the sound on the way from the loudspeaker to the tympanic membrane, including acoustical shadows, reflections, and diffraction due to the head and body, as well as the ear canal resonance. This is why the figure caption specifies the direction of the loudspeaker. These acoustical effects depend on the direction of the sound and are important cues for directional hearing, or the ability to identify the location of a sound source and differences in locations between different sound sources.

Sound direction is described in terms of angles around the head (**Fig. 2.17**). The easiest way to describe the location of a sound source is to give its angle relative to the head (e.g., so many degrees to the right, so many degrees straight up from the front, etc.). The **horizontal plane** refers to horizontal directions around the head toward the right and toward the left. Horizontal directions are usually called angles of **azimuth**. For example:

0° azimuth is *straight ahead* (in front of your nose); 90° azimuth is directly *to the right* (in front of your right ear);

180° azimuth is *straight back* (directly behind your head); and 270° azimuth is directly *to the left* (in front of your left ear).

Thus, an azimuth of 45° means 45° to the right, and 315° azimuth is the same as 45° to the left. By the way, it is perfectly fine to say "45° to the right" and "45° to the left." It is also acceptable (and often confusing) to use positive angles for one direction and negative values for the other, such as + 60° azimuth for 60° to the right, and –60° azimuth for 60° to the left.

Azimuths of 0° and 180° are both dead center between the two ears. The same thing is true for

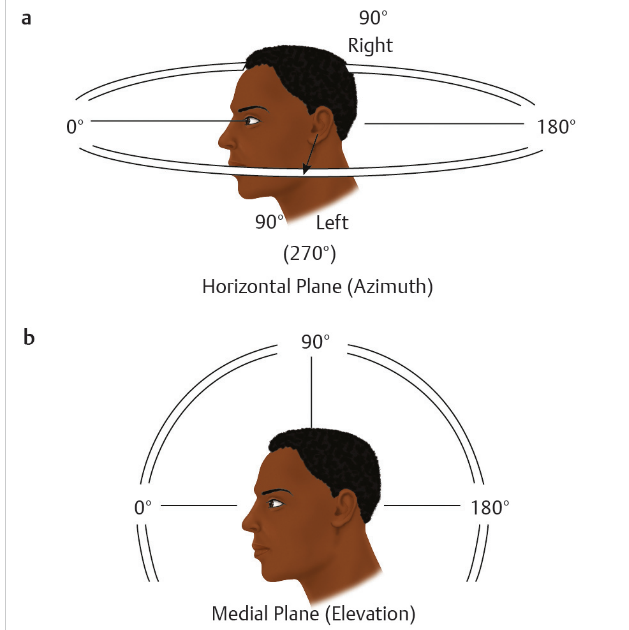


Fig. 2.17 Directions around the head. **(a)** The horizontal plane goes around the head horizontally from right to left, expressed in degrees of azimuth. **(b)** The medial (median) plane goes around the head vertically from front to back, expressed in degrees of elevation.

elevations such as straight up above the center of the head, 45° upward from directly in front, 30° downward from directly in front, and 70° upward from straight back. These are examples of directions in the **medial (or median) plane**. The medial plane is the same as going around the head vertically in the *mid-sagittal anatomical plane*, so that all locations on the medial plane are equidistant from the two ears. In the medial plane, 0° is straight ahead, 90° is straight up, and 180° is straight back.

To appreciate how directional differences are represented acoustically at the ears, let us suppose that a loudspeaker is located at an azimuth of 45° to the right, as depicted in **Fig. 2.18a**. Even though the original sound is the same, it reaches the left ("far") ear differently than it does the right ("near") ear. The far ear is subjected to an acoustical shadow when the head obstructs the path of the sound (**Fig. 2.18b**).

