

The Science of Musical Sound

Revised Edition

John R. Pierce



W. H. Freeman and Company
New York

*To Max Mathews, whose Music V and whose kind and
patient counsel started many things in many places.*

Library of Congress Cataloging-in-Publication Data

Pierce, John Robinson, 1910—

The science of musical sound—rev. ed.

Bibliography: p.

Includes index.

1. Music—Acoustics and physics. 2. Music—
Psychology. 3. Sound. 1. Title.

ML3807.P5 1983 781'.22 82-21427

ISBN 07167-6005-3

Copyright © 1983, 1992 by Scientific American Books

No part of this book may be reproduced by any mechanical, photographic,
or electronic process, or in the form of a phonographic recording, nor may
it be stored in a retrieval system, transmitted, or otherwise copied for
public or private use, without written permission from the publisher.

Printed in the United States of America

1 2 3 4 5 6 7 8 9 0 VB 9 9 8 7 6 5 4 3 2 1

Contents

Preface ix

- 1 Sound, Music, and Computers 1
- 2 Periodicity, Pitch, and Waves 14
- 3 Sine Waves and Resonance 38
- 4 Scales and Beats 64
- 5 Consonance and Dissonance 76
- 6 Consonance, Dissonance, and Harmony 87
- 7 Ears to Hear With 102
- 8 Power and Loudness 115
- 9 Masking 130
- 10 Other Phenomena of Hearing 140
- 11 Architectural Acoustics 150
- 12 Sound Reproduction 167
- 13 Musical Instruments, Analysis, and Synthesis 180
- 14 Perception, Illusion, and Effect 208

APPENDICES

- A Terminology 223
- B Mathematical Notation 225

C	Physical Quantities and Units	227
D	Mathematics and Waves	228
E	Reflection of Waves	234
F	Digital Generation of Sound	239
G	The MIDI Standard	248
H	MAX	252
I	Bibliography	257
	Sources of Illustrations	261
	Index	265

Preface

The first edition of this book, published in 1983, was written in a different world. Then, the computer generation of musical sound was a lusty but tiny infant. Commercial synthesizers were chiefly analog devices, costly and tricky. Digitally generated sound, now familiar through inexpensive digital keyboards, was limited by cost and capability.

Today analog sound is obsolete. Audio technology *is* digital technology. Digital keyboards are more common than pianos and are more suited to our limitations of space and money. Composers produce movie and television scores by digital means, without the intervention of instrumentalists.

Musical sound is musical sound, whether it is generated by acoustical instruments or by digital hardware, and musical sound is what this book is about. *The Science of Musical Sound* describes the physical and mathematical aspects of sound waves that underlie our experience of music as well as the psychoacoustics of musical sound — the relation of the physical aspects of sounds to their perceptual features.

The historical and scientific basis of this book isn't all that different from the first edition, for science and sound don't change that rapidly. Technology does; new technology and new work have added to our knowledge, and I have sometimes amplified and sometimes corrected what I said in the first edition. There are other changes in the book. All chapters have been updated but the most heavily reworked are Chapter 1 on the role of computers in music, Chapter 12 dealing with sound reproduction, and Chapter 13 on synthesized sound and equipment for producing it. I have added an appendix on MIDI, that surprising standard through which

any commercial keyboard can operate any commercial digital synthesizer, and which in general allows interconnection of digital sound gear of various manufacturers. Andrew Schloss, who urged the republication of this book, suggested this addition and drafted the appendix. It was revised with his aid and that of David Jaffe. Further, David Zicarelli has supplied an appendix on MAX, a popular programming language widely used in the control of MIDI-compatible synthesizers.

I have drastically revised, pruned, and added to a bibliographical appendix. I have cited books and pertinent recordings on compact discs of both digital sound examples and of a very few examples of computer music that make important points.

Sound examples are essential to our understanding of sound. Today, a fair amount of computer music has appeared on compact discs, through Wergo and other publishers. And, there are a few recordings of sound examples which are referred to in the bibliography. Beyond this, many commercial synthesizers are available, and many of these can be used in producing the sine waves and combinations of sine waves described in various chapters. Thus, diligent readers can find sound examples on various discs or, better yet, synthesize them on their own.

The preface to the first edition told how I used the fifth Marconi International Fellowship award in preparing the first edition and of my gratitude to Gioia Marconi Braga, Marconi's daughter, for the opportunity the award honoring her father gave me. It told of my indebtedness to Jean-Claude Risset, to Elizabeth Cohen, and to many at Stanford's CCRMA (pronounced *karma*, the Center for Computer Research in Music and Acoustics) for their work in preparing sound examples. It told of people who, in my division at Bell Laboratories, had inspired my interest in sound — of E. E. David, Jr., and Max Mathews, whose early work, culminating in his Music V program, launched computer music to the world in 1957. Today both Max Mathews and I are professors in the music department of Stanford University, invited to join CCRMA by its director, John Chowning. His invention of fm synthesis made possible the first reasonably priced digital keyboard (Yamaha's DX7).

In the preface to the first edition I mentioned the happy month that I spent at Pierre Boulez's IRCAM (Institute for Research and Coordination of Acoustics and Music) in Paris in 1979. I acknowledged the help and inspiration of many: including the late Gerald Strang, Manfred Schroeder, and Earl Schubert, Jr. And of Gerard Piel, Linda Chaput, and others at Scientific American Library who had made the book possible.

Several colleagues at CCRMA and elsewhere have been helpful in preparing this revised edition. I have already mentioned the contributions of Andrew Schloss, David Jaffe, and David Zicarelli. I have consulted

frequently with Earl Schubert. Jay Kadis, CCRMA's audio engineer, was very helpful in connection with Chapter 12, and so was James A. Moorer of Sonic Solutions. Above all, John Chowning, director, and Patte Wood, administrative director, have been essential to this book — as to everything else that happens at CCRMA.

I also wish to thank Linda Chaput, president of W. H. Freeman, for making this new edition possible, and Christine Hastings and others at Freeman for their patient and diligent help in the revision of the manuscript and in seeing it through to publication.

John R. Pierce

This book is about musical sound. Perhaps we should say *acoustics*, a word that entered the English language in the seventeenth century, the century in which Galileo wrote about musical sound.

Today papers in the *Journal of the Acoustical Society of America* are divided into many categories, some of which have little relation to music, such as "Underwater Sound," or "Ultrasonics, Quantum Acoustics, and Physical Effects of Sound." "Music and Musical Instruments" is a category, but much that is essential to an understanding of musical phenomena appears under "Psychological Acoustics" and "Physiological Acoustics." In today's world, knowledge has become so divided and extended that it sometimes seems hard to make sense of it.

In the Greek and Roman worlds music, including whatever was known about musical acoustics, held a high place in science and philosophy. In the liberal arts of the Middle Ages, music was a part of the higher *quadrivium*, along with arithmetic, geometry, and astronomy. The place of music in the liberal arts was above that of grammar, rhetoric, and logic, which constituted the lower *trivium* that dealt with words rather than numbers.

As time passed, music's status became more complicated. The romantics tended to associate music with the grandiose sentiments rather than the grand scientific insights of their age. Still, through the nineteenth century, scientists studied music and musical sound with insight as well as aesthetic appreciation. In 1862 Hermann von Helmholtz, physiologist, physicist, great scientist on a grand scale, published *On the Sensations of Tone as a Physiological Basis for the Theory of Music*. Not a word in the

title about acoustics, though we would say that the book is about musical acoustics and psychoacoustics and, in addition, about music in general.

In this century some musicians have looked to science and technology for new directions in music without concentrating on the word *acoustics*. These have included Edgard Varèse and Hermann Scherchen. This has not been the chief current of musical thought, nor has music been a part of the mainstream of science. The greatest influence of science on music has come through the development of means for recording and reproducing the sounds of music played on conventional musical instruments. The phonograph, with its later electronic advances, and radio revolutionized the role of music in our lives as radically as photography, motion pictures, and television have changed our world of visual experience. Today the computer and digital technology in general are working fantastic changes in the recording and transmission of sound, and in the generation of musical sounds.

This book is indeed about acoustics—both the physical acoustics pertinent to the understanding of conventional musical instruments and the sounds they produce, and the psychoacoustics that helps us to understand the perception of musical sounds. But it is about acoustics in relation to music and musical ideas.

The changes that computers and their descendants will continue to work in music will come partly through fresh insights of musicians who work with new sounds. But computers have opened up new ways of analyzing and experimenting with sounds, and new ways of investigation the response of human beings to sound, including musical sounds. Today we know far more about sounds and their perception than we did in the pre-electronic era. And we will know more in the future.

We can be sure from past experience that new sounds and new understanding of sound will affect the course of music profoundly. Better sounds have always produced different music.

Musical instruments have improved greatly in range and quality in the past few centuries, and certainly up to the beginning of this century. In part, this improvement resulted from (1) the development of better, more easily playable instruments, especially brass instruments with valves and woodwinds with better key mechanisms; (2) the increasing skill of instrumentalists; and (3) an expansion of the range of musical sound, as composers and performers developed and exploited new effects and new idioms.

Whether or not we wish to call such change progress, it brought an expansion in the variety of orchestral sound. Think for a moment of the sounds of Bach, Mozart, Wagner, Debussy, and Stravinsky. As successive generations of composers expanded into new territories of concept, orga-



Figure 1-1 Krummhorns, a Renaissance group of double-reed instruments.

nization and sound itself, the music that they produced was different in style and sound. They did not confront the past on its own ground.

In our century, electronic sounds in general have had a profound effect on some nonelectronic music. When Edgard Varèse wrote *Déserts* in 1954 for taped sound and orchestra, he was proud that he had provided such continuity of sound quality that it is hard to detect transitions from tape to orchestra. Some of Krzysztof Penderecki's music of the 1960s for conventional orchestra deliberately imitates "electronic" sound quality, as does some orchestral writing of Yannis Xenakis a little earlier. In such works, written at a time before synthesized and computer-produced



Figure 1-2 A modern B-flat bass clarinet.

sounds escaped from their early “electronic” timbre, we hear the orchestra refining sounds whose somewhat harsh qualities many listeners found objectionable.

The influence of electronic music on some composers has been more subtle. György Ligeti, who had worked with Karlheinz Stockhausen at the West German radio electronic studios in Cologne from 1958 to 1960, abandoned the limited and difficult electronic means of tone production available at that time. Nevertheless his music for voices and conventional instruments shows that he is acutely aware of the subtle qualities of electronic sounds and of the musical value of the sophisticated control of sound quality.



Figure 1-3 Edgard Varèse (left).

Computers and Music

When, in 1957, Max Mathews first used a digital computer to produce complex musical sounds, it seemed that this could have a liberating effect on composers. In principle, a computer can produce *any* sound. Its potential is limited only by the composer’s imagination. But of this Milton Babbitt said, “It’s like a grand piano in the hands of a group of savages. You know that wonderful sounds can come from it, but will they?” The computer will produce novel and fine music, but only through human skill and effort.

Today the ubiquitous personal computer can be adapted to generate musical sounds. Inexpensive digital keyboards are more widespread than pianos, and a seemingly endless variety of commercial digital hardware is available for producing, modifying, and analyzing musical sounds. Composers of many different styles of music are making themselves heard in homes, in concerts, and through TV and film scores.

Does the future of musical sound lie in digital synthesis? Will the computer have any influence on other aspects of music, for example form or organization? Some have thought it might.

In 1957 Lejarian A. Hiller, Jr., and his collaborator L. M. Isaacson took their inspiration from Johann Joseph Fux (1660–1741), who codified rules to describe the stylistic practices of earlier contrapuntalists, most notably Palestrina. Hiller and Isaacson programmed a computer to make random choices constrained by some of Fux’s rules for first-species counterpoint,

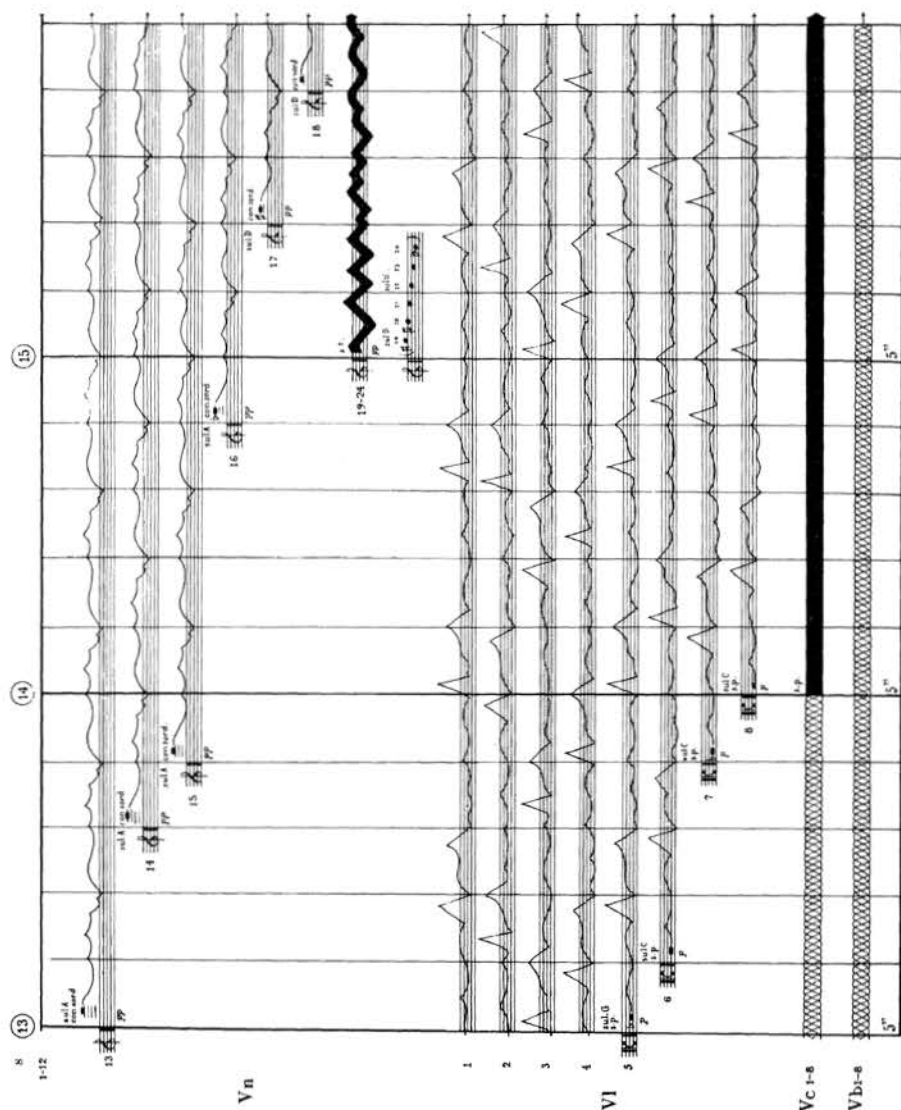


Figure 1-4 A page from the score of Krzysztof Penderecki's *Polymorphia*.



Figure 1-5 Milton Babbitt.

which sets notes of three countermelodies against each note of the original melody. A little of this music, called the *Iliac Suite*, sounds pleasant, but it wanders, and so does the listener's attention. Traditionally, the rules of counterpoint are not intended to tell you what to do, but what *not* to do. The chance element used in generating the *Iliac Suite* provided some surprise but no overall sense of direction.

A lack of musical purpose may be inherent in machine composition. Rules in music are not canned algorithms that we can use in making a



Figure 1-6 Lejaron A. Hiller.

computer solve cut-and-dried problems over and over again. What, indeed, is the place of computers in music?

Hiller and others have pursued the idea of the computer as a composer, or at least, as a tool in manipulating musical material. I will say little about the computer as a composer's aid, for that is a subject unrelated to understanding and generating musical sounds.

We may note, however, that some help that the composer needs can be and is supplied by a computer. One of the most useful tasks that a computer can do is to produce musical scores of high quality in a day when even clear hand copies of music have become excessively expensive. Leland Smith's pioneer SCORE program is now available for IBM (or IBM-compatible) personal computers. It produces scores of excellent, publishable quality. It is easy to make changes and to extract instrumental parts from a full orchestral score. The input to SCORE is the computer keyboard. Many later programs have been developed for producing musical scores on both Macintosh and IBM computers, including Finale, Professional Composer, and NoteWriter. Some allow playing on a pianolike keyboard as an input.

The role of the computer in composition goes beyond the production of a neat final score. Computers are used to store and manipulate musical materials, including lists of notes or their equivalent.

I think that the chief challenge of the computer lies in another direction, that of new sounds and their use. In the past, many composers have responded to the challenge of creating new tone colors. Harry Partch invented both a new scale and an entire orchestra of new instruments to play his music, but Partch's instruments were difficult to build and are not commercially available. In their search for new timbres, some composers have evoked strange sounds from conventional instruments. Perhaps the strangest was the sound of a violin burning on a New York stage, an event staged by Lamont Young and Charlotte Moorman.

Two early-twentieth-century analog instruments, the Theremin and the Ondes Martenot, were recognized as unique musical resources by a number of composers who wrote idiomatically for them. We have noted that electronic analog (as opposed to digital) synthesizers appeared in the mid-twentieth century. They played an appreciable role in music through the 1960s. Robert Moog's analog synthesizers had a distinct musical impact, for example, through Walter Carlos' *Switched On Bach*.

Electronics was an essential part of *musique concrète* of Pierre Schaeffer and others at the Studio d'Essai of the French radio system. Analog synthesis was pursued in the West German radio studios at Cologne by Karlheinz Stockhausen and others.

Analog electronics tended to be expensive and not to stay in adjust-



Figure 1-7 Herbert A. Deutch, with the earliest prototype of the Moog synthesizer, which he helped invent.

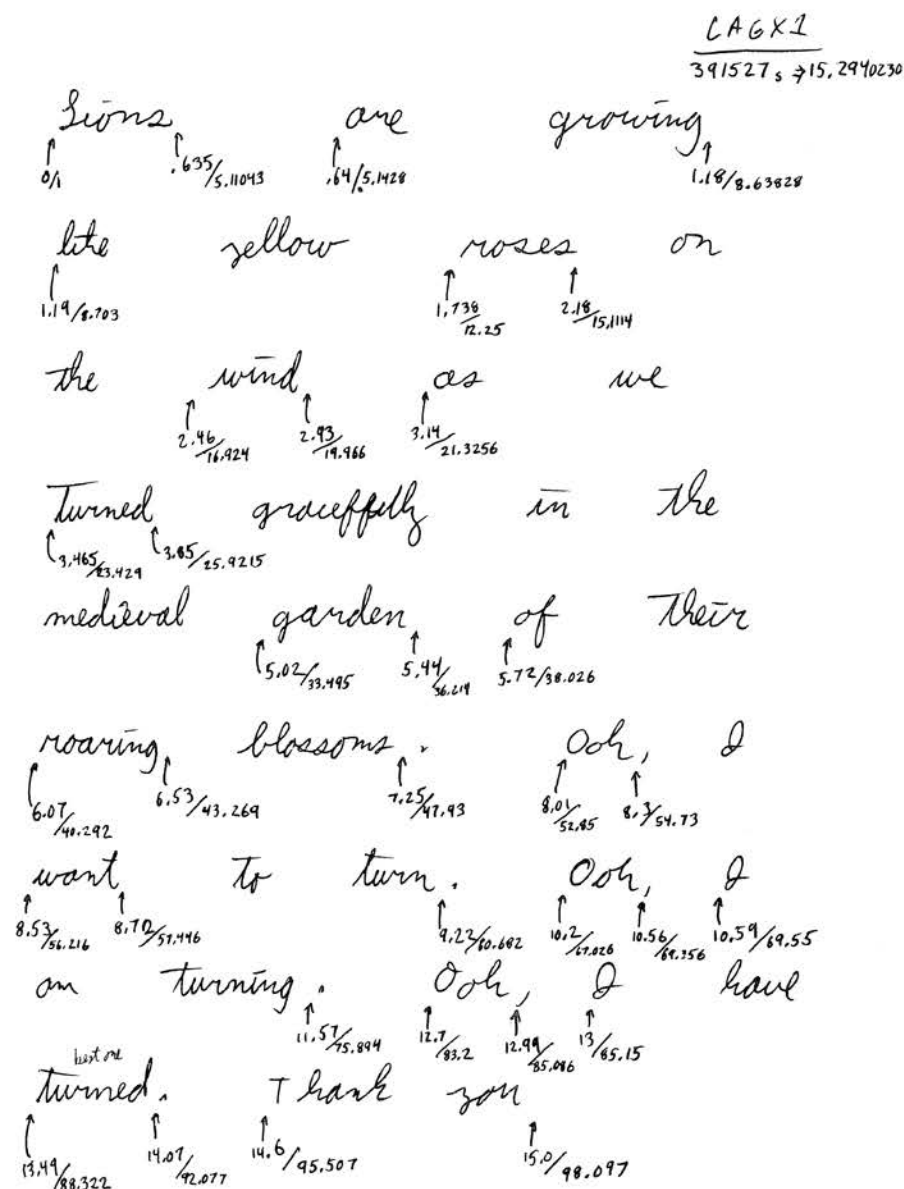
ment. Therefore, analog synthesis did not have a musical impact comparable to that which digital synthesis has had.

New Technology and Computer Music

The computer offers a wide range of sounds, along with the means for controlling them very accurately. The challenge is how to master a constantly changing medium of unlimited acoustic potential, and how to find aesthetic reasons for realizing these new capabilities.

In the early days of computer music, composers encountered a number of problems. There were no instruments available. Composers had to create their own instruments as computer software. And they had to play them; no performers were available either. The composer had to supply all input through a typewriterlike keyboard. All these factors proved awkward in performances. Once a composition was completed, who wanted to sit in an auditorium and listen to music coming from loudspeakers? An audience could not even be sure when to clap unless the composition gave a clue, or unless the house lights came up. Was there an alternative to the concert? Only a few commercial recordings were made, and none had a wide distribution.

Some of the troubles with concerts were overcome in various ways. One solution was to couple recorded sounds with projected images, as in Andy Moor's *Lions Are Growing*, a setting of a poem by Richard Brauti-

Figure 1-8 A working computer score for Andy Moor's *Lions are Growing*.

gan. In this effective piece composed in 1978, a computer-processed voice speaks, sings one line of music and chords, and roars, accompanied by appropriate slides. This made computer music into a component of a multimedia event, like later digitally synthesized film and TV scores. Today, computer film and TV scores are no more and no less recorded and replayed than are scores played by musicians on conventional instruments.

Another concert alternative has been to couple recorded synthesized sound with a singer or instrumentalist. This can be very successful, but it has not been the only resource. In this day of digital keyboards (Figure 1-10), parts or all of a score of digital sound can be evoked using a conventional keyboard. Through the world-standard MIDI interface, a standard commercial keyboard can control commercial digital synthesis hardware. Or a keyboard player can supplement or interact with or control and modify the process of digital sound synthesis.

Max Mathews's Radio Baton, a sort of computerized musical drum, offers another solution. In all modes of operation the player strikes or

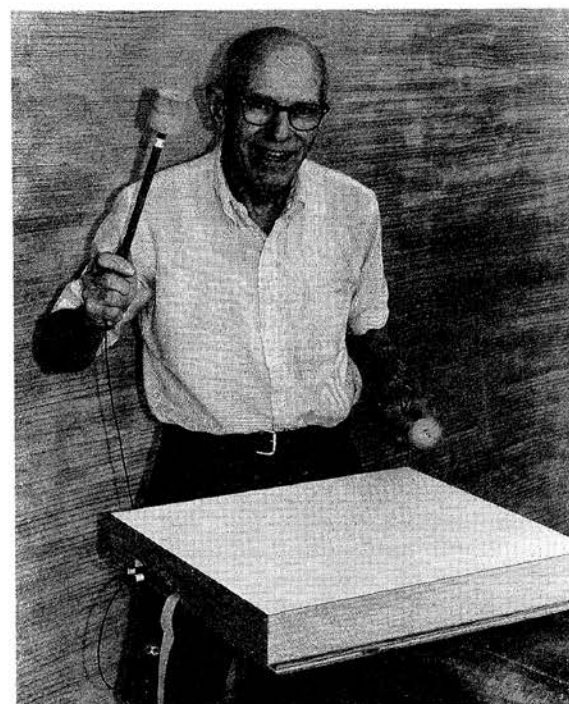


Figure 1-9 Max Mathews and his Radio Baton. Among other things, the Radio Baton enables the performer to beat out the rhythm, loudness, and sound balance of a piece whose notation is stored in computer memory.

strokes a surface with one or more drumsticks. The velocity of striking can control loudness; the position can control timbre in one direction and pitch in another. The successive pitches and durations can also be stored in the computer. Rate of striking or position of stroking can control tempo, and position of striking or stroking can change sound quality and instrumental balance. In this way the Radio Baton brings the performer into the realm of computer-generated sound, but the skills required are perhaps nearer to those of a conductor than of a traditional percussionist. A number of similar devices have appeared.

Today, digital synthesis of sound is used by composers all over the world, in universities, in conservatories, in computer-music institutions, and in commercial music. The MIT press publishes a quarterly *Computer Music Journal*. The worldwide Computer Music Association holds an annual International Computer Music Conference, as well as other meetings, and issues a publication called ARRAY. There are commercial publications, including *Keyboard* and *Electronic Musician*. When I look back over more than thirty years to the time when Max Mathews generated the first computer music piece in 1957 I am struck by obstacles overcome and progress made.

Many talented young composers use computers and digital synthesis hardware as experimental tools to study the intricacies of musical sounds. Some will produce digitally synthesized music. Others will compose music for conventional instruments. All will be influenced, all will learn many new and useful things. And so, I hope, will the reader of this book.

Although this book discusses computers and the digital analysis and synthesis of musical sounds, it is really about the well-known aspects of all aspects of musical sounds, about pitch, scales, consonance, harmony, and timbre, and about some less-known aspects of perception. We can't have a useful understanding of musical sounds without considering these aspects. We will start with periodicity, pitch, and waves.



Figure 1-10 The Yamaha SY99, a late model digital keyboard. The earlier Yamaha DX7, which reached the market in 1983, was the first completely digital keyboard synthesizer, and the first such synthesizer that sold at an affordable price (around \$2,000). Early “digital” keyboard devices sold for about ten times this price. The success of the DX7 was due partly to Japanese persistence and ingenuity, partly to the use of special integrated circuit chips, and partly to the use of fm (frequency modulation) synthesis, an invention of John Chowning. Now discontinued, the DX7 was a landmark that ushered in a new era. (Photo courtesy of Yamaha Corp. of America)

2 Periodicity, Pitch, and Waves

Although wind instruments have been known for nearly five thousand years, and harps for almost as long, sounds of a definite pitch are not necessary to music. The earliest musical instruments that archeologists have found in Egypt are clappers. Perhaps song accompanied their rhythmic beat, but the music may have been largely rhythmical. A knowledgeable friend of mine tells me that, in very primitive music, the chief interval used is the fifth (seven semitones), though sometimes an indefinite musical third (four semitones) is also used.

Rhythm by itself can make music. In our time, Carlos Chávez composed a toccata for percussion alone. A siren is heard in Edgard Varèse's *Ionisation*, but that fine work achieves its effect chiefly by rhythm and *timbre* (sound quality).

Pitch and Periodicity

What *is* pitch? Psychologists insist that pitch is a name for our subjective experience of periodic waveforms, rather than a physical property of the sound wave that reaches our ears. Loosely, we can use the word *pitch* to denote the shrillness of a sound. In this sense, the hissing sound *s* has a higher pitch than the shushing sound *sh*. In this chapter, pitch is considered to be a definite quality related to musical tones, such as those produced by the violin, the clarinet, the tuba, the piano, or the human voice. We hear such sounds as having definite pitches that correspond to particular notes of the musical scale. The present *musical* standard for

“concert pitch” is that the A above middle C sounds at 440 vibrations each second. More on pitch can be found on pages 36–37.

Sounds that have a definite, unambiguous pitch are called *periodic*, because something happens over and over again at a constant rate. Galileo found by accident that he could produce a sound having definite pitch by scraping a brass plate with a sharp iron chisel. The tiny parallel and equidistant ridges left on the brass were a permanent witness to the vibrations of the screeching chisel that had engraved them. An old book on musical acoustics relates that Galileo also produced a pitched sound by rubbing a knife rapidly around the edge of a milled coin. You can try this by scratching the edge of a quarter with your fingernail. The sound produced as your nail encounters the ridges around the edge of the coin does have some pitch. The faster you scratch the coin, the higher the pitch.

The siren provides a clear illustration of the periodicity related to musical pitch. The very siren that Varèse used in *Ionisation* stood in his New York studio. When I turned the crank faster, the pitch of the sound that the siren produced rose. Why was this? In order to understand, we must examine the mechanism of the siren, which was invented by Charles

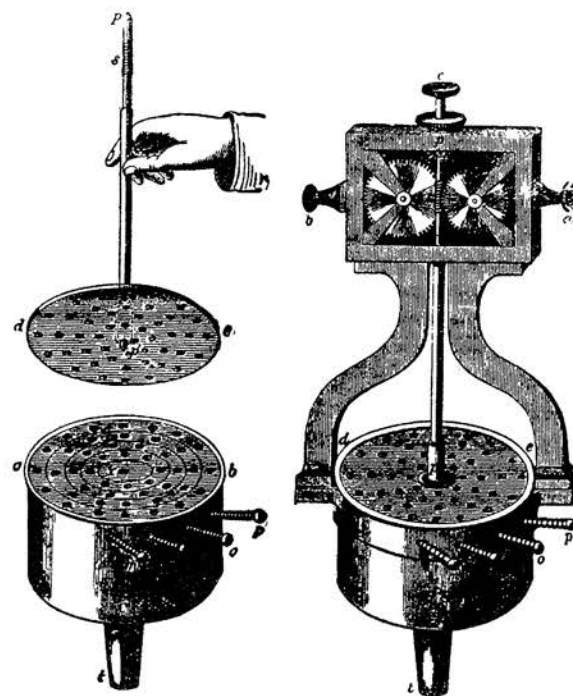


Figure 2-1 De la Tour's siren.

Cagniard de la Tour in 1819. Reduced to its essential parts, a siren consists of a rotating disk pierced by a number of equally spaced holes near its circumference, and a nozzle that directs a jet of air through the holes as they pass, as shown in the diagram in Figure 2-2. Such a siren emits a puff of air each time a hole on the disk passes the nozzle.

The graph in Figure 2-3 represents the succession of puffs of air from the siren. As time passes (moving from left to right), there are intervals when no air passes through a hole (the low, flat parts of the curve) and shorter intervals when puffs of air pass through a hole, shown by asterisks where the curve rises and falls. The succession of puffs of air that the siren produces is periodic: That is, successive puffs are produced at equal intervals, T seconds apart. The time T between successive puffs is called the *period*. The number of puffs in a standard unit of time (usually one second) is the *frequency*.

The periodic pulses of air that the siren emits set up a periodic disturbance called a *sound wave* that travels through the air. When it reaches our ears, we hear this periodic wave as a sound with a definite musical pitch. When the siren produces 440 pulses per second, we hear the A above middle C (in concert pitch). If there are 220 pulses per second, the sound is an octave lower. If there are 880 pulses per second, the sound is an octave higher. Figure 2-4 shows the frequency of various notes of the

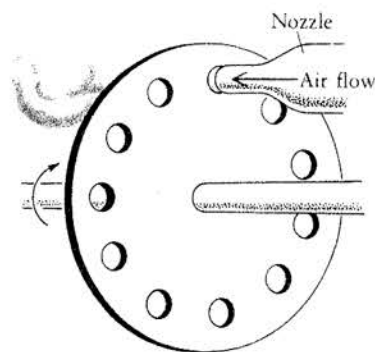


Figure 2-2 A simple siren. Compressed air from a nozzle passes through a circle of equally spaced holes in a rotating disk. As puffs of air pass through the holes, they excite a periodic vibration in the air. The number of puffs per second is the number of revolutions per second times the number of holes in the circle. This disk has 11 holes. If it revolves 40 times a second, it will produce 440 pulses a second, which corresponds to the pitch A above middle C.

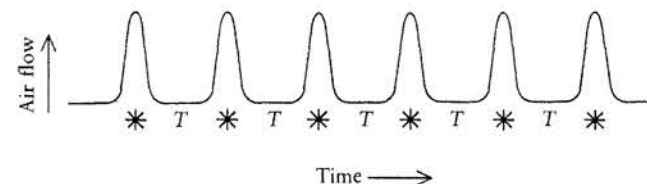


Figure 2-3 Periodic puffs of air produced by a siren. The time, T , between one puff and the next (indicated by asterisks) is called the period.

piano keyboard, together with the ranges of pitch, or *compasses*, of several other musical instruments.

Only sources of sound that are periodic have a clear, definite, unarguable pitch. In this chapter, and in several that follow, only periodic sounds and their pitches are considered, for such sounds form the basis of scales and traditional harmony.

Frequency and Pitch

We might think that the relation between frequency and musical pitch was discovered by means of the siren. The number of pulses of air that the siren produces in a second is the number of holes around the disk times the number of times that the disk revolves each second. We can drive the disk at a high speed by turning a crank attached to a train of gears. If we measure the number of times per second that we turn the crank, we can calculate the number of times per second that the disk rotates, and hence the frequency of the pulses at each pitch that the siren produces.

In fact, the relation between frequency and pitch was discovered much more slowly and much less directly, long before the siren was invented. The pitch of musical sounds that we now know to be periodic was a central aspect of music long before frequency had been related clearly to pitch. Many early cultures used pitch in an orderly and effective way. Like other peoples, the Greeks must have noticed from the earliest times that plucked strings vibrate. Various Greek philosophers associated high notes with swift motion and low notes with slow motion, though they gave no exact relation of motion to pitch. The new discovery that the Greeks made and passed on to posterity is the wonderful numerical relation between the lengths of strings and musical intervals. This discovery is commonly attributed to Pythagoras (c. 500 B.C.)

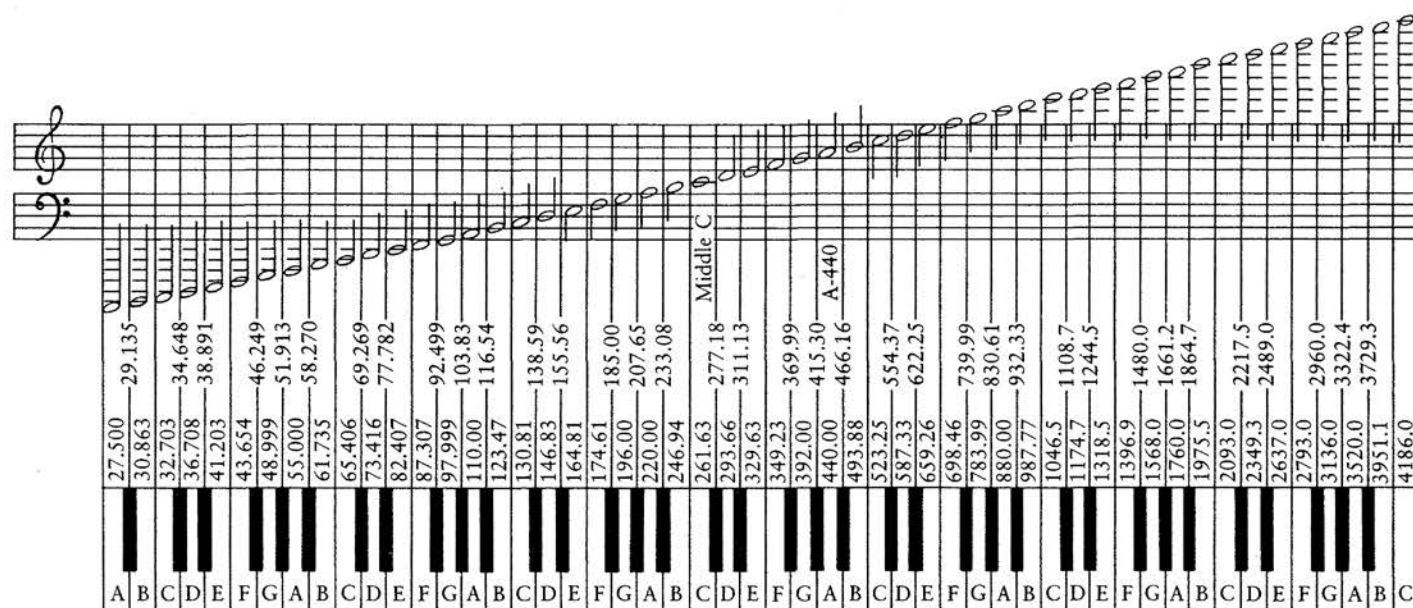


Figure 2-4 The pitch of periodic musical sounds is determined by their frequency, that is, by the number of periods per second. This diagram relates the notes of the musical scale, the positions of the keys of the piano, and the ranges of various musical instruments to the corresponding frequencies. Some periodic sounds have frequencies above or below the range of the piano keyboard. For such sounds, changes in frequency don't correspond to clear musical intervals, though the sensation of pitch does go up and down with frequency. Such sounds don't have a clear or useful *musical* pitch.

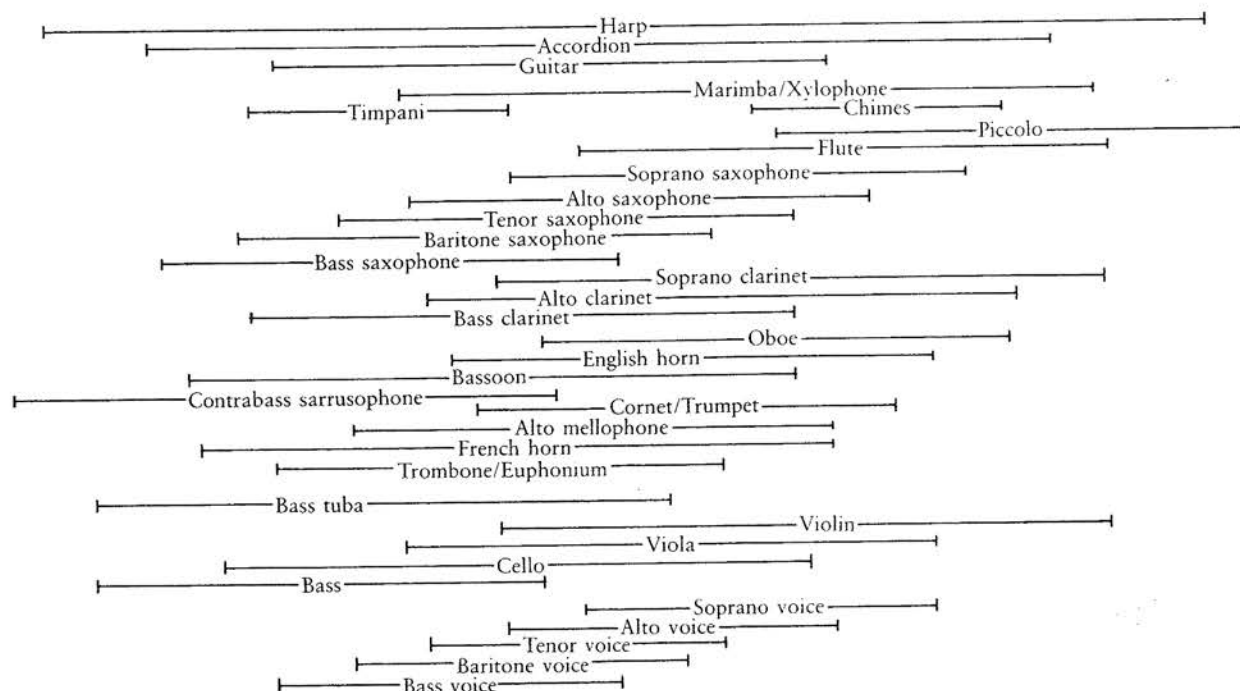




Figure 2-5 Greek vase painting of a citharist. The cithara was sacred to Apollo.

Imagine a stretched wire of length L , as shown at the top of the first part of Figure 2-6. If the wire is plucked, it will emit a sound with a definite pitch, say, middle C. If the tension of the wire is kept constant, but the length is shortened by placing a solid wedge somewhere along the wire, the pitch of the plucked wire is higher. The figure shows the relation between the lengths of the plucked wire and the pitches produced. For example, if the wire is shortened to two-thirds of its original length, it will sound the note that is a *fifth* (seven semitones) above the original note.

The Greeks had a rather mystical regard for number and proportion. They were gratified to find a relation between the ratios of the whole numbers and familiar musical intervals. They sought similar harmonious relations for the proportion of buildings. This was characteristic of their thought. Plato, for example, identified the five regular polyhedra (see Figure 2-7) with the four elements and the universe (tetrahedron, fire; cube, earth; octahedron, air; icosahedron, water; dodecahedron, the universe).

To more modern minds the relation between the lengths of strings and musical intervals is empirical, and people sought some physical explanation for it. From the Renaissance on, scientists wanted to get behind the regularities of complex phenomena and find their explanation in simple terms. In *Dialogues Concerning Two New Sciences*, published in 1636,

RAILOS

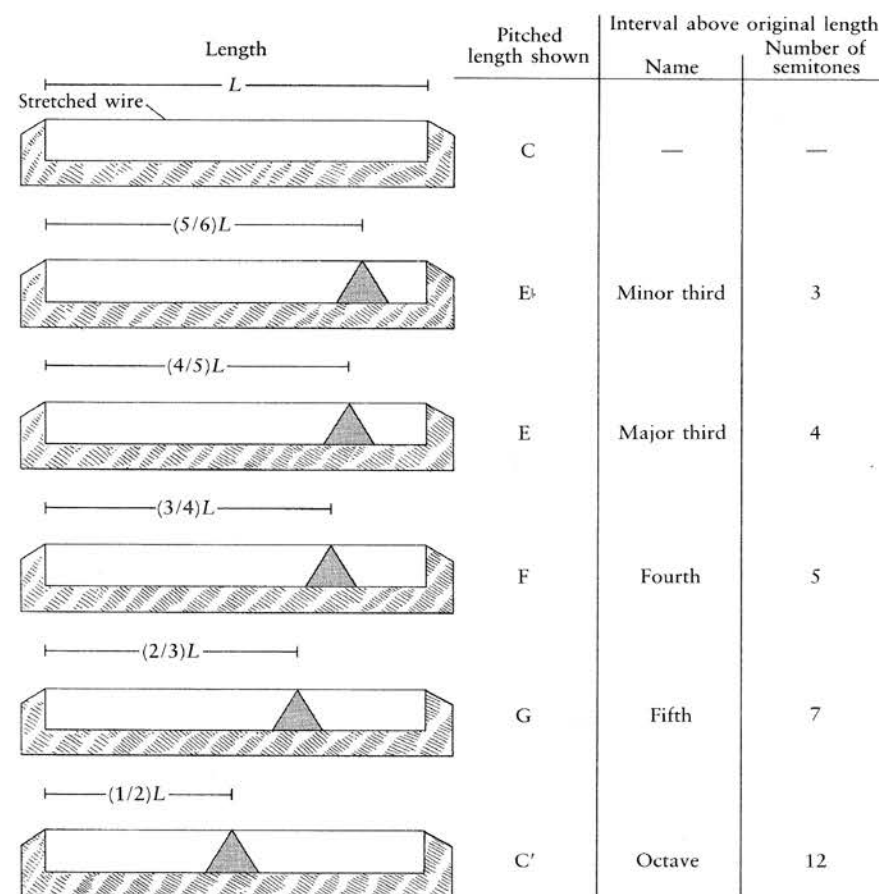


Figure 2-6 A stretched wire vibrates with a particular frequency and gives a tone of a particular pitch. Here we assume this pitch to be middle C. If we keep the tension the same, but use a wedge to reduce the vibrating length to $5/6$ of the original length, the frequency increases by $6/5$, and the pitch goes up a minor third, to E \flat (E-flat). Other fractional reductions in the length of the string result in other musical intervals: $4/5$, major third, E; $3/4$, fourth, F; $2/3$, fifth, G; $1/2$, octave C.