As a result, the sound reaches the far ear at a softer level than it reaches the near ear. This level disadvantage at the far ear is called the head shadow effect. The head shadow occurs for frequencies that can be obstructed by the head. This occurs for sounds with wavelengths that are short compared with the size of the head. Recall that wavelength gets shorter as frequency increases. Hence, the head shadow affects relatively higher frequencies, especially those over ~ 1500 Hz. This level difference

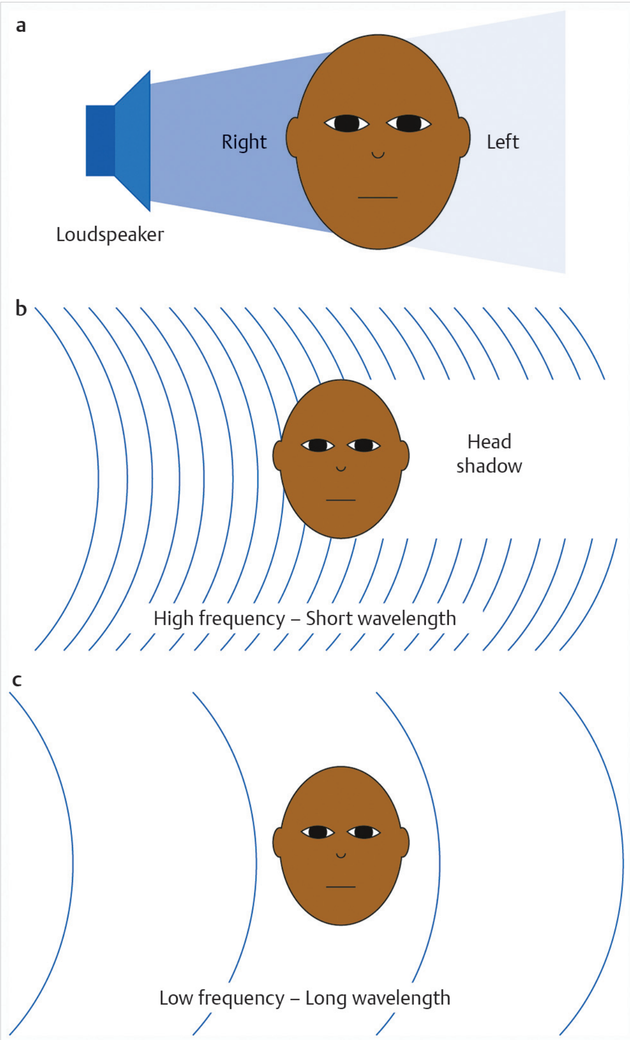


Fig. 2.18 **(a)** Sound coming from a loudspeaker off to the right side arrives differently at the two ears. **(b)** An acoustical shadow (the head shadow effect) occurs for high-frequency sounds because their wavelengths are short compared with the size of the head. **(c)** Because of their relatively longer wavelengths, low-frequency sounds are not subjected to a head shadow because they are able to bend around the head.

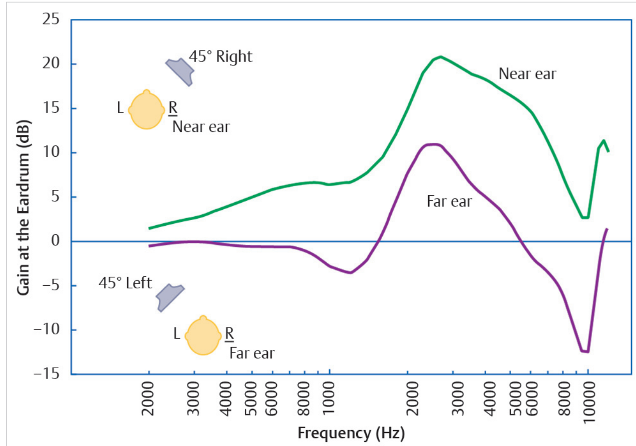


Fig. 2.19 Head-related transfer functions for the right ear (R) when the loudspeaker is located at azimuths of 45° to the right ("near ear" situation) versus 45° to the left ("far ear" situation). (Based on the data of Shaw 1974 and Shaw & Vaillancourt 1985.)

between the ears is called an **inter-ear (interaural) intensity difference**.

The result of these differences at the eardrum can be seen by comparing the **head-related transfer functions** obtained when a loudspeaker is positioned at different locations around the head. **Fig. 2.19** illustrates the **azimuth effect** by showing how the head-related transfer functions are different for the right ear when a loudspeaker is at an azimuth of 45° to the right compared with when is 45° to the left (i.e., 45° versus 315°). The sound level at the right eardrum is greater when the sound comes from 45° to the right ("near ear") compared with when it comes from 45° to the left ("far ear"). Notice that the shape of the curve also changes depending on the azimuth. This graph shows the results at two representative azimuths. Curves obtained from many azimuths all around the head would reveal a continuum of these kinds of differences (Shaw 1974).