Galileo clearly explained the relation between pitch and the frequency of vibration of a string, but he wrote only of the relative numbers of vibrations per second corresponding to various musical intervals.

In *Harmonic Universelle*, also published in 1636, the French cleric, philosopher, and mathematician Marin Mersenne related pitch to the

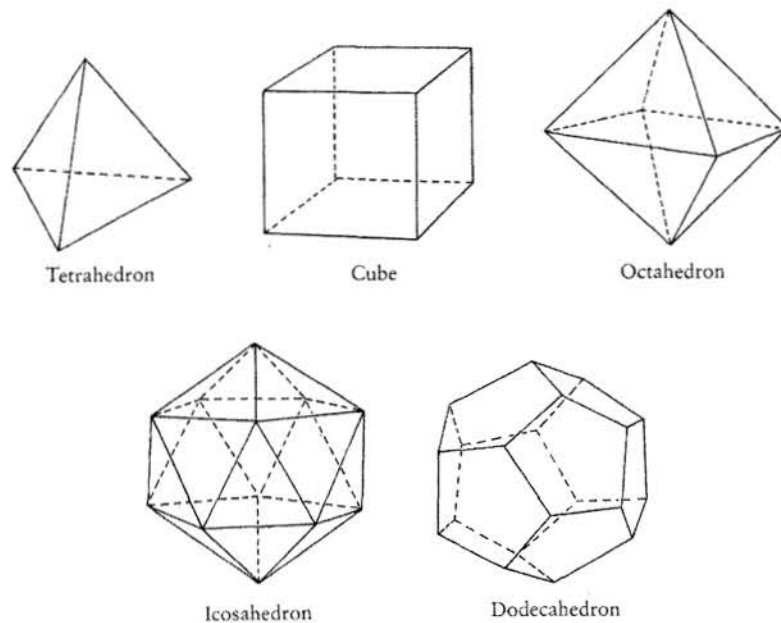


Figure 2-7 The five regular polyhedra.

actual number of vibrations per second. Like Galileo, with whose work he was familiar, Mersenne knew how the frequency of vibration varies with the length of a stretched string (frequency is proportional to the reciprocal of the length*), with tension (it is proportional to the square root of the tension), and with mass per unit length (it is proportional to the reciprocal of the square root of mass per unit length). Putting all this together, we find

$$\text{frequency} = k \times \frac{\sqrt{\text{tension}}}{\text{length} \times \sqrt{\text{mass per unit length}}}$$

But by what factor k must we multiply this product of quantities in order to get the actual number of vibrations per second? Mersenne found the correct factor by counting the number of vibrations per second of long strings, including a hemp cord 90 feet long and $\frac{1}{12}$ inch in diameter, and a brass wire 138 feet long and $\frac{1}{48}$ inch in diameter.

*For any number n , the reciprocal is $1/n$.

And what did he find for k ?

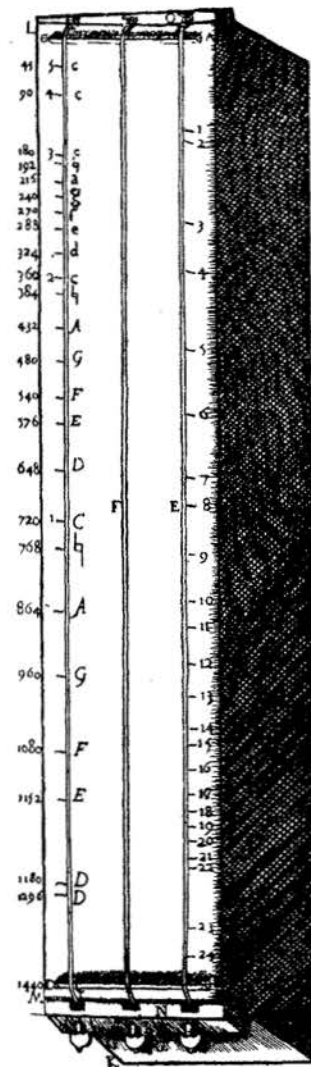


Figure 2-8 The monochord, used to relate lengths to pitch, as depicted in Mersenne's *Harmoni Universelle*.

Not everyone immediately appreciated or accepted these ideas. According to Samuel Pepys's diary entry for August 8, 1666, Robert Hooke told him that "he is able to tell how many strokes a fly makes with her wings (those that hum in their flying) by the note that it answers to in musique during their flying." Pepys characterized this as "a little too much refined." Nevertheless, we now know that the periodic nature of musical sounds arises from the nature of waves, in air, water, and strings.

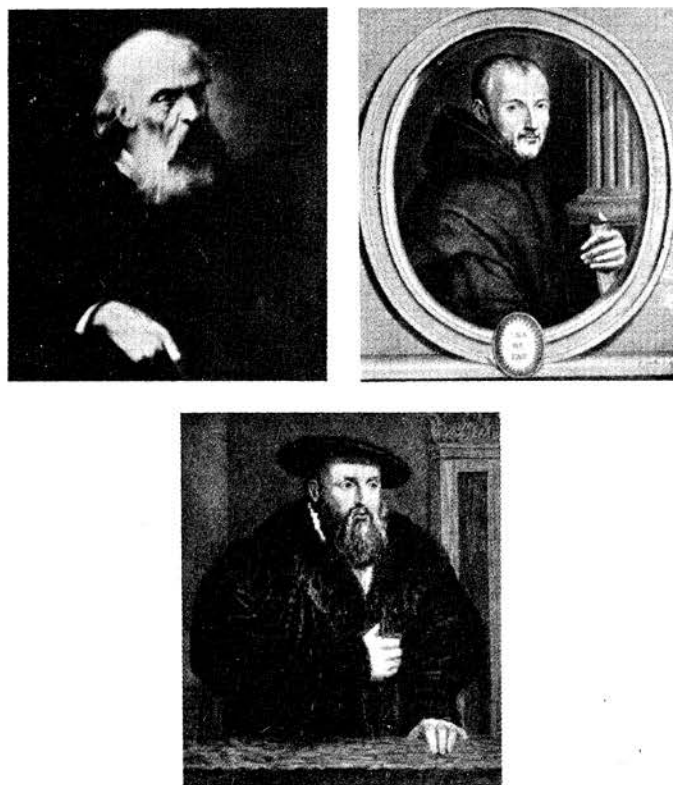


Figure 2-9 Portraits of Galileo, Mersenne, and Kepler.

Sound Waves

We have all seen the circles of ripples that move outward when a raindrop falls into a quiet pool, or when we drop a pebble into smooth water. In a similar way, a disturbance of the air moves out from the disturbing source. However, sound waves do not travel merely on a surface, but through the air in all directions. The air in a sound wave does not move bodily, as water flows in a stream, but only locally. One part of the air imparts motion to the air ahead, as might happen if, in a long line of closely spaced people, a person at the end gave a push to the one ahead, and that one, in turn, pushed the next. We can imagine a disturbance traveling to the head of the line without anyone taking a step forward.

We experience sound in air, but we can't see sound waves. Furthermore, they pulsate so rapidly that we can't feel their individual pulsations, except perhaps in the lowest notes of a pipe organ. We can visibly represent the motion



Figure 2-10 Ripples on water.

of the air in a sound wave with a contrivance made of a series of little weights connected by springs (see Figure 2-11). If we give a sharp blow to one end of this device, the weights down the line move forward and back in turn, and we see a wave travel along the series of weights and springs. The figure shows successive positions of the masses as a single wave, or pulse, travels to the right. This contrivance produces waves that, unlike sound waves, travel in only one dimension, a straight line. However, the springs and weights accurately represent the two properties of air that allow the propagation of sound waves — elasticity and mass.

We are familiar with the elasticity of air from experience with automobile tires and bicycle pumps. When we compress air, it gets smaller and fits into a smaller volume. Conversely, air under pressure expands if it can, as we see when we blow into a toy balloon. That air has mass is evident in every breeze and strong enough to make visible objects move. Sailboats and the leaves of trees move in the wind because the moving air imparts to them some of its momentum, a property of moving things that have mass.

A single sudden disturbance, such as the explosion of a firecracker, pushes the air next to the disturbing object. Because air has mass and elasticity, it resists and is compressed. The compressed air then expands again, pushing in all directions against the air around it. The surrounding air in turn becomes compressed, forming a shell of compressed air a little

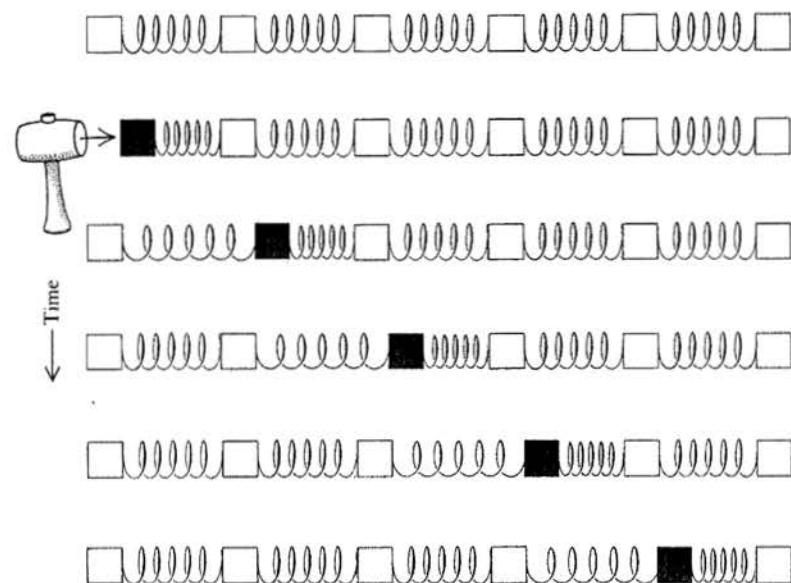


Figure 2-11 A wave can travel along a sequence of weights connected by springs. Waves travel through air in a similar manner, for air has both mass and springiness, or elasticity.

distance from the original disturbance. The expansion of the air in this shell creates yet another shell, farther out, and so on (see Figure 2-12). Thus, we can think of a single sound wave as an expanding shell of compression. Successive layers of air are compressed and decompressed as the wave moves outward from the source of disturbance, but each individual air molecule moves only a little distance out and back. A steadily vibrating object, such as a plucked string, starts an expanding shell of compression with each vibration. Thus, we can visualize the emitted tone as a series of expanding shells (see Figure 2-13).

In Figure 2-11 and in actual sound waves, the movement of the medium, whether springs and weights or air molecules, is back and forth along the direction in which the wave itself travels. Such waves are *longitudinal*. In contrast, waves in water are *transverse*: That is, most of the motion of the water is up and down, *sideways* to the direction of the visible wave. Plucked strings also exhibit transverse waves.

The tension in a stretched string is a force that tends to straighten out bends in the string. Because of the tension, a stretched string, when pulled aside at one point, pulls back at that point. However, because it has mass

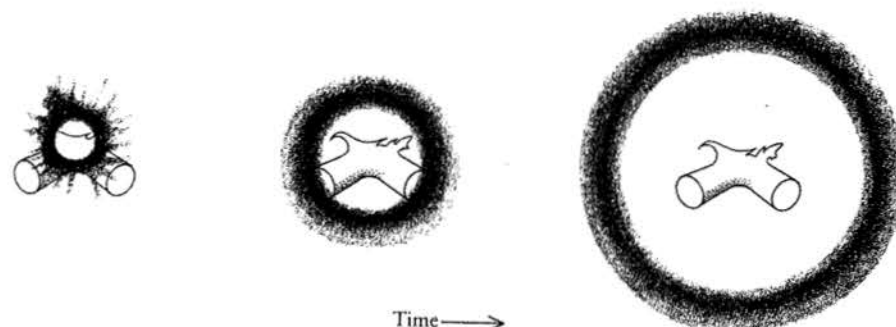


Figure 2-12 The sound wave excited by an exploding firecracker is a single expanding spherical shell of compression in the air.

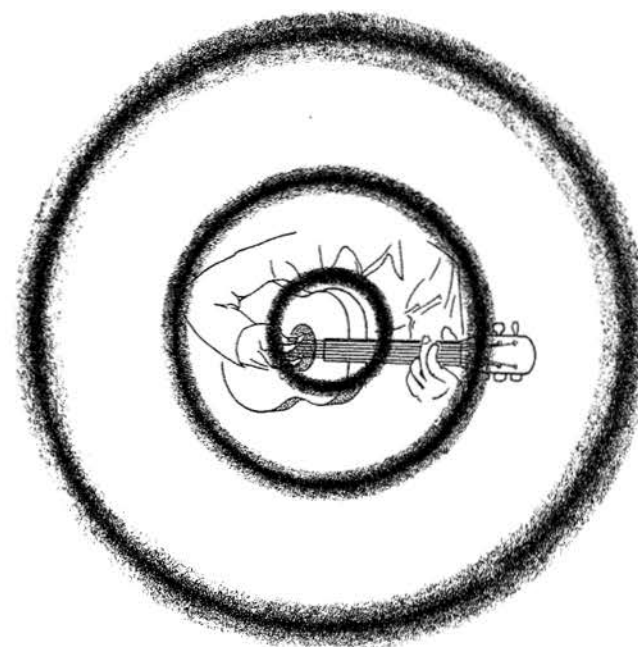


Figure 2-13 A steadily vibrating object, such as a vibrating string, sends out a series of expanding shells of compression. If the object vibrates fast enough and with enough force, we can hear the succession of shells as a tone.

and, thus, inertia, it does not simply return to a straight line after it has been released, but moves on, to bend on the opposite side. Here the force of tension acts in the opposite direction and pulls the string back again. Thus the tension and mass of the string act together to keep it vibrating from side to side (see Figure 2-14).

We can look at the way in which vibrating strings move from side to side in another way: We can consider the vibrations to consist of waves that travel along the string and are reflected at its ends. If we displace and release a string, we set up a wave that travels down the string to the end, where it is reflected as a sort of echo; then it starts back along the string in the other direction as a bend on the other side (see Figure 2-15).

Under favorable conditions, waves in strings can be both seen and felt. Inside the glass façade of the New York State Theater at Lincoln Center hangs a sort of curtain made of separate strands of metal beads strung on wires or cords. During an intermission, I was standing on a walkway near the very top of these strings of beads, and I idly plucked a string to see how a wave would travel down it. A noticeable transverse wave did indeed travel down and then back up the string of beads.

I wondered if I could observe such a wave at home. First, I fastened a heavy thread to a door knob and unwound several tens of feet. Then I plucked the string. The results were inconclusive: I couldn't be sure that I

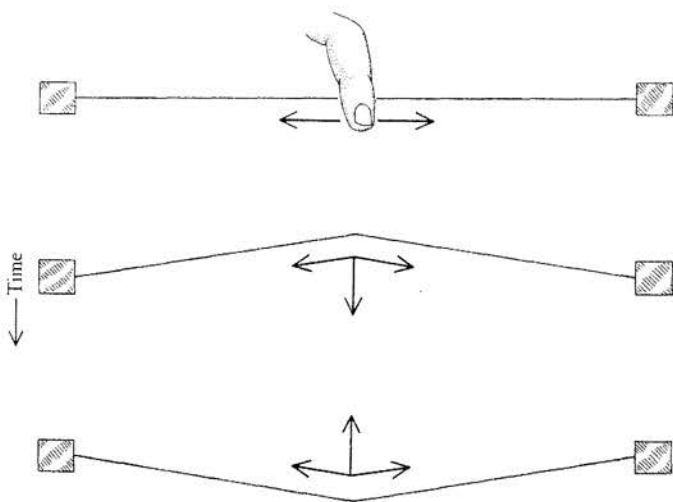


Figure 2-14 The force of tension (arrows) and the inertia of a plucked string keep the string vibrating back and forth.

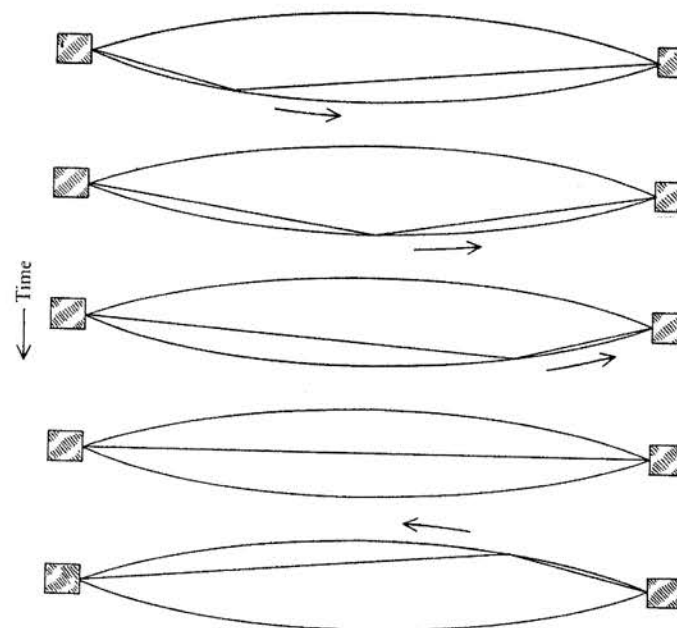


Figure 2-15 A plucked string appears to widen into a curved ribbon whose curved edges narrow to a point at each end. However, the actual shape of the string at any moment is a straight line sharply bent at one point. These “snapshots” of the bend show it traveling as a transverse wave along the stretched string. In the last three snapshots, the wave has been reflected at the right end of the string, and the bend is moving leftward on the opposite side of the string.

had observed waves, and the thread was a terrible mess when I had rewound it on the spool. I had better luck outside, with a nylon fishline. After tying one end to a tree, I unwound about sixty feet of line, stretched it lightly, and plucked the end that I held. I couldn't convince myself that I could see a wave, but I could certainly feel it as a series of jerks or pulses. The wave that I had set up traveled to the tree, was reflected there as an echo, and was reflected again at the spool that I held, traveling back and forth several times before dying out. I felt a little jerk each time the wave reached me.

Whether waves travel longitudinally in air, or transverse along stretched strings or in water, they are reflected when they encounter a solid, immovable obstacle. (The reflection of waves is discussed in Appen-

dix E.) A reflected wave is called an *echo*. How long does it take an echo to return? That depends on the distance and on the speed of the wave. In a stretched string, the velocity, v , of the waves, measured in meters per second, is given by

$$v = \sqrt{\frac{T}{M}}$$

In the equation, T is the tension of the string, measured in newtons, and M is the mass per unit length of the string, measured in kilograms per meter. (Appendix C lists and comments on such units. Appendix D tells more about waves.) Movement of the waves is slower in a more massive string and faster in a string that is stretched more tightly.

A violin dealer once came to Max Mathews for advice about choosing strings for a violin. He knew that a steel string required more tension to produce the same pitch as a gut string, but he didn't know how much more, and he feared that trying the steel strings might damage the violin. Mathews told him to estimate the increase in tension by weighing the strings. If a steel string is more massive than a gut string, its tension must be increased in the same ratio to give the same frequency and pitch.

The velocity of sound waves is not affected by changes in air pressure, because compressed air has both greater density, which slows down the waves, and greater elasticity (resistance to compression), which speeds up the waves. However, the velocity of sound waves does increase with a rise in temperature, because the elasticity of air results from the motion of its molecules, which move faster as the temperature is increased.

At room temperature (conventionally, 20 degrees Celsius or 68 degrees Fahrenheit), the velocity of sound is 344 meters per second, 1,128 feet per second, or 769 miles per hour. This velocity was first measured, somewhat inaccurately, by Mersenne in about 1636. He measured the time interval between seeing the flash of a gun at a known distance and hearing the sound. A more satisfactory measurement was made later, in about 1750, under the direction of the Academy of Sciences in Paris. Knowing the velocity of sound in air, if we count the seconds between seeing a flash of lightning and hearing the associated thunder, we can easily tell how far away the lightning struck. The distance in feet is about 1,000 times the number of seconds—about five seconds for a mile, or three seconds for a kilometer.

Echoes, Frequency, and Pitch

Under certain circumstances, echoes demonstrate the relation between frequency and pitch. Perhaps the simplest way to hear this is to stand in front of a receding sequence of regularly spaced vertical surfaces. These could be a sequence of solid seats, as in a stadium, or the seats of a Greek-style theater (see Figure 2-16), or even a long flight of broad concrete steps. A handclap or any other sharp sound will be reflected from successively more distant surfaces, so that it returns as a series of echoes. What is the time interval between successive echoes?