Low frequencies have wavelengths that are long relative to the size of the head, so that diffraction can occur. In other words, they are able to bend around the head to the far ear with little if any loss of level (**Fig. 2.18c**). However, the sound will arrive at the near ear earlier than at the far ear, constituting an **inter-ear (interaural) time difference**. Interaural intensity and time differences provide the principal cues needed for directional hearing.

Pinna Effect

What does the pinna do for us? It has long been known that any amplification provided by the pinna is essentially negligible in spite of its funnel-like appearance, and that its main contribution to hearing is in the realm of sound source localization (Bekesy & Rosenblith 1958). (To appreciate the importance of the pinna in directional hearing, one has only to watch a cat orient its pinnae toward a sound source.) The pinna provides sound localization cues because its asymmetrical and irregular shape, ridges, and depressions modify the spectrum of a sound in a way that depends on the direction of the source (Blauert 1983). The simplest example is that sounds coming from the rear are obstructed by the pinna so that some of the high-frequency components of their spectra are attenuated compared with the same sounds arriving from the front. These kinds of pinna-related spectral differences are particularly

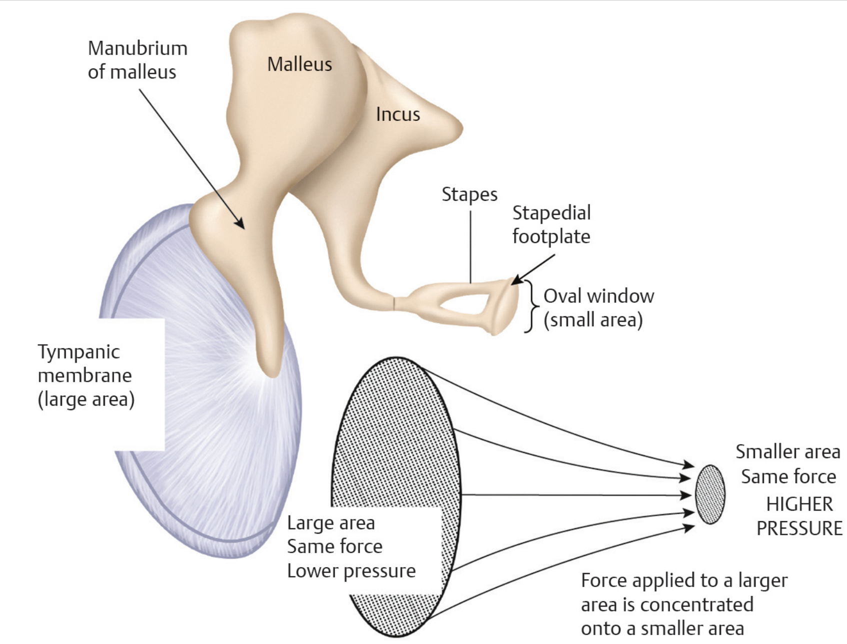


Fig. 2.20 The area advantage involves concentrating the force applied over the tympanic membrane to the smaller area of the oval window.

important when inter-ear sound differences are negligible or absent. This is the case for localizations in the medial plane and/or when trying to localize sounds with just one ear.

The Middle Ear Transformer

The sound signal that reaches the ear in the form of air vibrations must be transmitted to the cochlea, which is a fluid-filled system. The impedance of the cochlear fluids is much greater than the impedance of the air. As a result, most of the sound energy would be reflected back if airborne sound were to impinge directly on the cochlear fluids, and only ~ 0.1% of it would actually be transmitted. This situation is analogous to the reflection of airborne sound energy off the water's surface at the beach, which is why you cannot hear your friends talking when your head is under the water. The middle ear system overcomes this impedance mismatch by acting as a mechanical *transformer* that boosts the original signal so that energy can be efficiently transmitted to the cochlea.