Suppose that the reflecting surfaces are spaced a distance W apart. In reflection from the next farther surface, the sound has to travel an added distance W to reach the surface and an added distance W to get back (see Figure 2-17). For an observer at the sound source, the time interval T between one echo and the next is two times W divided by the velocity v of sound:

$$T = \frac{2W}{v}$$

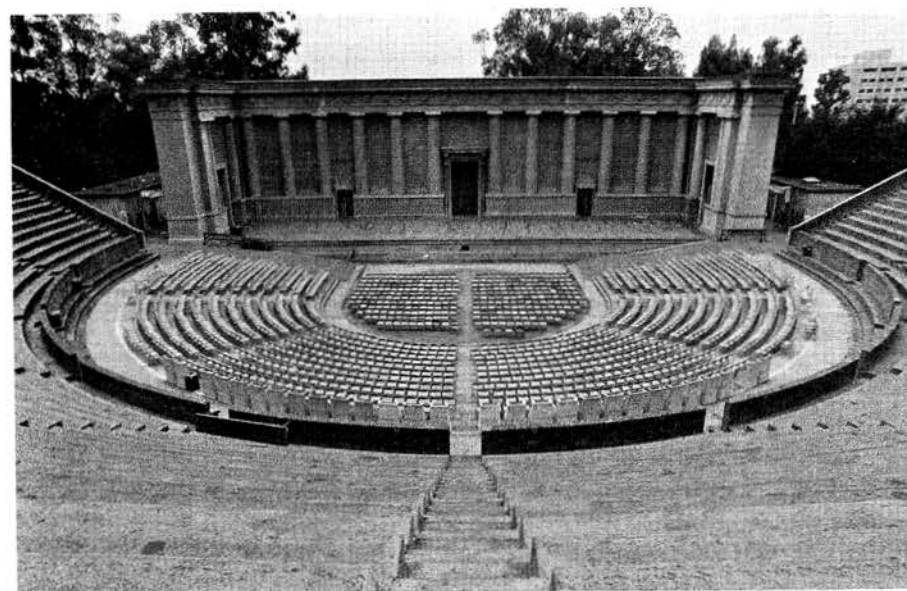


Figure 2-16 Greek-style theater at the University of California, Berkeley.

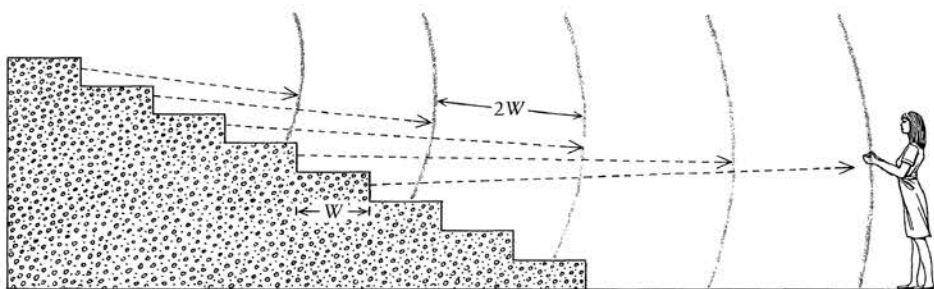


Figure 2-17 A single handclap in front of a receding sequence of steps or seats, as in a football stadium or a Greek theater, produces a periodic succession of echoes. If they are loud enough, they can be heard as a sound having a definite pitch.

The frequency f , which is the number of echoes per second, is thus

$$f = \frac{1}{T} = \frac{v}{2W}$$

If the distance W between seats is 3 feet, the frequency would be 188 echoes per second. This corresponds to a pitch of about the first G below middle C. For steps having a standard depth of 10 inches, the frequency would be 677 echoes per second, or a pitch between those of the second E and F above middle C.

It is easier to hear successive echoes from two hard, parallel surfaces. I succeeded in hearing a sequence of echoes as I stood on a concrete walkway under a projecting canopy of the Drama Building of the University of California at Santa Cruz (see Figure 2-18). When I clapped my hands, I heard a distinct sequence of claps, or pulses, of decreasing intensity. The overhang was about twenty feet above the concrete on which I stood so the reflections repeated with a frequency of about 28 per second. I didn't hear them as a pitch, but as a sequence of sharp sounds (see Figure 2-19). The hall of our house is four feet wide; for someone standing in the middle of the hall, the reflections coming alternately from *both* walls would have a frequency of

$$f = \frac{v}{w} = \frac{1128}{4}$$

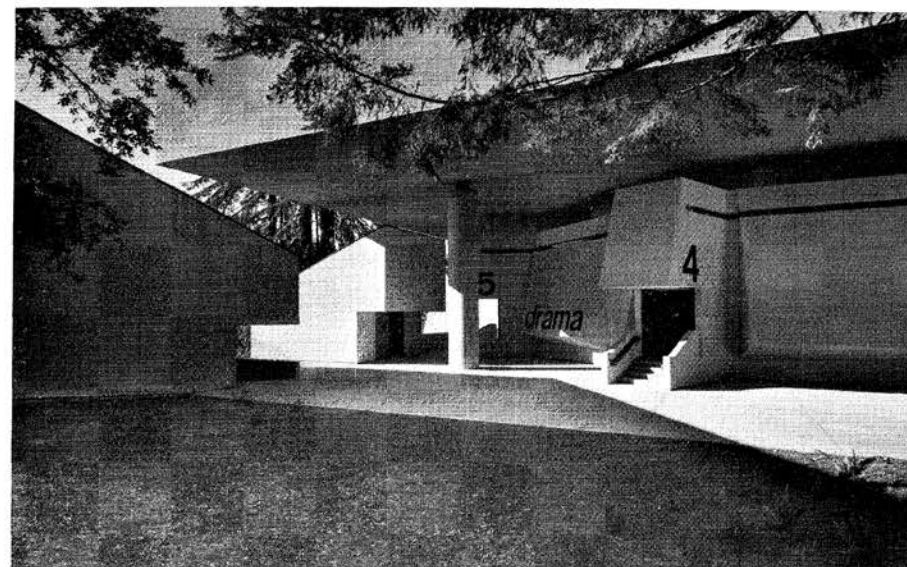


Figure 2-18 The overhang at the Drama Building, University of California at Santa Cruz.

This calculation gives a frequency of 282 echoes per second—roughly, the C# (C-sharp) above middle C. I can't say that I actually hear this pitch when I clap my hands in that hall, but on occasion I have heard a tonelike flutter when I clapped my hands in some rooms and halls that had hard, uninterrupted parallel walls.

Vibrations in Musical Instruments

As we have seen, with each vibration of a stretched string, a transverse wave travels the length of the string twice—once in each direction. Thus the frequency, the number of vibrations per second, is given by

$$f = \frac{v}{2L}$$

Here L is the length of the freely vibrating string, and v is the velocity with which a transverse wave travels along the string. It has already been noted that this velocity is the square root of the ratio of the tension T to the mass

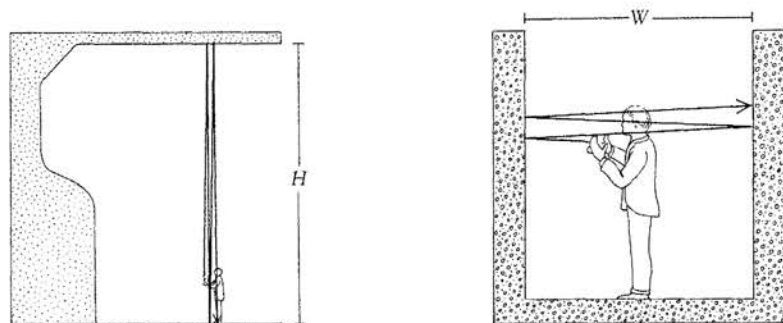


Figure 2-19 The echo from a handclap under a high overhang must travel up to the overhang and back, a distance approximately $2H$, every time it returns to the observer. The echo from a handclap in the middle of a hallway must travel to the wall and back, a distance W , every time it returns.

per unit length, M . Thus, the frequency of vibration of the stretched string of a musical instrument can also be written.

$$f = \frac{\sqrt{T/M}}{2L}$$

All stringed instruments are tuned by adjusting the tension T of the strings. Thus, tightening a guitar string, for example, raises the pitch of the note that it produces, because an increase in T causes an increase in f . While playing the guitar, however, the musician continually changes the pitch of the strings by changing their length L . When a string is depressed on one side of a fret, the string does not vibrate on that side; the rest of the string, being shorter than the whole, then produces a higher note when it is plucked. Although the main change in pitch is caused by reducing the length of the string, the guitarist can also introduce minor variations into the sound of a vibrating string by wiggling the finger that holds the string down. This causes very slight fluctuations in T , and thus in f , so that the string does not produce a steady tone, but a vibrato.

Each string of any instrument must be chosen carefully so that, at a reasonable tension, the string will produce a sound of the desired pitch and loudness. In a piano, the bass strings are heavily overwound to increase their mass, besides being longer than the treble strings. Because of the great mass of the bass strings, a wave travels along them slowly; because of their long length, it has to travel far. The massive bass strings also give a loud sound. In order to make higher notes comparably loud, two strings (also

overwound) are used per note in the midrange and three strings (plain) in the upper range.

In wind instruments, the frequency is determined by the time it takes a sound wave in air to travel from one end of a tube to another. A pipe of length L that is open at both ends vibrates at a frequency

$$f = \frac{v}{2L}$$

A pipe closed at one end gives a sound one octave lower than an open pipe of the same length. In an organ, for example, an open 8-foot pipe has a frequency of about 70 vibrations per second, roughly two octaves below middle C. A 16-foot open pipe or an 8-foot pipe closed at one end would sound another octave lower.

Organ pipes are arranged in ranks of pipes (also called *stops* or registers). A stop contains one pipe for each key of the keyboard and can be referred to by the length of its longest pipe. Thus all the pipes of a 16-foot stop are twice as long as the corresponding pipes of an 8-foot stop, and all their pitches are an octave lower.

The pipes of different stops differ not only in length but also in shape and composition, which gives them different *timbres*, or qualities of sound. Furthermore, two or more stops can be played simultaneously to give special timbres. *Mutation stops* are designed for just this purpose: the nazard, or $2^{2/3}$ -foot stop, gives sounds pitched an octave and a fifth above those of an 8-foot stop; the tierce, or $1^{3/5}$ -foot stop, gives sounds pitched two octaves and a major third above those of an 8-foot stop; the larigot, or $1^{1/3}$ -foot stop, gives sounds pitched two octaves and a fifth above those of an 8-foot stop. By employing mutation stops together with 8-, 4-, or 2-foot stops, organists produce strange, juicy timbres that may sound to the uninitiated like a lot of wrong notes going along with the melody.

We have seen in this chapter that the pitch of a sound depends on the frequency of the vibrations that produce it. In a siren, the frequency, and hence the pitch, of the sound depends on the number of holes that pass the nozzle each second. Echoes from a sequence of equally spaced surfaces can produce a tone whose frequency is determined by the number of echoes that reach the observer each second. In a vibrating string or a speaking organ pipe, the pitch is determined by how much time it takes a transverse wave or a sound wave to travel back and forth along the string or pipe. The next chapter explores ideas of frequency and periodicity in more detail.

A Note on Musical Pitch

Pitch originated as a musical term and has become a psychological term used to designate a perceived quality of sound.

In music, the pitch of musical sounds was perceived long before the physical basis for pitch was understood. One of the great musical (and psychological) discoveries is that for periodic musical sounds, such as those produced by the organ, strings, winds, and the human voice, pitch is tied unalterably to the periodicity or frequency with which the waveform of the sound repeats.*

Periodic musical sounds are made up of many harmonically related frequency components, or *partials*, of frequencies f_0 , $2f_0$, $3f_0$, $4f_0$, and so forth. Such sounds have many perceived qualities besides pitch. One of these other qualities is shrillness, or brightness. A sound with intense high-frequency partials is bright, or shrill. A sound in which low-frequency partials predominate is not bright, but dull.

When you listen to periodic musical sounds on a stereo system, you can change the brightness by turning the tone control, but this doesn't change the pitch. The brightness depends on the relative intensities of partials of various frequencies. Turning the tone control can change the relations of the partials, but won't change the periodicity of the sound, which is the same as the *fundamental*, the frequency of the first partial, f_0 .†

Sounds that are not periodic musical sounds are not as clear and distinguishable in pitch and brightness, but some of them can be granted pitch by a sort of musical courtesy. Among these are sine waves (pure tones), the tones of bells, the clucking sound that we can make with the tongue and the roof of the mouth, the somewhat related sound of the Jew's harp, and the sound of a band of noise.

Sine waves are peculiar in that they consist of a single harmonic partial. The sense of pitch that they give is not as certain as that of other periodic sounds; it can differ a little with intensity, and between the two ears. For other periodic sounds, the sense of the octave is very strong, for the partials of a sound a' (that is, an octave above sound a) are all present in sound a .‡ The sense of the octave is not strong with sine waves.

try this!

*There can be an ambiguity of an octave in the pitch of a periodic musical sound, because musicians sometimes (though rarely) report the pitch as being an octave away from where it falls according to its periodicity.

†The first partial need not be physically present in the sound wave, but its absence doesn't alter the pitch. This psychological phenomenon will be discussed in Chapter 6.

‡If a sound has frequency a , its partials have frequencies $2a$, $3a$, $4a$, and so forth. The octave of a has frequency $2a$, which has partials $4a$, $6a$, $8a$, and so forth. All frequencies in this latter series also occur in the first series.

Furthermore, because sine waves contain only one frequency component, their brightness is tied inextricably to their pitch.

Musically trained people react to sine waves and their pitches much as they react to periodic musical sounds. Naive people may react differently. By asking naive subjects to relate frequency changes of sine waves to a halving of pitch, psychologists found a *mel scale* of pitch (for sine waves). In the mel scale there is no simple relation between frequency and pitch; nothing like the octave shows up. I think the mel scale is a scale of brightness, not of pitch. It might be possible to check this by using musical sounds whose brightness and pitch could be varied independently.

The sounds of orchestral bells and of tuned bells are not periodic, and these sounds do not have all the properties of periodic musical sounds. One can play tunes with bells, and the pitches that are assigned to bells can be explained largely in terms of the frequencies of prominent, almost-harmonic partials.

Clucking sounds and shushing sounds (bands of noise) have a brightness, but no periodicity. Oddly, *we can play a recognizable tune with these sounds*, even though they cannot be heard as combining into chords or harmony. Apparently, in the absence of a clear pitch, brightness can suggest pitch. This seems natural. When we play a scale on a musical instrument, the brightness increases as we go up the scale. But the "pitch" of clucks or bands of noise is only a *suggestion* of pitch. It depends on the frequency at which the brightness spectrum peaks, and this (and therefore the "pitch") changes when we turn the tone control.

Even periodic sounds can be constructed to give unusual pitch effects, but such sounds are *not* produced by musical instruments. For periodic musical sounds, the pitch is tied firmly to their periodicity, the frequency of the first harmonic partial. The only mistake we can make in "confusing" pitch, a sensation, with periodicity, the numerical frequency of the fundamental, is that of offending psychologists.

3 Sine Waves and Resonance

We have seen that sound waves travel from the source of sound to our ears as fluctuations in air pressure. Different sorts of fluctuations cause us to hear a wonderful variety of different and interesting sounds, sounds that we can identify and appreciate. Among these are musical sounds, in which the air pressure rises and falls almost periodically with time and which have a pitch that corresponds to the frequency of this nearly periodic rise and fall in air pressure.

We can pick up the variation of sound pressure with a microphone, and we can use a cathode-ray oscilloscope to trace out the way that the sound pressure varies with time. Thus we can see the waveform of a sound wave. Is there a method that allows us to analyze sound waves and understand why the ear responds differently to different sounds? There is, indeed, and that is the subject of this chapter.

Sine Waves

Among the sounds used in the laboratory there is one, called a *pure tone*, or *sine wave*, in which the air pressure rises and falls sinusoidally with time. A mathematical explanation of sine waves is given on pages 61–63.

A sine wave is simply a tone for which the air pressure varies sinusoidally with time. It is a mathematical function that has unique and important properties. We can represent any periodic variation of air pressure over time as a sum of sinusoidal components (as we will see later). Furthermore, in responding to various sounds, the mechanism of the ear partly sorts out sinusoidal components of different ranges of frequency. When

we listen to a musical sound, such as that of a voice or of a clarinet, different nerve fibers that go from the ear to the brain are excited by different ranges of the sinusoidal components. We can find these ranges by analysis of the sound wave.

The shape of a sine wave is completely described by three attributes: *amplitude*, *period*, and *phase*. First, the *amplitude* represents the maximum displacement of the varying quantity from its average value. In Figure 3-1, the amplitude is h . In a sound wave, air pressure periodically rises to a high pressure (P) and falls to a low pressure ($-P$), above and below the average air pressure.

Second, the *period*, T , of a sine wave is the time between amplitude peaks, usually measured in seconds. The reciprocal of T , $1/T$, therefore gives the number of peaks per second, which is the *frequency*, f , of the sine wave:

$$f = 1/T$$

Frequencies used to be started as cycles per second, or cycles for short. Today the term *hertz*, honoring the physicist Heinrich Hertz (1857–1894), is used to designate cycles per second. The abbreviation of hertz is Hz.* For example, a sine wave with a frequency of 440 Hz has a period given by

*Often people use *frequency* to describe the rate of oscillation of complex waveforms that are made up of many sinusoidal components of different frequencies. In this book I try to use *frequency* and *hertz* only for sine waves, and *periodicity* for the number of cycles per second of complex waveforms. This may seem awkward, but it is unambiguous.

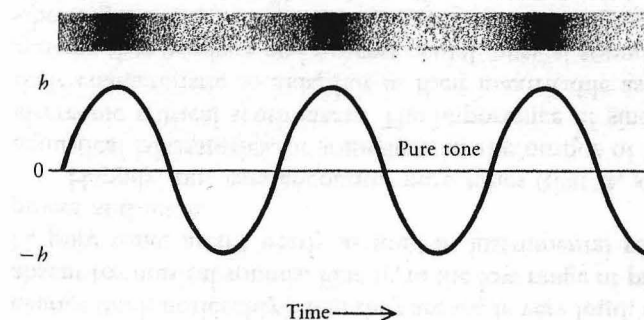


Figure 3-1 A sine wave, or pure tone.

$$T = 1/440 = .0022727 \text{ seconds}$$

Similarly, a sine wave with a period of $1/1000$ or 0.001 second will have a frequency of

$$f = 1/0.001 = 1,000 \text{ Hz}$$

Third, a sine wave has a *phase*. The two sine waves shown in Figure 3-2 have the same frequency and amplitude, but have different phases because they reach their amplitude peaks and cross the horizontal axis at different times.

Why are pure tones, with their sinusoidal variation of air pressure, important for understanding musical sound? We almost never hear pure tones, except in the laboratory, or when we listen to a tuning fork that isn't struck too hard. Pure tones sound very uninteresting. When the pitch is low, they sound like the hum of a malfunctioning radio. Pure tones of higher pitch are steady but not bright or interesting whistles. Pure tones are also, in many ways, odd and unnatural sounds. In a reverberant room we can't sense the direction from which they come. Some people hear the pitch of a pure tone differently in the left and right ears; this is called *diplacsis binauralis*. Even for people with normal hearing, pure tones

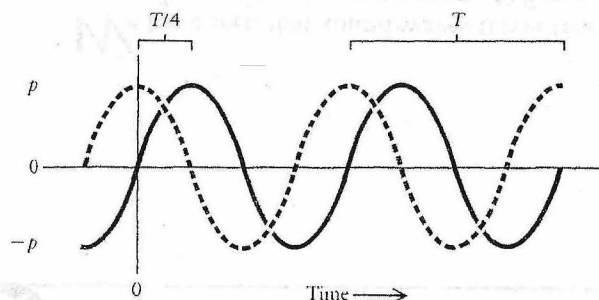


Figure 3-2 A sine wave is characterized completely by three quantities: its *amplitude* or extreme height, its *period* (the time between one peak and the next), and its *phase*, which we can take as the time when the wave crosses the axis when going upward. In this figure, the solid curve crosses the axis going upward at time $t = 0$. The dashed curve peaks at $t = 0$ and crosses the axis going upward at a time $T/4$ earlier (at $t = -T/4$, in which T is the period of the wave). The solid curve is called a *sine* curve, and the dashed curve is called a *cosine* curve. A sine curve and a cosine curve differ only in their phases.

change pitch noticeably when they are made very loud; this effect is nearly absent for musical sounds. Finally, in the low range of pitch, below middle C, pure tones aren't nearly as loud as instrumental sounds of the same power and pitch.

Happily, our ears encounter pure tones (that is, sine waves) only in acoustical laboratories, or sometimes as the output of inferior or ill-used electronic musical synthesizers. The importance of sine waves lies not in their characteristic sounds, but in their inextricable association with the systems that produce and process actual musical sounds, systems such as vibrating strings, or columns of air, or bells and gongs, or drumheads. These are, approximately at least, *linear systems*, and sine waves are the mathematical way in which the time variations of the vibrations of linear systems are expressed.

We are all familiar with stereo amplifiers. Such amplifiers are linear. If you put one frequency in (via a radio tuner or CD player for example), you get only that frequency out, not other sounds. If you double the amplitude of the input signal, you double the amplitude of the output signal. These are more than mere characteristics of linear systems; a system *is* linear if these things are true.

But if you put too intense a sine wave into the input of any actual amplifier, the sine wave will produce extraneous sounds in the output. The amplifier overloads, and you will hear frequencies in the output that weren't in the input. Further, the amplitude of the output signal is no longer directly proportional to that of the input signal. Physical systems aren't exactly linear, whether they are amplifiers or the mechanical parts of the middle and inner ear that vibrate in accord with the sounds about us. But at small amplitudes, actual physical systems do behave in a linear fashion.

This is true of strings that vibrate when struck or plucked, and of columns of air in organ pipes or horns, and of bells and gongs and drumheads. Sine waves also have a crucial relation to such vibrations. When such a resonant system is set in motion, the overall vibration is found to be made up of a host of different patterns of vibration called *modes*. The vibration of each mode varies sinusoidally with time. Or rather, the vibration functions as a decaying sine wave, a wave that loses some constant fraction of its amplitude in each short successive period of time (such decay is called *exponential*). Thus, at their very source, musical sounds are associated with sine waves—but with a collection of sine waves of many different frequencies.

The French mathematician François Marie Charles Fourier (1772–1837) invented a type of mathematical analysis by which it can be proved that any periodic wave, however it may have been generated, can be

represented as the sum of sine waves having the appropriate amplitude, frequency, and phase. Furthermore, the frequencies of the component waves are related in a simple way: They are all whole-number multiples of a single frequency, for example, f_0 , $2f_0$, $3f_0$, and so on.

A Fourier representation of a complicated wave can require very many components, even an infinite number. However, to approximate a wave to a desired degree of accuracy, fewer components may be adequate. A deceptively simple-looking wave whose Fourier representation requires an infinite number of components is the *square wave* (see Figure 3-3). A square wave having amplitude 1 and frequency f_0 can be represented as the sum of sine waves having frequencies f_0 , $3f_0$, $5f_0$, $7f_0$ (and so on, indefinitely), amplitudes 1, $1/3$, $1/5$, $1/7$ (and so on, indefinitely), and the proper phases (see Figure 3-4). (If the number of components being added is finite, as it will be in any real amplifier, the resulting square wave will not have exactly sharp corners; that is, it will be somewhat distorted.) We should note that square waves are peculiar in being made up only of odd frequency components. Most periodic sound waves consist of both odd and even frequency components (although closed organ pipes and some wind instruments do have predominantly odd frequency components).

There are three different systems for naming the sinusoidal frequency components of a periodic sound, as shown in Table 3-1. Notice that the number of the harmonic or partial is the same as that of the relative frequency; for example, $5f_0$ is the fifth harmonic or the fifth partial. However, $5f_0$, the *fifth* harmonic, is the *fourth* overtone, so it is important not to confuse these three terms.

It is convenient to use the term *harmonic* to deal with strictly periodic sounds. However, some musical sounds (usually percussion) consist of frequencies that are not harmonic, that is, they are not integral multiples of the lowest frequency. For example, the frequencies of a “free” (lightly supported) vibrating rod or bar might be

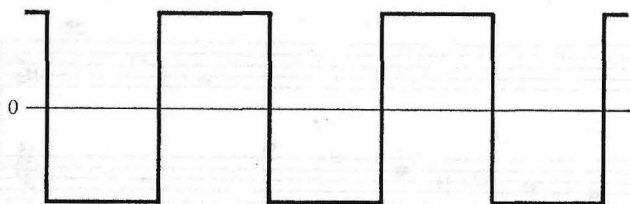


Figure 3-3 A square wave.

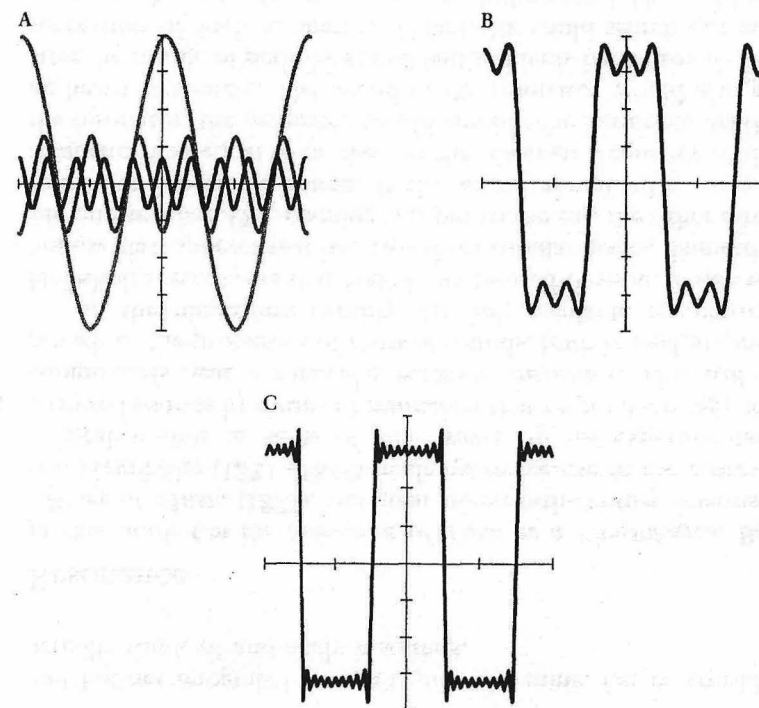


Figure 3-4 A. Three waves, of frequencies f_0 , $3f_0$, and $5f_0$. B. The wave that results from adding together the three waves in part A. They roughly approximate a square wave. C. The approximation to a square wave produced by adding together the first nineteen components in the same series as in part A.

Table 3-1. Systems for Naming Frequency Components

Frequency	Harmonics	Overtones	Partials
f_0	Fundamental	Fundamental	First partial
$2f_0$	Second harmonic	First overtone	Second partial
$3f_0$	Third harmonic	Second overtone	Third partial
$4f_0$	Fourth harmonic	Third overtone	Fourth partial

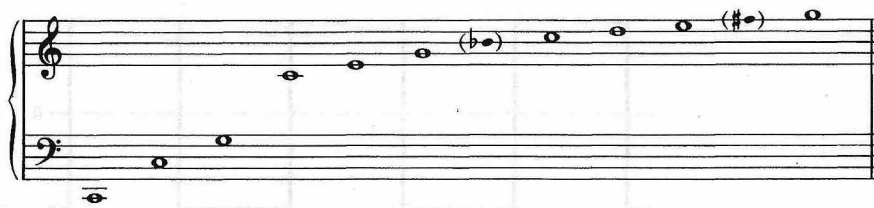


Figure 3-5 The first twelve harmonics of C_2 . The representations of the pitches of the seventh and eleventh harmonics are approximate.

$$f_0, 2.756f_0, 5.404f_0, 8.933f_0, \text{ etc.}$$

It would sound rather illogical to call these higher frequencies “nonharmonic harmonics,” so instead they are called “nonharmonic partials.” It is also convenient that the numbering of the partials agrees with the numerical order of the frequency. The lowest frequency is always the first partial, the next higher frequency is the second partial, and so on.

In principle, a sum of sine waves can be used to represent any sound. If the sound is of finite duration, as actual musical sounds are, it cannot be exactly periodic over all time, as a Fourier series of harmonic partials implies. In representing exactly sounds of finite duration, we must include components of *all* frequencies by using a *Fourier integral*.

Let us also consider “noisy” sounds, such as the hiss of escaping air or the *sh* or *s* sounds of speech. Such sounds can be represented as the sum of sine waves that have slightly different frequencies. If you sound the *sh* twice, the waveforms won’t look exactly alike. The power of the sound in any narrow range of frequencies will be about the same, but the amplitudes and phases of the individual frequency components won’t be identical. Nevertheless, the two *sh* sounds will *sound* just the same; we will *hear* them as being the same.

Trying to represent actual sounds as sums of *true* sine waves, which persist from the infinite past to the infinite future, is a mathematical artifice. Consider the (nearly) periodic sounds produced by musical instruments. A sum of harmonically related sine waves doesn’t correctly represent such a sound, because the sound starts, persists a while, and dies away.

In practice, we use the *ideas* of sine waves and their frequencies and amplitudes to characterize musical sounds, and other sounds as well. The measurements we *really* make are those suitable for our purposes, and are as accurate as they need be. They are similar in concept to Fourier series

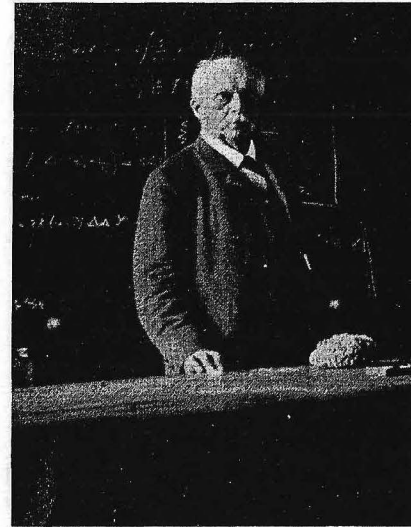


Figure 3-6 Hermann von Helmholtz.

and Fourier integrals but aren’t quite the same. Let us consider how we actually think of and analyze sounds.

Resonance

In this book *On the Sensation of Tone as a Physiological Basis for the Theory of Music* (1877), the great nineteenth-century scientist Hermann von Helmholtz (1821–1894) made extensive use of the representation of musical sounds as sums of sine waves. In his experimental work, he analyzed sounds by means of *resonators* that respond strongly to sinusoidal components near a particular *resonant frequency*. This will be our approach to the properties of musical sounds, both in analysis and synthesis.

In the nineteenth century the only available resonators were the Helmholtz resonators that Helmholtz himself devised. These were usually hollow glass spheres that had two short tubular necks, diametrically opposite one another. One opening was put to the ear, the other directed at the source of a periodic sound. If the sound contained a harmonic whose frequency was equal to or close to the resonant frequency of the cavity of the resonator, the resonator would amplify the harmonic, so that it could be heard separately. The sound in the resonator would also persist even after the source of periodic sound had suddenly been turned off. By using a succession of such resonators, Helmholtz could search out and estimate the strengths of the harmonics of a periodic sound. He could also find the

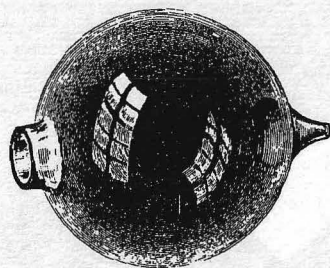


Figure 3-7 A Helmholtz resonator.

frequencies of nonharmonic partials, such as those of bells and gongs (see Chapter 13).

We can use a piano to experience something of what Helmholtz did. Hold the damper or sustaining pedal (the “loud” pedal”) down (this lifts the felt dampers off the strings), and whistle near the strings. After you stop whistling, you will hear a ghostly persistence of the note that you have whistled. This dies away with time. The piano strings act as resonators. Those that can vibrate at the frequencies of the whistled note do so; those that can’t, don’t.

A taut piano string can vibrate at more than one frequency. The lowest, or *fundamental*, mode of oscillation was described in Chapter 2 (see Figure 2-14). These vibrations can be regarded as sine waves traveling along the string with a constant velocity v , and being reflected repeatedly at the ends of the string. If L is the length of the string, and if the wavelength (distance between peaks of the sine waves) is $2L$, we get the pattern of vibration shown in part A of Figure 3-8, in which the resonant frequency f_0 is

$$f_0 = \frac{v}{2L}$$

The next three resonant frequencies, $2f_0$, $3f_0$, $4f_0$, have the patterns or modes of vibration shown in parts B, C, and D, respectively.

Plucking or striking a stretched string excites many of these vibrations and produces a complex sound wave made up of many harmonics. The length, mass, and tension of the string determine the periodicity and pitch of the sound that is produced, but do not determine the exact waveform. The relative strengths of the various harmonics depend on whether we pluck or strike a string, and on where along its length we pluck or strike it. We can hear the difference in the sound quality or timbre. For example,

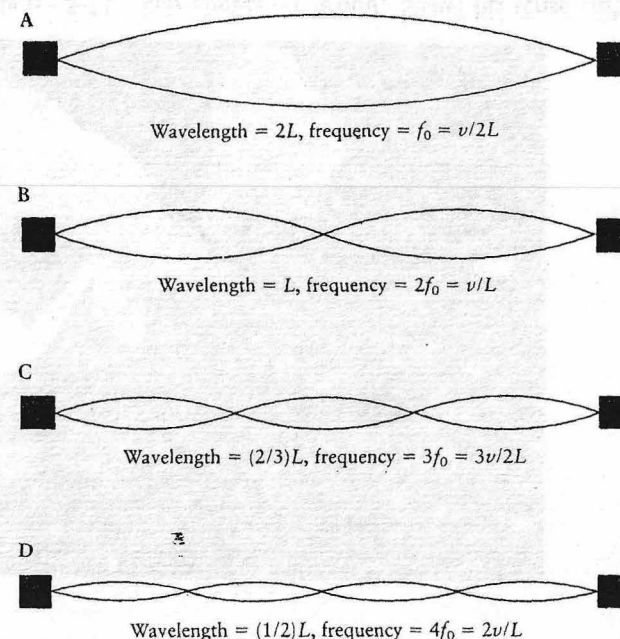


Figure 3-8 A vibrating string can have several resonant modes, which correspond to standing waves of different wavelengths and frequencies. These are the natural harmonics of the string. Parts A through D show the wavelength and frequency of the first four harmonics.

the plucked string of a harpsichord sounds quite different from the struck string of a piano. When we excite a string by bowing it, we get a persistent sound of yet another character, even though the resonances of the string may be the same.

Physical resonators are used in some musical instruments. The vertical metal tubes under the wooden bars of xylophones and marimbas, and the bamboo tubes under the brass bars of gamelan instruments, act as resonators that intensify and prolong certain of the partials generated by striking the bars, specifically, the partials that correspond to pitches of the scale. Many European baroque instruments, such as the viola d’amore, the viola bastarda, and the baryton, made use of resonators called *sympathetic strings*. These affected the relative intensities of the partials produced by bowing a string, and emitted a sustained tone even after bowing was stopped or fingering was altered.



Figure 3-9 Gamelan.



Figure 3-10 Jazz trumpeters Howard McGee (right) and Miles Davis; McGee's embouchure is relaxed; he is playing a relatively low note.

The long tubes of brass instruments resonate at harmonically related frequencies. Which frequency is excited when the instrument is played depends on how the player constricts his or her lips. The tubular structures of woodwind instruments are resonators whose resonant frequencies are controlled by opening or closing various holes.

Resonances also have a noticeable effect on the timbre of musical sounds. The vocal tract has several resonances that emphasize various ranges of frequency in the sound produced by the vibration of the vocal folds. By changing the shape of the vocal tract, we change the frequencies of these resonances, or *formants*, which determine what vowel sound we produce. Table 3-2 gives the first three resonance or formant frequencies for some vowel sounds. The resonances of the soundboard of the violin greatly affect its timbre. Figure 3-12 shows, as a function of frequency, the ratio of the sound radiated by a good violin to the motion of the bridge caused by bowing the string. The rising and falling curve reduces the intensities of some partials and increases those of others. The suppression of some partials is important for the musical quality of the violin tone. In poor violins, the sound is harsh; a curve like that shown here does not rise and fall so markedly.

Conventional musical instruments use mechanical or acoustical resonators in producing musical sounds. When analyzing a sound, we pick the



Figure 3-11 Jazz trumpeter Woody Shaw: his tense embouchure shows that he is playing a relatively high note.

Table 3-2. The First Three Formants of Selected Vowels

Formant number	Cycles per Second per Vowel							
	Heed	Hid	Head	Had	Hod	Hawd	Hood	Who'd
1	200	400	600	800	700	400	300	200
2	2,300	2,100	1,900	1,800	1,200	1,000	800	800
3	3,200	2,700	2,600	2,400	2,300	2,200	2,100	2,050

sound up with a microphone and determine its frequency content electronically. Early analyzing devices made use of electric circuits called *filters*. A filter can be described by a curve that shows how the ratio of output amplitude to input amplitude varies with frequency (see Figure 3-13).

Present-day devices for frequency analysis make use of digital techniques. In a digital *spectrum analyzer*, the Fourier spectra of short sections of the sound wave are computed by a specialized digital computer and displayed on a televisionlike screen. Or the analysis can be made by a computer with a sound-processing board. The curve on the screen shows intensity as height, and frequency as distance to the right (as shown in Figure 3-14). The curve in this figure has one peak or *line*, representing a

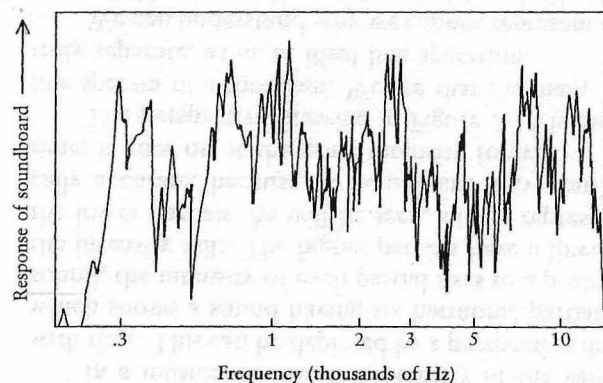


Figure 3-12 The body of a violin resonates more at some frequencies than at others. This is a plot of the intensity of the sound wave produced by exciting the bridge of a famous Guarneri violin with a sinusoidal source of constant amplitude and increasing frequency.

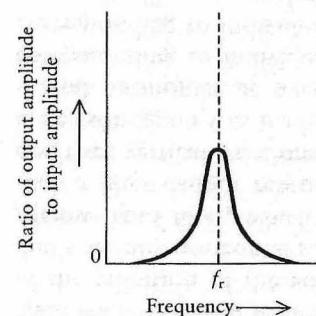


Figure 3-13 A resonator or electronic filter responds only to a narrow range of frequencies. In this curve the ratio of output amplitude to input amplitude is plotted against the frequency of the input wave. The output peaks the resonant frequency, f_r .

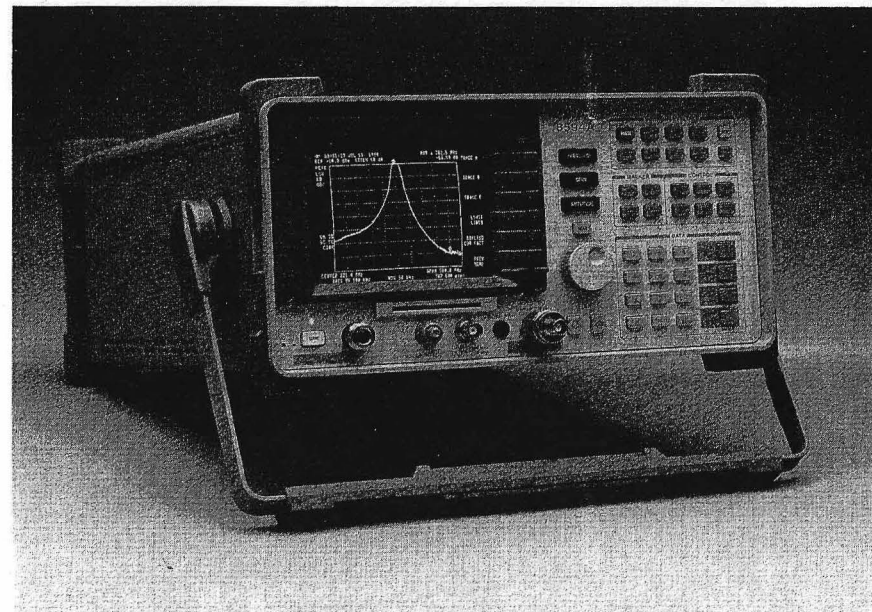


Figure 3-14 A digital spectrum analyzer displays a graph of intensity against frequency on a cathode-ray tube. (Photo of HP 8594A portable spectrum analyzer courtesy of Hewlett-Packard, Signal Analysis Division, Rohnert Park, California)

single frequency. The line spectrum of a musical sound has many sharp peaks, each corresponding to a single partial or frequency component.

In a musical sound, the intensity of the various partials can change with time. This can be depicted by a perspective drawing, as in Figure 3-15, which shows a sound having six harmonic partials. With the onset of the sound, the intensity of each partial rises to a peak; as the sound dies away, the intensity falls. The higher partials have a lower peak intensity than do the lower partials. As will be seen, such a representation isn't mathematically accurate, because an actual sine wave can't change intensity with time; it goes on at the same intensity forever.

The perspective drawing in Figure 3-16 is made up of a sequence of line spectra of a tom-tom. We see that the many frequency peaks are not truly separate, as in an ideal line spectrum.

We can understand why we cannot represent both frequency and time precisely if we consider another sound-analyzing device called a *sonograph* or *sound spectrograph*, which produces sonograms. Frequency is displayed vertically, and time is displayed horizontally; intensity is indicated by the darkness of the shading. The sonograms in Figure 3-17 are those of a human voice.

The broad horizontal bands in both sonograms represent the resonant frequencies of the vocal tract, or formants. Such resonances change rela-

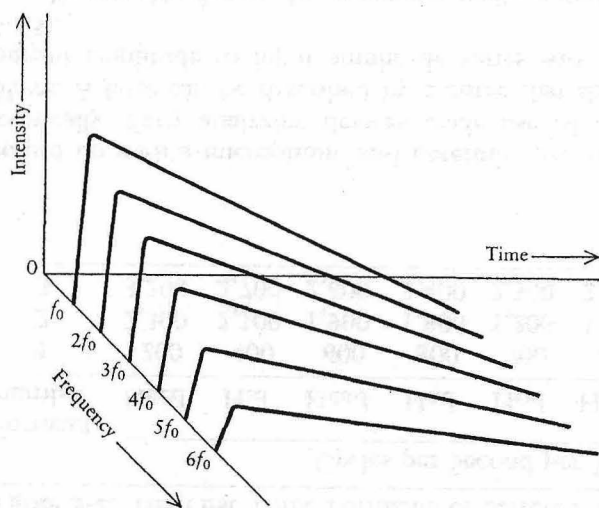


Figure 3-15 A perspective drawing of a sound in which the intensities of the six harmonics shown rise and fall with time.