The transformer function of the middle ear is accomplished by the combination of three mechanisms: (1) the area ratio advantage of the eardrum to the oval window, (2) the curved membrane buckling effect of the tympanic membrane, and (3) the lever action of the ossicular chain. The largest contribution comes from the **area advantage**. Here, the force that is exerted over the larger area of the tympanic membrane is transmitted to the smaller area of the oval window, just as the force applied over the head of a thumbtack is concentrated onto the tiny area of its point (**Fig. 2.20**). Pressure is force divided by area ($p = F/A$), so that concentrating the same force from a larger area of the eardrum down to a smaller area of the oval window results in a proportional boost in the pressure at the oval window.

The curved **membrane buckling** mechanism is illustrated in **Fig. 2.21**. The eardrum curves from its rim at both ends to its attachment to the manubrium toward the middle. As a result, eardrum vibration involves greater displacement for the curved membranes and less displacement for the manubrium, which might be envisioned as a buckling effect. A boost in force accompanies the smaller displacement of the manubrium because the product of force and displacement must be the same on both sides of a lever ($F1 \times D1 = F2 \times D2$). The third and smallest contributor to the middle ear transformer mechanism is the **lever action of the ossicular chain**. **Fig. 2.22** shows how the malleus constitutes the longer leg of this lever and the incus is the shorter leg, as well as the axis of rotation.

How much of a boost does this transformer mechanism provide? To answer this question we can plug some representative values into the relationships just discussed. The area of the eardrum is roughly 85 mm^2; however, only about two-thirds of this area vibrates effectively (Bekesy 1960), so that the effective area of the eardrum is something like 56.7 mm^2. The area of the oval window is roughly 3.2 mm^2. Hence, the area ratio is 56.7 to 3.2, or 17.7 to 1. The ossicular lever ratio is ~ 1.3 to 1. So far, the total advantage is $17.7 \times 1.3 = 23$ to 1. In terms of pressure the decibel value of this ratio would be $20 \times \log\ (23/1) = 27$ dB. However, if we add the

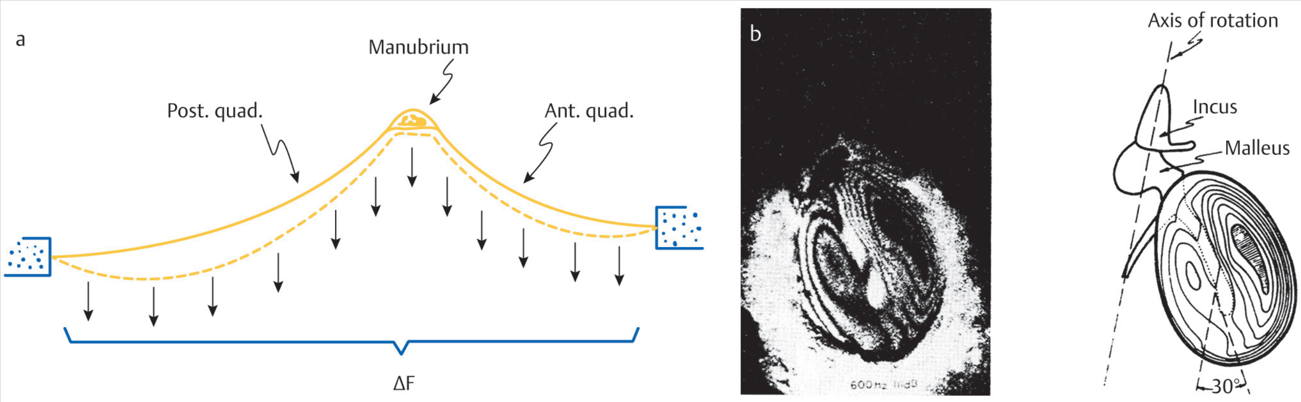


Fig. 2.21 **(a)** The curved membrane buckling principle involves a boost in force (*F*) at the manubrium because it moves with less displacement than the curved eardrum membrane. (Post. Quad., posterior quadrant; Ant. Quad., anterior quadrant). **(b)** Variations in the amount of displacement are shown by concentric contours in this photograph and corresponding drawing of the vibration pattern of the cat's eardrum at 600 Hz. The two areas of concentric contours on both sides of the manubrium show that the eardrum's vibration pattern agrees with the curved membrane principle. (Adapted from Tonndorf & Khanna 1970, with permission of the *Annals of Otology, Rhinology, and Laryngology*.)