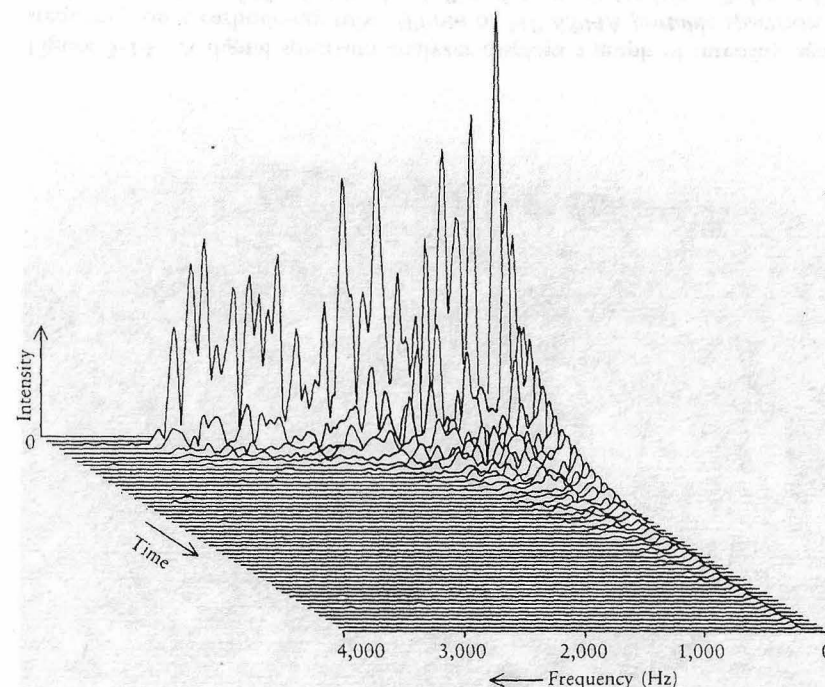


Figure 3-16 Depiction of how the spectrum of the sound of a tom-tom changes with time after the tom-tom is struck.

tively slowly with time as we speak. However, the upper and lower sonograms differ in detail. In the upper sonogram we see horizontal striations. They represent the individual harmonic partials of the voice, and their frequency separation is equal to the pitch frequency. In the lower sonogram we see vertical striations, separated by a time equal to the period, T , of the vibration of the vocal folds (the pitch frequency is $1/T$); but we don't see any horizontal striations. The upper sonogram was made with a narrow-band filter, which can sort out individual harmonics. However, such a filter cannot respond to rapid changes of sound pressure, so we don't see vertical striations corresponding to individual pitch periods. The lower sonogram was made with a broader-band filter, which responds to several harmonics at once; hence we do not see horizontal striations corresponding to individual harmonics, but we do see vertical striations corresponding to individual pitch periods.

Using filters to analyze sound means that, if we depict frequency in fine detail, we can't depict time in fine detail, or, if we depict time in fine

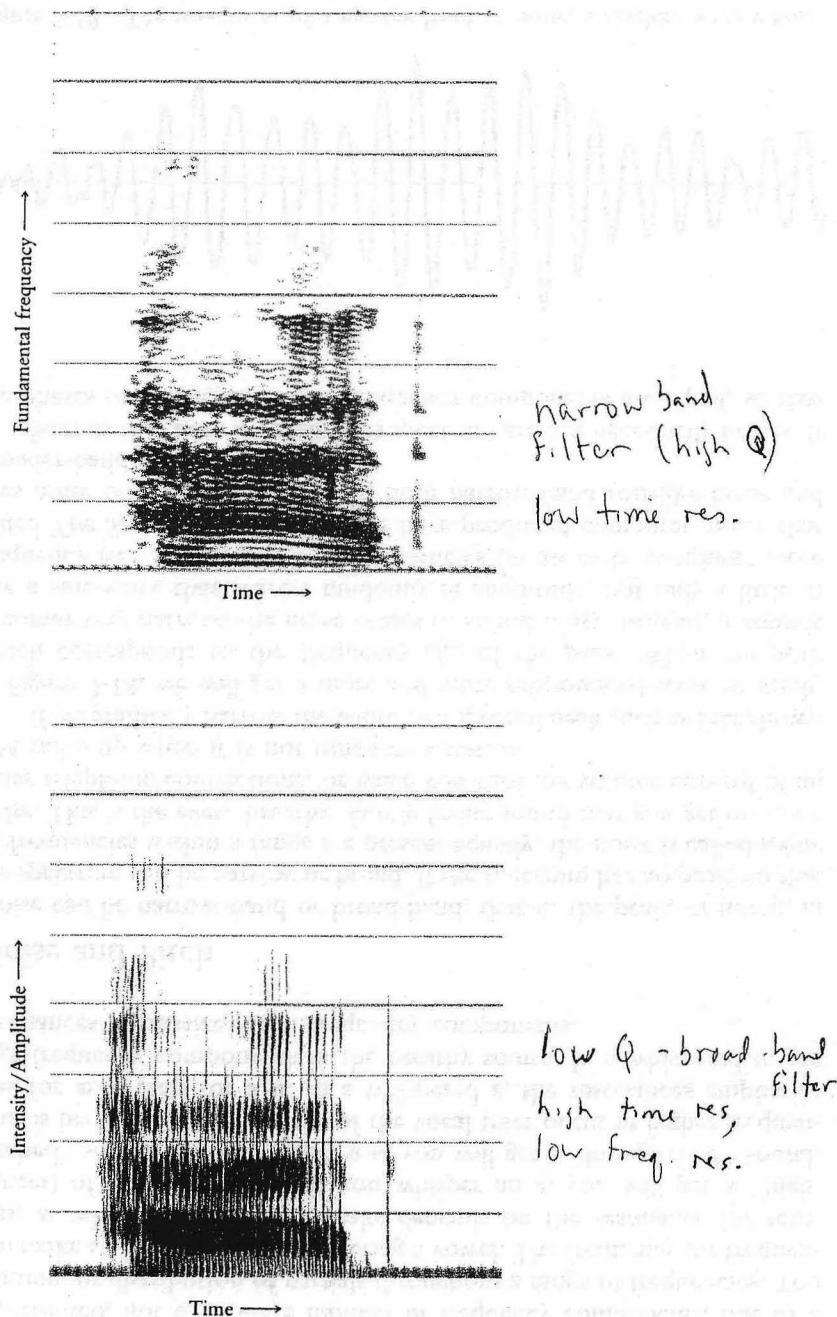


Figure 3-17 Sonograms of a human voice, saying the vowel *a* in "had."

detail, we can't depict frequency in fine detail. This is why Figure 3-15 is only qualitatively (not quantitatively) correct, for it seems to show both frequency and time very accurately.

Sonograms are usually used to get an approximate idea of how amplitude or power varies with frequency and time. However, if we take successive spectra of properly overlapping portions of a waveform in the right way we can reconstruct the waveform from the successive spectra. This is easier if in taking the successive spectra we compute the phases of the sinusoidal components as well as their amplitudes.

A device that is built or programmed for such analysis into successive spectra and resynthesis into the original waveform is called a *phase vocoder*. Besides being used in studying musical sounds, phase vocoders are used to alter sounds in useful ways by modifying the spectral representation before resynthesis. Phase vocoders can be used to shift the pitch of sounds, or to stretch them out or compress them in time, or both at once.

Although no musical sound is truly periodic in persisting unchanged forever, most musical sounds are nearly periodic. They can be approximated very closely by a fairly small number of sinelike waves whose amplitudes rise and fall slowly with time, and whose frequencies are nearly harmonically related; so they can be represented by a succession of changing line spectra. The vibration of a piano string dies away slowly; in its line spectrum, the peaks that represent the various harmonics decrease in height as the intensity of the sound decreases. Bells and gongs (see Chapter 13) do not produce periodic sounds, but they do have line spectra that decrease in amplitude gradually as the sound dies away.

Suppose we measure the spectrum of a sound and get a smooth curve rather than a series of spikes, as in Figure 3-18. Such a sound can be

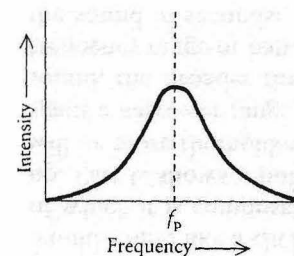


Figure 3-18 Spectrogram of a noise that is made up of an infinite number of sinusoidal frequency components whose phases are random and therefore don't peak at the same time. Instead, they give the bell-curve distribution familiar from statistics. A noise with a spectrum like this one will give a sense of pitch at the frequency f_p , at which the spectrogram peaks.

represented, not by a finite number of frequency components, but by a continuous distribution of partials throughout a range of frequencies. You can make such a sound by whispering a vowel. The frequency (or frequencies) at which the spectrum peaks depends on the resonance (or resonances) of the vocal tract. If you whisper an *e*, you will get a “high-pitched” sound; if you whisper a *u*, you will get a “low-pitched” sound. This is because the resonances of the vocal tract occur at higher frequencies for an *e* than for a *u*. In a whispered *e*, the resonances emphasize high-frequency components in the breathy sound. In a whispered *u*, the resonances emphasize lower-frequency components.

Noise and Pitch

Noise can be narrow-band or broad-band, that is, the peak, or hump, in the spectrum can be narrow or broad. If the spectrum has no peak, so that all frequencies within a range are present equally, the noise is called *white noise*. This is the even, breathy, faintly frying sound that you get on some noisy telephone connections, or when you turn the volume control of an AM radio up when it is not tuned to a station.

If we gradually narrow the width of a spectral peak such as that shown in Figure 3-18, we will get a more and more pronounced sense of pitch, which corresponds to the frequency (f_p) of the peak. When the peak becomes very narrow, the noise ceases to sound noisy. Instead, it sounds like a sine wave that wavers randomly in amplitude, but only a little in frequency (see Figure 3-19). Max Mathews (in an early computer piece called *The Second Law*) and others have produced computer music that uses noise of various bandwidths, both narrow-band tonelike noise and broader-band shushing noise.

Sounds that have a continuous spectrum are not necessarily noises. If the phases of all the sinusoidal frequency components are equal, so that

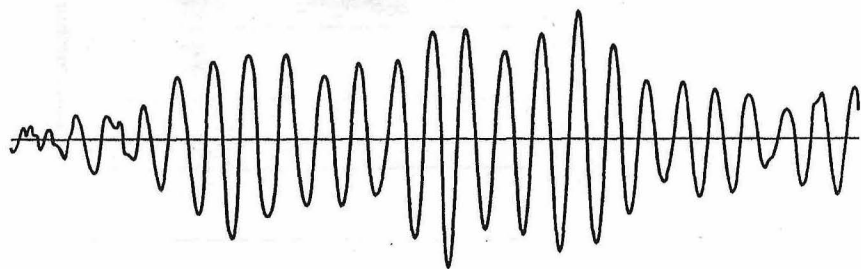


Figure 3-19 The waveform of a narrow band of noise, a sinelike wave whose amplitude varies randomly with time, but whose frequency is nearly constant.

they are peak at the same instant, or if the phases change slowly and smoothly as the frequency changes, we get a single pulse rather than a chaotic, persistent noise.

Part A of Figure 3-20 shows a short pulse or burst of pressure that would sound like a click, and part B shows that the sinusoidal components of which it is composed span a broad range of frequencies, from zero on up. Part C shows a pulse that has a broader rise and fall of amplitude. It will be more thumplike than clicklike, and part D shows that its spectrum spans a narrower range of frequencies. In general, the shorter the pulse of sound, the broader the range of frequencies. Roughly, the width of the frequency range or band (B) is inversely proportional to the duration (D) of the sound in seconds:

$$B = 1/D$$

Most short sounds aren't simple rises and falls in pressure like those in this figure. When we strike a block of wood, we get a sort of click, but the

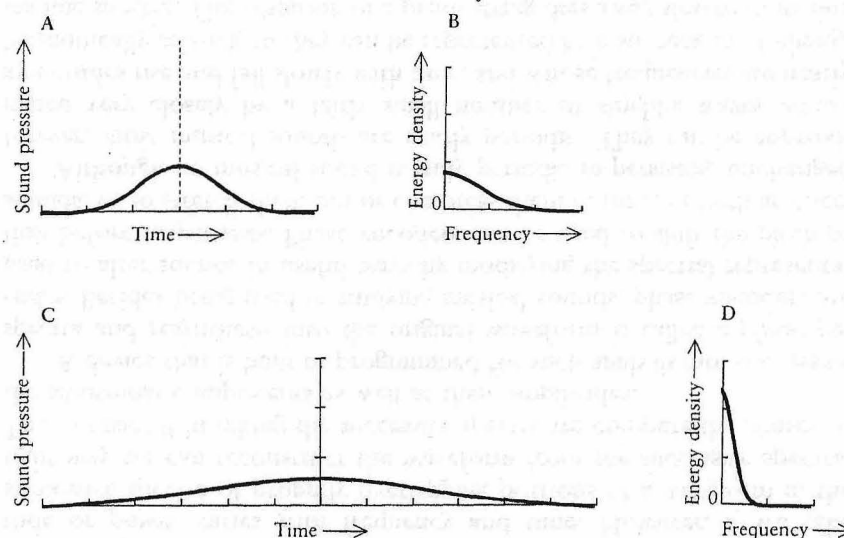


Figure 3-20 The spectrum of a single pulse of air pressure is not a line spectrum, but a continuous spectrum made up of components of a continuous range of frequencies. If the pulse is short (part A), the spectrum is broad (part B), and the pulse sounds like a click. If the pulse is longer (part C), the spectrum is narrower (part D), and the pulse sounds like a thump.

pressure tends to oscillate up and down for a short time. Similarly, when we strike a bass drum, we get a sort of thump, but the drumhead oscillates for a short time. The bandwidth, or range of frequencies, of such sounds is still proportional to the duration of the sound, but the peak of the spectrum is at the frequency with which the struck object oscillates: the larger the object, the lower this frequency of oscillation, and the duller the sound. For example, in part A of Figure 3-21, we have a short section of a waveform, two cycles long, and in part C a longer section, four cycles long, twice the duration. Such short sections of sine waves give a sense of whether the sound is “dull” or “bright.” In parts B and D, the peak of the spectrum is at the frequency of f_0 of the sine wave, but the spectrum of the longer, four-cycle section is half as wide as the spectrum of the short, two-cycle section.

When a sine wave is turned off or on abruptly, as shown in Figure 3-21, this excites resonant devices of many frequencies. Indeed, we hear a click when the wave is turned on or off abruptly. Anyone who listens to the output of an audio oscillator, and makes and breaks the connection

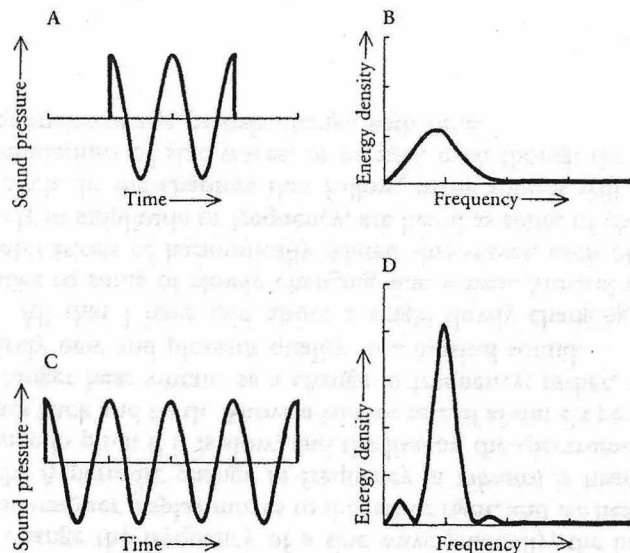


Figure 3-21 A short section of a sine wave of frequency f_0 , as in part A, has a broad spectrum centered on the frequency f_0 , as in part B. A longer section of the same sine wave, as in part C, has a narrower spectrum, as in part D, still centered on the frequency f_0 . When the sine wave is turned on abruptly, as shown in parts A and C, we hear a click.

with the speaker by flipping a switch, will hear such a click. We can get the effect of hearing a short pure tone or sine wave without a click by increasing and decreasing the intensity gradually, as shown in Figure 3-22. If the slowly rising and falling sine wave has a total of about 40 or more complete cycles, it sounds like a short burst of tone without a click. If it has 4 cycles or fewer it sounds like a click without a clear pitch, though, like a noise, it can sound “high” or “low,” or “bright” or “dull,” depending on the frequency of the sine wave. For intermediate numbers of cycles, we hear both a click and a pitch. The click part of the sensation becomes more prominent as the number of cycles is reduced, and the sensation of clear pitch fades into just “high” or “low,” or “bright” or “dull.” The appearance of the click occurs at a somewhat smaller number of cycles for sounds heard in a reverberant room than for sounds heard through headphones.

The practical result here is that we can add “slowly changing sine waves” to produce periodic, pitched sounds of finite duration. Mathematically, such a “slowly changing sine wave” is not a sine wave at all, but it approximates many of the properties of a sine wave. Both instruments for measuring and the human ear respond to such a “slowly changing sine wave” as if it were a pure tone that changes slowly in amplitude, frequency, or phase. However, if we change sine waves too fast, we will hear a click or the twang of a plucked string.

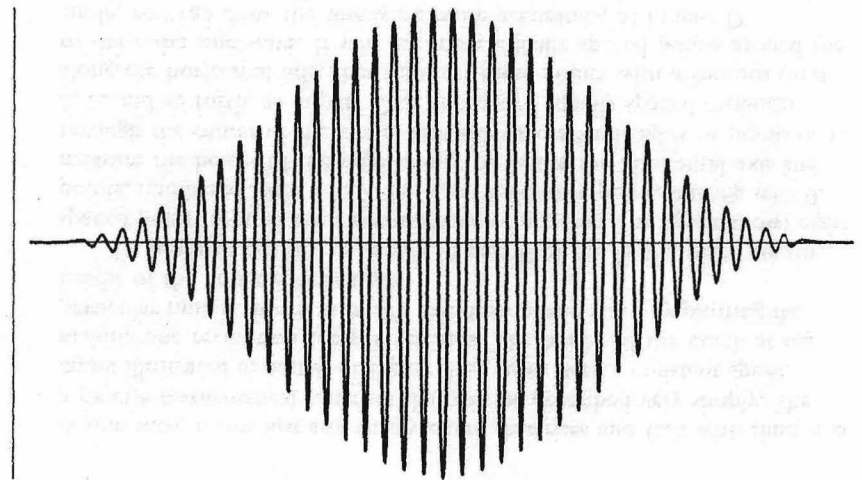


Figure 3-22 If a sine wave is turned on and off slowly, we don’t hear a click. (Mathematically, this is not a sine wave but, for practical purposes, it sounds like one.)

On the screen of a spectrum analyzer, a pure tone of slowly increasing amplitude appears as a line or peak of increasing height, and the corresponding sound grows louder and louder without changing pitch (much). If we change the frequency of a sine wave gradually, the line on the spectrum-analyzer display moves to the left or right, and we hear a fall or rise in pitch. A periodic change in frequency (a *vibrato*) is heard as a periodic change in pitch if it is slow, and the line on the spectrum-analyzer display jiggles back and forth. Above a vibrato rate of about six per second, we can no longer hear vibrato as a change in frequency; rather, vibrato gives an entirely new and pleasant quality to a musical sound.

All that I have said about a single slowly changing sine wave also applies to sums of slowly changing sine waves. Musical tones which are combinations of harmonically related sine waves, each of which changes slowly in amplitude or frequency, are heard as tones of changing loudness or pitch. In the chapters that follow, many sounds will be described as combinations of sine waves, or partials, even though the amplitudes and frequencies of the partials change with time.

A Note on Sine Waves

A sine wave is not just any wiggly curve that rises and falls with time; it is a precise mathematical function that can be described very simply. The crank illustrated in part A of Figure 3-23 turns with a constant speed, making one revolution every T seconds. The height of the crank at any particular time is shown as h . We can draw a sine wave by plotting the height of the crank against time.

To begin to do this, as shown in part B of the figure, place equally spaced points around the circumference of a circle. The figure shows eight points, numbered from 0 to 7. For each successive point, starting with 0, measure the height of the point above or below the horizontal axis line through the center of the circle. Plot the successive heights of points 0, 1, 2, 3, and so forth, as in part C, at successive equally spaced distances along the horizontal line; then connect these points with a smooth curve to draw the sine wave. If you use more equally spaced points around the circle, you can draw the sine wave more accurately, as in part D.

Mathematically, the relation between h (or amplitude) and time t is written $h = \sin \omega t$, in which ω is a constant, equal to 2π times the frequency f , and $f = 1/T$, T being the period. The higher the frequency (that is, the shorter the period), the more times the sine wave goes up and down each second.

There is another type of sine wave, the *cosine*, which is simply a sine wave with a different phase. The cosine of ωt is written $h = \cos \omega t$. The cosine reaches its peak a quarter of a period ($T/4$) before the sine does. The cosine is traced out by the height of the dashed crank in part A of Figure 3-23, which is at right angles to the solid crank, and is shown as the dashed curve in part C. To draw the cosine we simply start at point 2 instead of point 0, and plot successively the heights of points 2, 3, 4, 5, 6, and so forth.