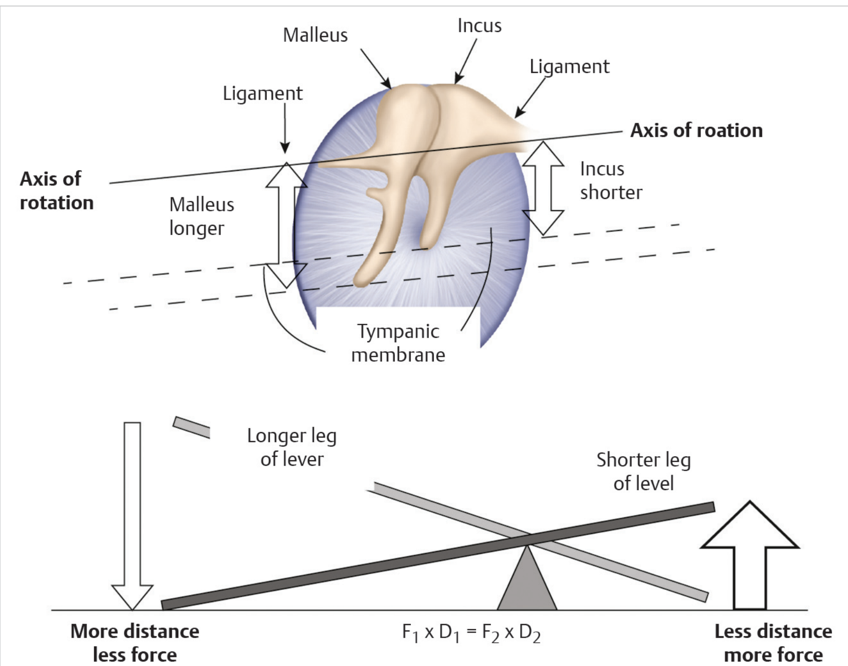


Fig. 2.22 Axis of rotation and relative lengths of the longer (malleus) and shorter (incus) legs of the ossicular lever. (Based in part on Bekesy 1941.)

curved membrane buckling advantage of 2 to 1, the ratio becomes 23 × 2 = 46 to 1. In decibels of pressure, the total advantage now becomes 20 × log (46/1) = 33 dB. This is only an approximation; the actual size of the pressure advantage varies considerably with frequency (Nedzelnitsky 1980).

Inner Ear

The Cochlea

Recall the inner ear is set up like a duct inside of a duct. The outer duct is called the **osseous** or **bony labyrinth** because its walls are made of the surrounding bone. The inside duct is made of membranous materials and is thus called the **membranous labyrinth**. The osseous and membranous labyrinths are represented in Fig. 2.23.

It is useful to "build" a simple model of this apparently complicated system, as in Fig. 2.24a. This model will represent the auditory part of the labyrinth, the cochlea. The outer, bony duct is represented by a steel pipe that is closed at the back end. The inner, membranous duct is represented by a pliable rubber hose (or a long balloon) with a closed-off end that is inserted almost all the way