This simple "mechanical" picture can tell us all we need to know about sine waves. Those who like mathematical puzzles can work out all sorts of interesting, useful, and important relations, such as

$$(\sin \omega t)^2 + (\cos \omega t)^2 = 1$$

A particularly interesting and important relation is

$$(\sin pt)(\sin \omega t) = (1/2) [\cos (\omega - p)t - \cos (\omega + p)t]$$

Here we can think of the sine wave $\sin pt$ as controlling the intensity or amplitude of the sine wave $\sin \omega t$. What can we discover from this simple

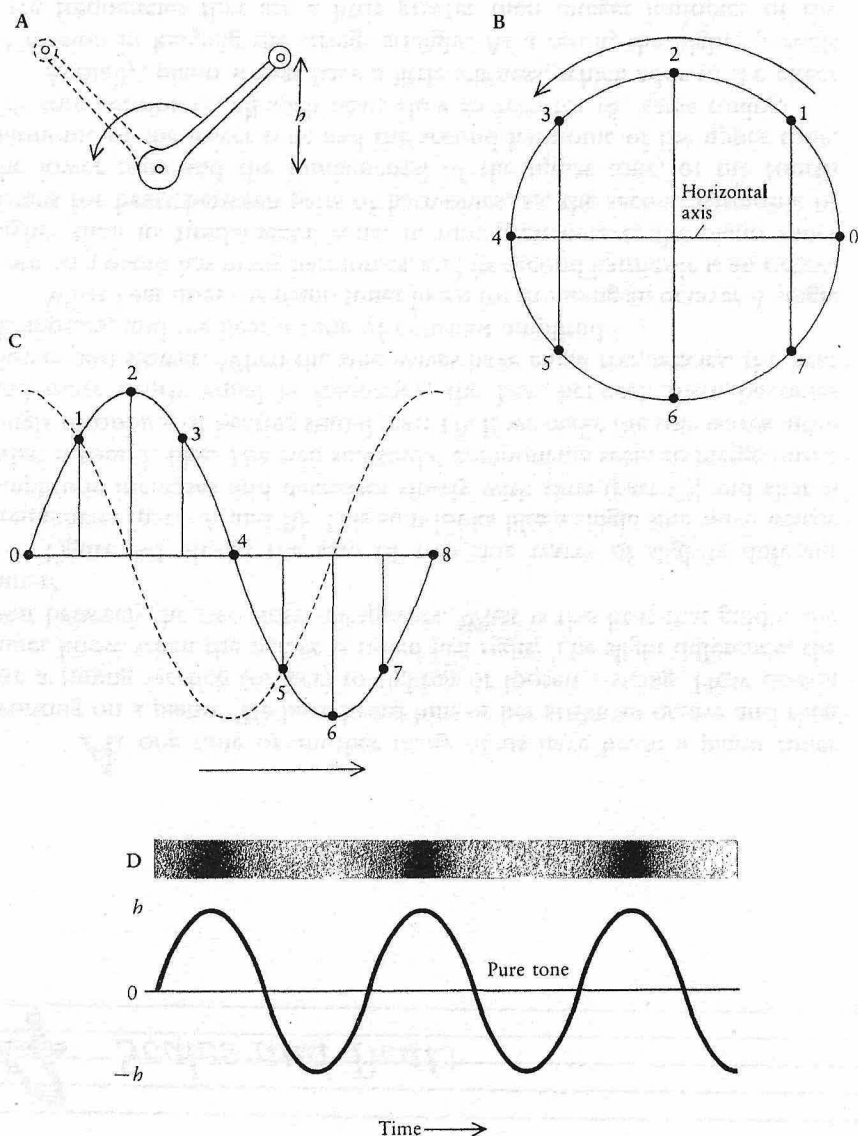


Figure 3-23 A. The height, h , of the handle of a crank that rotates at a constant speed varies sinusoidally with time. B. Equally spaced points around the circumference of a circle. C. When the points in part B are translated to a graph in which the steady passage of time is represented as movement to the right, they outline a sine wave. D. A smooth sine wave would be produced if every point on the circle were translated to the graph.

relation? First, we see that the product of two sine waves is simply two sine waves (since cosine waves are just sine waves that peak at different times, that is, have different *phases*). The average value of the product of two sine waves is zero, because the average value of each of the two sine waves is zero. If the sine waves have exactly the same frequency, one of the two components on the right is the cosine of zero, which is equal to one. In mathematically representing a function of time as a sum of sine waves, we make use of this property.

We can deduce another important result from this last formula. Imagine that p is very small compared with w , say, $p = 1$ and $w = 100$. Then the factor $\sin pt$ will cause the amplitude of the oscillation $\sin wt$ to vary slowly with time. Our formula tells us that

$$(\sin t)(\sin 100f) = (1/2)(\cos 99t) - (1/2)(\cos 101t)$$

That is, a slowly varying sine wave doesn't have only one frequency; mathematically it is made up of two or more frequencies, but these frequency components are very near the nominal frequency of the rapidly varying sine wave. This is very important, for when we turn a sine wave on or off, we still think of it as having the same frequency that it would if it persisted forever. In practice, to our sense of hearing, this is true *as long as we turn the sine wave on and off slowly enough* to avoid getting a click.

4 Scales and Beats

At one time or another many of us have heard a piano tuner working on a piano. We have heard him or her strike an octave and then use a tuning wrench (or key) to tighten or loosen a string. How does a tuner know when the octave is tuned just right? The slight difference, the *beat* between the two notes disappears. What is this beat that guides the tuner?

Figure 4-1 shows the sum of two sine waves of slightly different frequencies (parts A and B). This sum looks like a single sine wave whose amplitude increases and decreases slowly with time (part C), and that is what it sounds like. The two sinusoidal components seem to merge into a single throbbing or beating sound (part D). If we make the sine waves more and more nearly equal in frequency, the beat between them becomes slower and slower. When the sine waves have equal frequencies, the beat disappears, and we hear a tone of constant amplitude.

What beat does the piano tuner listen for in tuning an octave? A single note on a piano has many harmonics, and its second harmonic is an octave higher than its fundamental tone. In tuning an octave, the piano tuner listens for beats between pairs of harmonics, as, the second harmonic of the lower tone and the fundamental of the upper tone, or the fourth harmonic of the lower tone and the second harmonic of the upper tone. For true harmonics, all such beats slow to zero for the same tuning.

Actually, piano strings have a little stiffness, which adds to the effect of tension in keeping the strings straight. As a result, the higher partials have frequencies that are a little greater than integer multiples of the

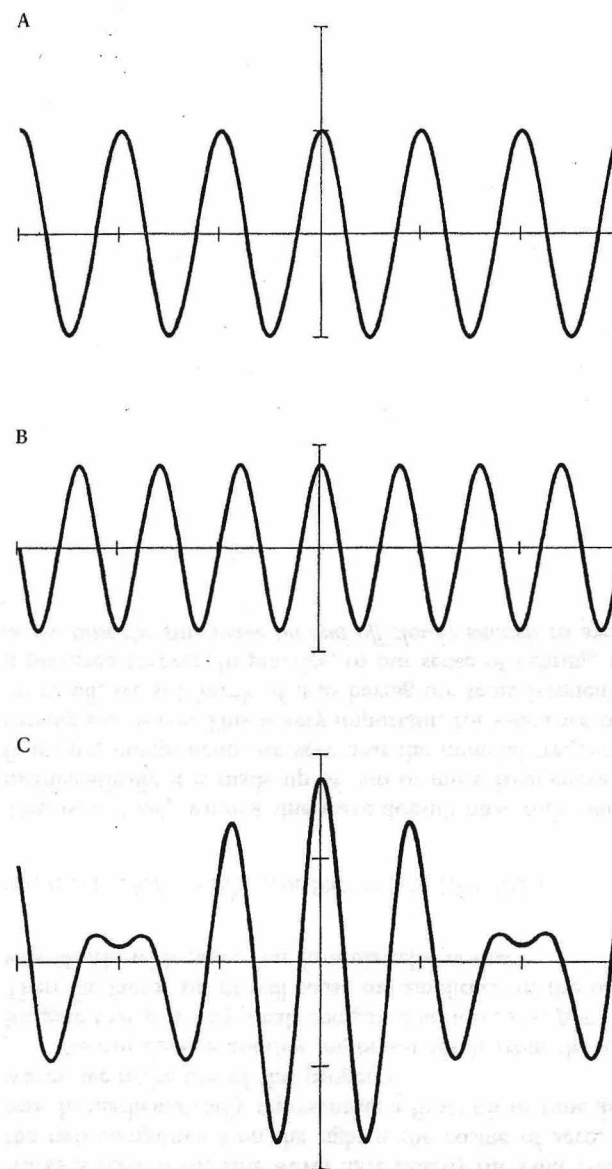


Figure 4-1 The phenomenon of beats. The amplitudes of the two sine waves in parts A and B, whose frequencies are very close to one another, add up to give the wave in part C, which looks very much like a sine wave of slowly varying amplitude.

Figure continued on following page.

D

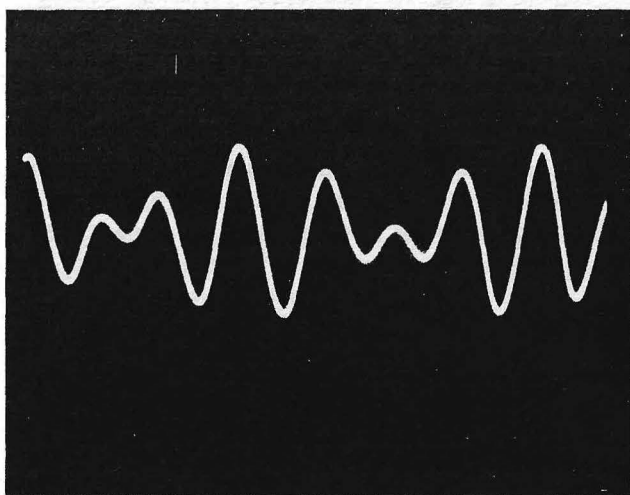


Figure 4-1 (Continued) The photograph of an oscilloscope screen in part D shows the beat between tones of 4,000 and 4,100 Hz. When the frequency difference is small, it sounds like a sine wave or pure tone whose intensity rises and falls slowly with time. As the frequencies of the sine waves are made more and more nearly equal, the rate at which the amplitude rises and falls (the frequency of the beat, or the beat frequency) gets slower and slower until, when the frequencies are equal, the beat disappears.

fundamental frequency f_0 ; that is, they are slightly larger than $2f_0$, $3f_0$, $4f_0$, $5f_0$, and so forth.

When octaves in a piano are tuned by the method of eliminating beats, the octaves will therefore be *stretched* a very little; that is, they will have frequency ratios slightly greater than 2. Pianos *are* often tuned with slightly stretched octaves, sometimes because of the stiffness of the strings, sometimes because the pianist prefers the brighter timbre that results. In this discussion, we will ignore the stiffness of piano strings and other practical realities, and assume that *all* musical tones have partials that are exactly harmonic, that is, that are integer multiples of the frequency of the first partial.

Frequency Ratios

Intervals other than octaves can also be tuned by means of beats. To see how, consider the piano keyboard shown in Figure 4-2. We go up a *semitone* when we go from any key (white or black) to the next key (black

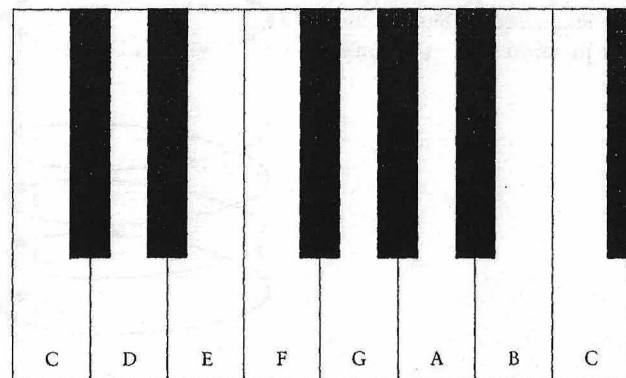


Figure 4-2 The white keys of the piano give the seven notes of the C-major diatonic scale.

or white). Certain intervals are called *consonant intervals*, because musicians and listeners consider pitches separated by these intervals to be pleasing when sounded together. These intervals are given in Table 4-1, which shows both the ideal frequency ratio between the two notes and the number of semitones of difference between them on the piano keyboard.

Figure 4-3 shows the harmonics of a C with a frequency f_0 and the G above it, which has a frequency $(3/2)f_0$. We see that the third harmonic of C has the same frequency, $3f_0$, as the second harmonic of G. If the C and G

Table 4-1 Consonant Intervals.

Name of interval	Notes (in Key of C Major)	<i>Just</i> Ideal Frequency Ratio	Number of Semitones
Octave	C-C	2	12
Fifth	C-G	$3/2$	7
Fourth	C-F	$4/3$	5
Major third	C-E	$5/4$	4
Minor third	E-G	$6/5$	3
Major sixth	C-A	$5/3$	9
Minor sixth	E-C	$8/5$	8

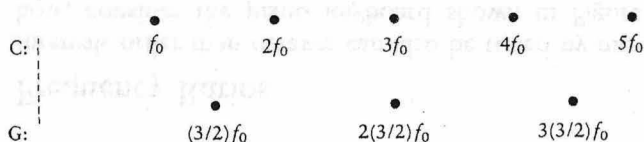


Figure 4-3 The frequencies of two pitches (such as C and G) that are an interval of a fifth apart are always in the ratio of 100:150. Their harmonics will be in the same ratio; so the second, fourth, and so forth, harmonics of G will coincide with the third, sixth, and so forth, harmonics of C.

are a little mistuned, these harmonics will produce an audible beat when the two notes are struck together. By tuning G so that it doesn't beat with C, we can assure that the fundamental frequency of G is just $3/2$ that of C, in musical terms a perfect fifth.

A little thought will show that other pitches can be tuned in this way as well. A fourth has a frequency ratio of $4/3$, so the third harmonic of F should have the same frequency as the fourth harmonic of C (which is C'' , two octaves above the first C).^{*} A major third has a frequency ratio of $5/4$, so the fourth harmonic of E should have the same frequency as the fifth harmonic of C. A minor third has a frequency ratio $6/5$, so the fifth harmonic of E should have the same frequency as the sixth harmonic of C (a G).

A friend of mine once tried to tune his piano according to such relations, thinking that by later adjustments he could get to the *equal-tempered* scale, to which a piano is really tuned. He underestimated the difficulty. Piano tuners have a systematic method of tuning, in which they tune the intervals within a single octave to have prescribed numbers of beats per second. Once they have tuned these twelve notes, they tune all the other notes in octaves above and below by the method of octaves, with no beats. This isn't what my friend did.

Here we come to the dilemma of the diatonic, or major, scale, represented by the white keys of the piano. Is there a *sensible* explanation for this scale, which is used in so many cultures? There *is* an explanation, but you must judge for yourself how sensible it is.

^{*}The primes are used to flag *relative* positions of notes. As we start up, we have the notes C, D, E, F, G, A, B; in the next octave up the notes are C' , D' , E' , F' , G' , A' , B' ; in the next, C'' , D'' , and so forth. This convention helps us keep track of *which* C we are talking about as we jump up and down the octaves in a discussion of harmonics, say. This designation is *relative*, not absolute; any note we choose can be the starting point for the primes in a given discussion.

First, we should observe that notes or tones an octave apart sound very similar. In primitive cultures, men and women sing in octaves without realizing that they are singing different notes. With some timbres, it is not uncommon to make an error of an octave in judging pitch. The psychologist Roger Shepard has likened the similarity of octaves to points on an ascending spiral. As shown in Figure 4-4, the notes C, D, E, F, G, A, B repeat again and again around the spiral as we go up the scale. But C' an octave up always lies close (in perception) to C an octave below, and so it is for all the other notes of the scale.

With this in mind, consider Figure 4-5, which shows notes in the bass and treble clefs. In part I, we see the first six harmonics of the C an octave below middle C.^{*} They are, successively, C, (middle) C' , G' , C'' , E'' , and G'' . The intervals between these successive notes are consonant intervals. They are the octave (C to C'), the fifth (C' to G'), the fourth (G' to C''), the major third (C'' to E''), and the minor third (E'' to G'').

So far we have only three of the seven pitches of the C-major diatonic scale, C (the *tonic*), E (the *mediant*), and G (the *dominant*). A plausible way in which to get one more note is to drop down a fifth from C'' , as shown in part II of the figure. This gives us F' (the *subdominant*); the frequency ratio of F' to C' is $4/3$.

In going from C'' to E'' to G'' , as in part I, we go up first by a major third (a frequency ratio of $5/4$), then by a minor third (a frequency ratio of $6/5$), and so arrive at an overall frequency ratio between G and C of $3/2$.

^{*}It is called "middle C" because it is written on the line in the middle, between the treble and bass clefs.

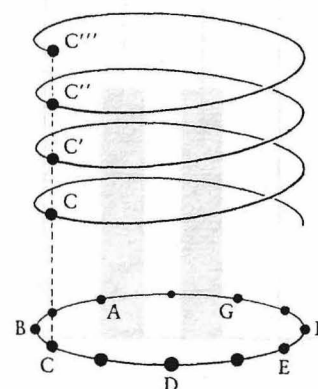


Figure 4-4 The notes of the diatonic scale as points along an ascending spiral. Here all Cs are close together and hence alike; all Ds are close together and hence, alike; and so on.

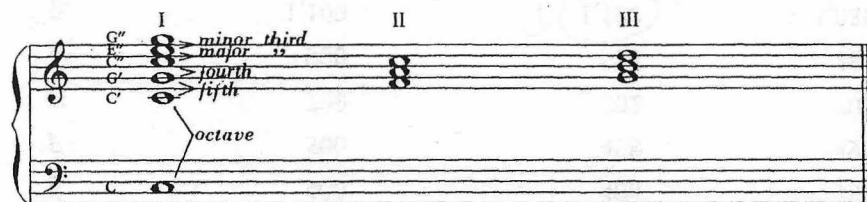


Figure 4-5 The successive harmonics of a note C give us the notes C', G', C'', E'', and G'', and include the intervals of the octave, the fourth, the major third, and the minor third. In any octave, if we go up a major third from G', we get B', and if we go up a minor third from B, we get D'. If we go down from C'' by a fifth we get F', and if we go up a major third and then a minor third from F' we get, successively, A' and C''. We have now derived all the pitches of the diatonic scale. In the notation shown, C' is middle C.

The fourth, or subdominant (F in the C major scale), plays an important part in Western music. It seems reasonable to go up from F' in the same pattern, first by a major third, then by a minor third, as in part II, thus basing a *major triad* on F'. This gives us A' (the *submedian*) and brings us back to C''. Here A' is a minor third below C'', that is $(6/5) A' = C''$ and, because $C'' = 2C'$, the frequency ratio of A' to C' is $5/3$.

The fifth, or dominant (G in the C major scale), plays an even more important role in Western music. Let us also base a major triad on G', again going up by a major third, then by a minor third, as shown in part III. This gives us the seventh, or *leading tone*, B', and the *supertonic*, D'.

We now have all seven pitches of the diatonic scale in *just intonation* in the key of C; C, being the first note of the scale, is the *tonic*.

Equal-Tempered Tuning

Musical intervals with frequency ratios given by the ratio of integers, such as $3/2$ (fifth) and $5/4$ (major third), are the basis of the diatonic scale. Such musical intervals are, of course, important in music itself. Let us concentrate on the black notes in the short passage of music shown in Figure 4-6. They are C, A, D', G, C. Let us designate the frequency of C as f_0 . In going up to A, we go to a frequency of $(5/3)f_0$. The interval from A up to D' is a fourth, with a frequency ratio of $4/3$; so when we arrive at D', the frequency should be $4/3$ that of A, which is $5/3$ times f_0 , or $(4.3)(5/3)f_0 = (20/9)f_0$. We now go down a fifth to G, which should give us a frequency of $(2/3)(20/9)f_0 = (40/27)f_0$. We now go down another fifth to C, which should give us $(2/3)(40/27)f_0 = (80/81)f_0$. But we started



Figure 4-6 Consider the black notes in this series of chords. They proceed from C to A to D' and back down to G to C, or by intervals of a sixth up, a fourth up, a fifth down, and a fifth down. If these were perfect intervals, in which the frequency ratio of the sixth is $5/3$, of the fourth is $4/3$, and of the fifth is $3/2$, we would not get back to the same pitch from which we started.

out by setting the frequency of C equal to f_0 , and we haven't gotten back there!

This short passage of music shows that no fixed scale of pitches can accommodate all upward and downward motions of pitch by ideal musical intervals, that is, musical intervals with frequency ratios given by the ratios of small integers. A number of failures to return to the initial pitch after such sequences of ideal intervals have been noted and named. They are shown in Table 4-2.

The comma of Pythagoras is of particular interest. If we go up repeatedly by the interval of a perfect fifth ($3:2 = 1.5$) we will reach octaves of

Table 4-2 Final "Error" in Sequences of Ideal Intervals That *Almost* return to Initial Pitch.

Name	Ratio of Difference	Difference in Cents	Motions of Pitch
Comma of Didymos, or synotic comma	$81/80$	21.5 cents	Up four fifths, down two octaves, down one major third
Schisma	$32,805/32,768$	2.0 cents	Up eight fifths, up one major third, down five octaves
Comma of Pythagoras	$521,441/524,288$	23.5 cents	Up twelve fifths, down seven octaves
Diesis	$128/125$	41.0 cents	Up one octave, down three major thirds

$$3^{12} = 531,441 / 2^{19} = 524,288$$

all the white and black keys of the piano keyboard—almost. If we go up repeatedly in a ratio close to that of a perfect fifth, a ratio approximately 1.4983, we do reach the octaves of all black and white keys, and we can get the pitches of all other keys by subsequently going down or up by octaves. This may be thought of as the source of equal-tempered tuning.