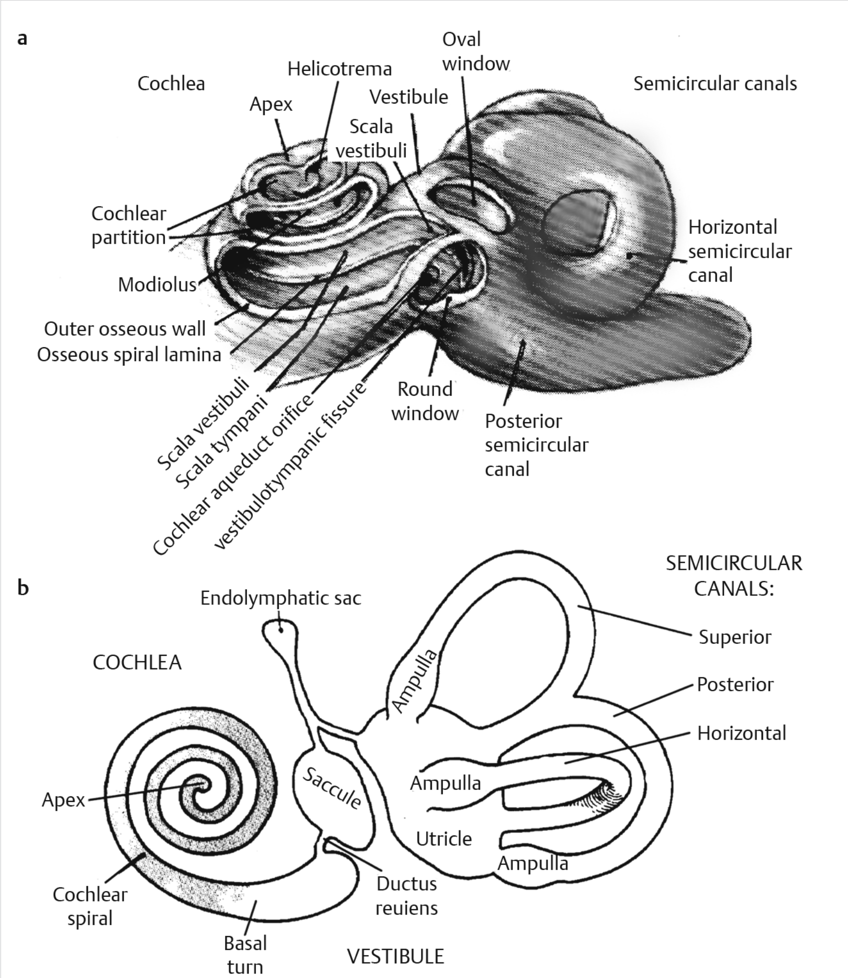


Fig. 2.23 **(a, b)** The major landmarks of the osseous and membranous labyrinths. (Adapted from Proctor 1989, with permission.)

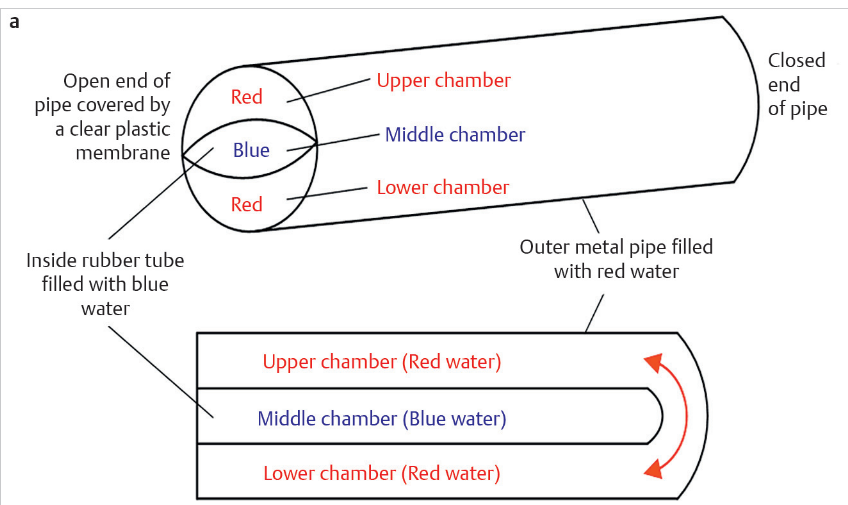


Fig. 2.24 **(a)** A model made of a rubber hose glued within a steel pipe to represent an uncurled cochlea.

into the pipe. The left and right sides of the pliable hose are glued to the inner right and left sides of the pipe, forming three chambers. The middle chamber is completely enclosed by the rubber hose. The upper chamber is above the rubber hose and the lower chamber is below the rubber hose. We now pour water that contains blue food coloring into the rubber hose. Hence, the middle chamber contains blue water. Then we pour water that contains red food coloring into one of the outer chambers, and

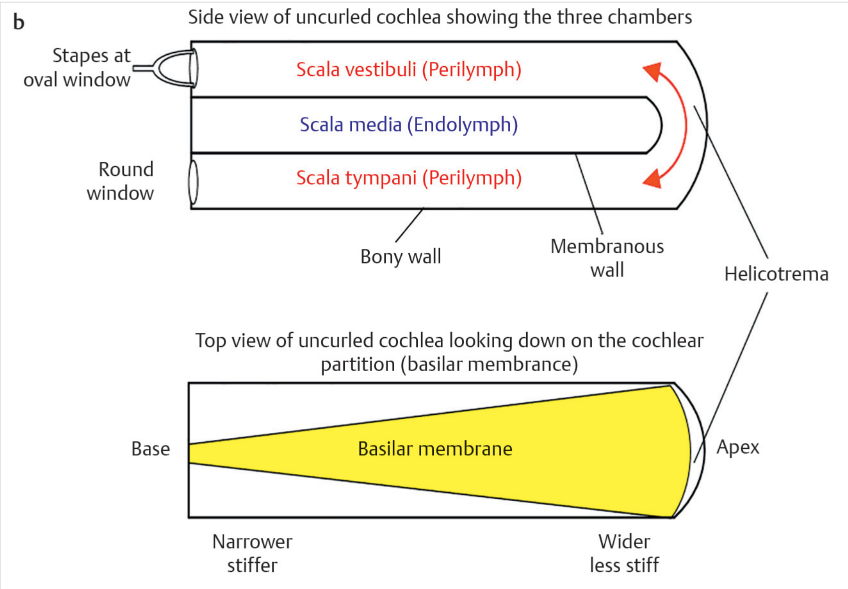


Fig. 2.24 *(Continued)* **(b)** Conceptual drawing of an uncoiled cochlea.

we find that the red water fills both of the outer chambers. This occurs because the rubber hose does not extend all the way to the far end of the pipe, so that the two outer chambers are jointed at that end. We then close off the open ends of the ducts with transparent plastic so that the water does not leak out. We now have a model of a duct inside of a duct. It has two outer chambers filled with red water that are separated by an inner chamber filled with blue water. In addition, the outer chambers are continuous at the back end of the pipe.

Notice the similarities between our pipe model in **Fig. 2.23a** and an artist's representation of a cochlea (**Fig. 2.24b**) that has been "uncurled" from its snail shell configuration. Here the upper duct is called the scala vestibuli and the lower duct is called the scala tympani. They are separated by a middle duct, called the scala media. A liquid called **endolymphatic fluid** or **endolymph** fills the scala media, whereas the two outer ducts are filled with a different liquid called **perilymphatic fluid** or **perilymph**. The two outer ducts meet at the far end of the tube at an opening called the **helicotrema**. The stapes at the oval window is at the base of scala vestibuli, and the round window is at the basal end of the scala tympani. The scala media is separated from the scala vestibuli above it by Reissner's membrane, and from the scala tympani below it by the **basilar membrane**. Notice that the basilar membrane is narrowest at the base and becomes progressively wider toward the apex.

Perilymph is similar to most extracellular fluids in the sense that it contains a large concentration of sodium. On the other hand, endolymph is virtually unique among extracellular fluids because it contains a high concentration of potassium. The same fluids are contained in the balance portion of the inner ear, as discussed below. The membranous labyrinths of the hearing and balance systems are connected by a tiny duct called the **ductus reuniens**, forming one continuous endolymph-filled system. Another conduit, called the **endolymphatic duct**, leads from the membranous labyrinth in the vestibule to the **endolymphatic** sac partly located in a niche in the petrous part of the temporal bone and between the layers of dura in the posterior cranial fossa. In addition, the **cochlear aqueduct** leads from an opening in the scala tympani to the subarachnoid space.

The cochlear duct is ~ 35 mm long and is coiled in the form of a cone-shaped spiral staircase. The resulting arrangement looks like a 5 mm high snail shell that is 9 mm wide at its base and tapers toward the apex (**Fig. 2.23**). The superstructure of the spiral duct is a bony shelf called the **osseous spiral lamina** that makes ~$2^3/_4$ turns around a central core called the **modiolus**, as shown in **Fig. 2.25**. **Fig. 2.25a** also shows that the medial side of the membranous duct (scala media) is attached to the lateral lip of the bony shelf, and follows it from base to apex. (Inside the cochlea, "medial" means toward the center core of the spiral, and "lateral" means toward the outer perimeter of the spiral, or away from the center core.) **Fig. 2.25b** shows how the cochlea would appear if cut down the center and opened to reveal the inside arrangement. Here we see several cross-sectional views of the cochlear duct going around the modiolus. The osseous spiral lamina forms a shelf that divides each section of the duct into an upper and lower part. The section of each duct above the shelf is the scala vestibuli and the section below the shelf is the scala tympani. The membranous duct that is