In a very real sense, perfect intervals, whose frequency ratios are the ratios of small integers, are the very foundation of music. These intervals derive from the harmonics present in musical tones. They are important to the ear. Yet they cannot be exactly realized in all the intervals of any one scale, any one system of tuning. The history of Western music is also the history of the attempts to resolve these discrepancies.

In discussing departures of intervals of a scale from ideal intervals, the *cent* is very useful. The cent is a hundredth of a semitone, or $1/1,200$ of an octave. The frequency ratio of the cent is the $1/1,200$ power (the 1,200 root) of 2, or 1.00057779. If we multiply 1,200 of this decimal number together (that is, if we raise it to the 1,200 power), we get 2, the frequency ratio of the octave. If we raise the number to the 100 power, we get the frequency ratio of the equal-temperament semitone, 1.059463.

Table 4-3 shows the intervals of the diatonic scale in cents for equal-tempered tuning, for *Pythagorean* tuning (which is designed to give perfect fifths), and for *just* tuning. Table 4-4 shows the “error” in these compromise tunings relative to the ideal ratios.

Some musicians love just temperament dearly. For example, Harry Partch built many strange instruments whose sounds were as wonderful as their names, and which were often of somewhat uncertain pitch. He had a

Table 4-3 The Intervals of Differently Tempered Scales, Measured in Cents.

Note	Equal-tempered	Pythagorean	Just
C	0	0	0
D	200	204	204
E	400	408	386
F	500	498	498
G	700	702	702
A	900	906	884
B	1,100	1,100	1,088
C	1,200	1,200	1,200

Table 4-4 Errors in Tempered Scales Relative to Ideal Ratios.

Interval	Name	Equal-tempered Error	Pythagorean Error	Just Error
C-E	Major third	+15 <i>+14</i>	+22	0
D-F	Minor third	-16	-22	-22
C-F	Fourth	+2	0	0
C-G	Fifth	-2	0	0
C-A	Sixth	+16	+22	0
D-A	Fifth	-2	0	-22

harmonium justly tuned in the key of C. It sounded excellent in C, but dreadful when played in any other key.

Let us consider just temperament closely. Figure 4-6 shows that in just temperament the musical interval of a fifth from D to A is in error by 22 cents; that is why we are able to get back to C in just temperament, but at a considerable cost. This 22-cent error is 6 cents larger than any error in equal temperament.

Furthermore, in equal temperament we have dealt very simply with the problem of sharps and flats. Ideally, in the interval of the minor sixth, *Ab* (A-flat) should have a frequency ratio of 1.6 to be a perfect major third from the C above it. In order to make the interval *E-G#* (G-sharp) a just major third, *G#* should have a frequency ratio $25/16 = 1.5625$. In our equal-tempered scale, both *Ab* and *G#* are represented by the same black key on the piano, and by a frequency ratio of 1.5874.

There are flats and sharps in diatonic scales in which the tonic or first note is not C. A fifth above C is G, the first note of a closely related key, G major. G major has six notes in common with C major. (The key of G major employs *F#* instead of F.) A fifth above G is D. The scale starting on D has two sharps, *C#* and *F#*, and has only five notes in common with the C scale.

Going up through a cycle of successive fifths on the piano, we reach all keys, white and black, and on each we can build a scale. Or we can start down from C by an interval of a fifth, to reach F. The scale in which F is the first note, or tonic, has six notes in common with the C scale; in F, *Bb* is used instead of B.

Johann Sebastian Bach was one of the first strong advocates of a tuning such that the intervals, though slightly “in error,” are tolerable in all

keys. Bach's *Well-Tempered Clavier* includes pieces in all keys. With equal-tempered tuning, all are equally in tune (or out of tune).

It is often argued that equal temperament can be offensive to musicians with keen ears. During the nineteenth century instruments with complex, almost unplayable keyboards were devised in order to allow the production of close approximations to ideal intervals; some are discussed in Helmholtz's book. But such keyboards cannot cure the inevitable wandering in pitch when one goes up and down by ideal intervals only.

It is comparatively easy to hear in sustained chords that the thirds of equal temperament are "out of tune," because they are slightly higher than just tuning. However, experiments by Max Mathews show that some people like major triads with out-of-tune thirds; they sound "brighter." Good musicians can distinguish equal temperament from just tuning in listening to the successive notes of a scale. But the tuning of the piano is acceptable to most, and violinists in orchestras ordinarily seem to approximate just intervals no more closely than the piano does.

Sometimes just intervals are important, as in some passages in early choral music, and in barbershop quartets. Good singers will sing just intervals in such passages, and maintain the initial key through using unjust intervals where these are acceptable.

Tonal Centers of Scales

There is one other important characteristic of a scale. On what pitch does it start?

The diatonic C-major scale, represented by the white keys of the piano, consists of seven notes, or eight if we span a whole octave. Within the octave of the major scale, five of the intervals are whole tones (two semitones), and two (the third and the seventh) are semitones. The succession of intervals that we will encounter in going up the scale depends on what note we choose as the beginning of the scale. Sharps or flats (called chromatics) are added to preserve the sequence of whole tones and semitones, since this sequence defines the major scale.

Classical Greek music was based on a mathematically derived scale made up of tetrachords, groups of four descending pitches spanning a perfect fourth. The two middle pitches of each tetrachord could be chromatically altered; the unaltered version was referred to as diatonic. The entire collection of tetrachords, plus one added pitch at the bottom, formed a single two-octave scale. It probably always had as its center tone the A below middle C.

The (unrelated) medieval Church modes (sometimes erroneously referred to as the "Greek" modes) were based on ascending diatonic octaves. The tonic center of a mode is always the lowest pitch in the octave. The

modes were codified in a process begun by Pope Gregory I in an attempt to standardize the great number of chants that had accumulated in early Christian liturgy (hence the name Gregorian chant). The Church modes flourished from about 800 to about 1500.

Although the Church modes largely fell into disuse in about the seventeenth century, two of them—which we now call the major and minor modes or scales—survived and became the tonal basis of Western European music. Some modern composers have written works or parts of works in a neomodal style. Various other types of chromatically altered scales are the basis of certain types of European folk music, most notably Eastern Europe and gypsy music.

If we start at C on the piano and play the white keys to C', we play a major scale, in which the pattern of whole tones (1) and semitones ($1/2$) is 1, $1/2$, 1, 1, 1, $1/2$.

If we start at A and play the white keys to A', we play a minor scale, in which the interval pattern is 1, $1/2$, 1, 1, $1/2$, 1, 1. There are two other variants on the minor scale as well, in which the last three intervals in the pattern can be 1, 1, $1/2$, or even $1/2$, $1 1/2$, $1/2$ produced by using (in the key of A minor) F#, or G#, or both, instead of F and G.

We may observe that the major scale (think of the scale that starts on C) has major triads on the tonic, the fifth, and the fourth, all important tones of the scale. In the first minor scale mentioned (think of the white keys starting on A) the triads on the tonic, the fifth, and the fourth are all minor triads, a sort of complement of the case of the major scale.

In this book, the central concern is with the science of musical sound. This chapter has noted that the musical intervals produced by instruments can be tuned by listening for beats between the harmonics of two notes that are sounded simultaneously. We have seen that one cannot choose the frequencies of a seven-tone (or twelve-tone) scale in such a way that the ratios of the frequencies of the notes of the intervals are ratios of integers. We have seen that equal-tempered tuning, based on semitones of frequency ratio 1.059463, gives excellent approximations to ideal intervals, and that these approximations are the same for scales in all keys.

When we venture beyond the matters of intervals and tuning to discuss modes, we are wandering a little beyond the science of musical sound into the territory of music itself. This we cannot wholly avoid. But we will concern ourselves only with the two surviving modes: the major scale, because it is important to Rameau's ideas concerning harmony; and the minor scale, because it is hard to avoid.

5 Consonance and Dissonance

Melody is characteristic of almost all music. Strangely, so is the diatonic scale (or some subset of it), with intervals based on frequencies in the integer ratios $2/1$, $3/2$, $4/3$, $5/4$, $6/5$, and so on. No doubt, musicians of many ages and cultures have plucked strings simultaneously and tuned their instruments by observing beats, as piano tuners do today. However, sounding notes together in what we conceive of as harmony is peculiar to Western music.

The notes of one instrument or several can be sounded together for a variety of purposes. In giving instructions for the representation of cannon shots in his piece *Les Caractères de la guerre*, written in 1724, François Dandrieu recommended that the harpsichordist “strike the lowest notes on the keyboard with the entire length of the hand.” By 1800, another French composer recommended striking the lowest three octaves with the flats of both hands in his rendition of the sound of a cannon.

Whatever we think of as the essence of harmony, it does not lie in such sounds, or in the near-tone clusters of Percy Grainger, or in the tone clusters of Henry Cowell, or even in the haunting reminiscence of church bells that Charles Ives evokes in his *Concord Sonata* by instructing the pianist to strike simultaneously many black keys, softly, with a bar of wood (see Figure 5-1).

We think of harmony as notes that are sounded together smoothly and sweetly, or as rough-sounding or dissonant combinations of notes, full of tension, that miraculously resolve into a succeeding consonant chord. Beyond this, we think of well-known (and sometimes well-worn) progressions from chord to chord that serve as phrases or words of music. We

The image shows a page from the score of Charles Ives's *Concord Sonata*. The score is written for piano and includes various dynamic markings (pp, mp, mf, ff, p, pp, ppp) and tempo markings (Slowly, faster, somewhat faster, tempo). The music features complex, dense clusters of notes, particularly in the right hand, which are played with a stick. The score is written for piano and includes various dynamic markings (pp, mp, mf, ff, p, pp, ppp) and tempo markings (Slowly, faster, somewhat faster, tempo).

Figure 5-1 Page from score of Charles Ives's *Concord Sonata*, showing instructions for playing black keys with a stick.

think of modulation, the shifting from key to key, which is sometimes forthright, sometimes elusive or ambiguous.

About the harmonic language of music this book has little to say. Our concern will be chiefly with the consonance or dissonance one experiences when two, three, or more notes are sounded together. It is these experiences of consonance and dissonance that underlie the evolution of the musical theory of harmony.

Consonance, Dissonance, and the Critical Bandwidth

Of consonance, dissonance, and chords, there can be several views. One view might be that we are culturally conditioned to regard certain combinations of pitches as consonant, others as dissonant. However, composers of our century present us with an extraordinary range of combinations of notes, and with some rules for making such combinations. Can consonance and dissonance spring from nothing more than rules and customs? As will be seen, in a sense they can.

I believe, however, that to look at concepts of consonance and dissonance as arising from rules is to look at things the wrong way round. Rather, I am sure that the rules and customs are based on experiences of consonance and dissonance that are inherent in normal hearing. Of course, musical training and sophistication will color the subjective experience. The trained ear hears much in harmony that escapes the musically untrained. Sometimes the trained ear hears things that aren't there, for a musical friend of mine hears $G\sharp$ and $A\flat$ as different on the piano. All this we will consider later.

Let us return to the piano tuner of Chapter 4. Beats and "roughness" are phenomena that are crucial to consonance and harmony. We are already acquainted with beats. We know that the piano tuner tunes intervals for the absence of beats, or else for the number of beats per second necessary for the equal-tempered scale. In the nineteenth century, Helmholtz tried to explain consonance and harmony entirely in terms of beats. He thought that intervals were consonant if there were no (or few) beats between their partials. To explain dissonant intervals, he proposed that partials of different tones were so close together in frequency that the beating between them was perceived as dissonance.

The work of Rainier Plomp and of others at the Institute for Perception Research (IPO), which Jan Schouten founded in the Netherlands, has shown that this is too simple a view. Slow beats do not give a sense of dissonance, but merely a rising and falling of amplitude. Further, as we gradually separate the frequencies of two sine waves or "pure tones," we hear a disagreeable roughness even when the frequencies are so far apart that we no longer distinguish beats. The range of frequencies in which we

hear beats or roughness is called the *critical bandwidth* and is shown in Figure 5-2.

The critical bandwidth is an important experimental fact of hearing. To some degree, in listening to sounds we can tune in on a narrow band of frequencies, much as we tune in to a radio channel. When frequency components are separated by more than a critical bandwidth, we can hear them separately (this is called *hearing out*). But frequency components that lie within a critical bandwidth interact, and give us sensations of beats, roughness, or noise.

The critical bandwidth is important in the perception of loudness, in the perception that a sound is noise, and in the *masking* or hiding of one sound by another. In essence, the critical bandwidth results from the way that the ear resolves frequencies. At low and moderate sound levels, frequency components lying farther apart than a critical bandwidth send signals to the brain over separate nerve fibers, but frequency components lying within a critical bandwidth send a mixed signal over the same fibers.

The graph in Figure 5-3 shows the relation between critical bandwidth and consonance for two pure tones in another way. We can see from it that the maximum dissonance occurs at about a quarter of a critical bandwidth. With greater frequency separations, consonance also increases and becomes almost perfect for *all* separations greater than a critical bandwidth.

In order to know what frequency intervals between pure tones are consonant, we must know how the critical bandwidth varies with fre-

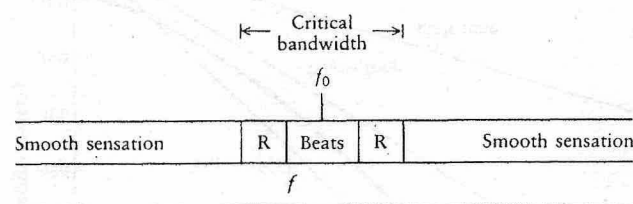


Figure 5-2 Pure tones (sine waves) that are close enough together in frequency give rise to audible beats. Even when the frequencies of such tones are too far apart for us to distinguish beats, we can still hear a certain "roughness." If the frequencies differ still more, we hear each tone separately and smoothly. Imagine one pure tone to have a frequency f_0 and another to have an adjustable frequency f . As we vary f from much below f_0 to much above f_0 we pass from smooth to rough to beats to rough to smooth again, as shown. The range of frequencies within which we hear roughness or beats is called the *critical bandwidth*.

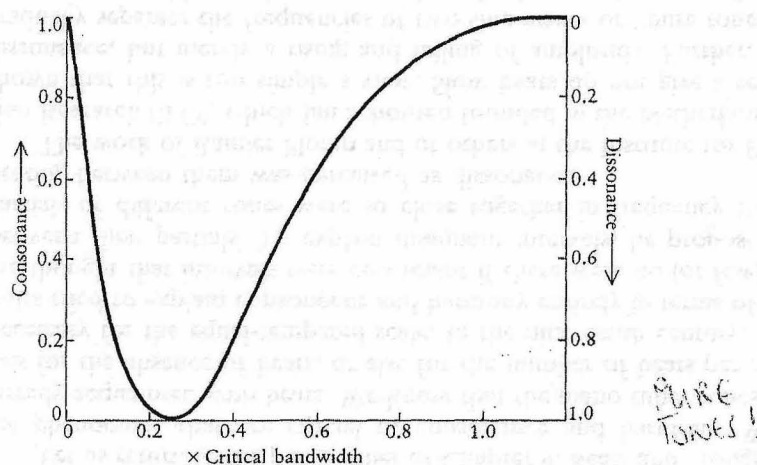


Figure 5-3 Plomp's curve representing the consonance of two pure tones sounded together as a function of their frequency separation (expressed as a fraction of their critical bandwidth). The tones are consonant if the frequency separation is so small that we hear slow beats, and consonant if the frequency separation is a critical bandwidth or greater, so that we hear the tones separately. The consonance and dissonance scales are chosen arbitrarily, to give consonance increasing from 0 to 1 (maximum) upward on the left, and, conversely, dissonance increasing from 0 to 1 downward on the right. The horizontal scale at the bottom represents frequency separation in fractions of a critical bandwidth. While this curve is derived from experimental data, its accuracy is not high.

quency. This is shown in Figure 5-4. Vertically it shows, for a critical bandwidth, the difference between the two frequencies that lie at its edges, f_1 and f_2 . Horizontally the graph shows the average value of two frequencies, that is, $(f_1 + f_2)/2$.

We see that, for most of the frequency range shown, the critical bandwidth lies between a minor third and a whole tone. For frequencies below 440 Hz (the A above middle C), the critical bandwidth is larger. Thus we might expect that, in order to be consonant, notes of low frequencies that are sounded together would have to be more widely separated than notes of high frequencies that are sounded together. Indeed, in piano music the notes of the chords in the bass are commonly more widely separated than the notes of the chords in the treble. However, for many musical sounds lower frequencies cannot be very important to consonance, since we hear mostly the higher partials. This is partly because

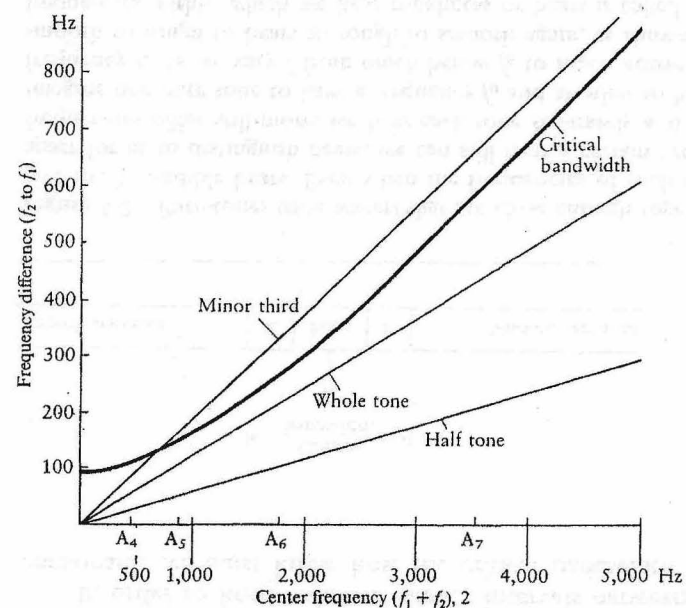


Figure 5-4 The width of a critical bandwidth as a function of frequency. Above a frequency of 500 Hz the critical bandwidth is roughly proportional to frequency. The frequency differences for three intervals are shown. A good approximation is that, if two pure tones are separated by a minor third or more, they will sound consonant together.

a good deal of the energy is in the higher partials (almost all when we listen to a transistor radio), and partly because the ear is more sensitive to higher frequencies than to lower frequencies.

A good rule of thumb is that pure tones less than a minor third apart are dissonant, but tones a minor third or more apart are consonant; so, for pure sine tones, *any* interval greater than a minor third will be judged as consonant, however odd the ratio of frequencies. But, of course, this isn't so for musical tones, such as piano tones, which have many harmonic partials. And musicians tend to judge the consonance or dissonance of pairs of sinusoidal tones by first recognizing the musical interval, and then calling the pair of tones consonant or dissonant on the basis of their past experience with pairs of nonsinusoidal tones and on the basis of what they have been taught.

Figure 5-5 shows the result of a calculation of the relative dissonance (or consonance) according to Plomp for two complex tones, each of which

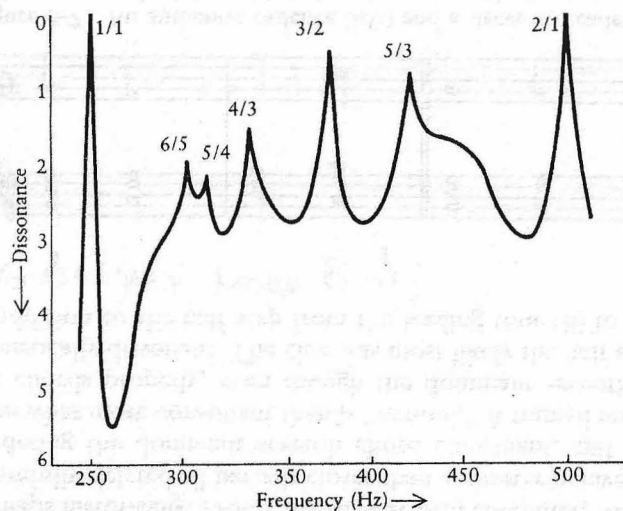


Figure 5-5 If we know the critical bandwidth, which depends on frequency, and use the curve in Figure 5-3, we can calculate the consonance of a pair of tones, each having six harmonic partials, as a function of the frequency separation of their fundamentals. The curve in this figure is based on such a calculation. The frequency of the fundamental of the lower tone is fixed at 250 Hz, and that for the upper tone varies from a little below 250 Hz to a little above 500 Hz (an octave above 250 Hz). The ratios above the peaks of consonance show the frequency ratios of the traditionally consonant intervals: 1/1, unison; 6/5, minor third; 5/4, major third; 4/3, fourth; 3/2, fifth; 5/3, sixth; 2/1, octave.

consists of a fundamental frequency and six harmonic partials. Here we see peaks of consonance at the familiar musical intervals. Whether or not the result represents exactly what we hear, it exhibits features characteristic of harmonic relations. The consonance of the octave is the easiest to explain.

If we designate the fundamental of the lower tone as f_0 , the partials of the lower note of the octave are $f_0, 2f_0, 3f_0, 4f_0, 5f_0$, and $6f_0$, and those of the upper note are $2f_0, 4f_0, 6f_0, 8f_0, 10f_0$, and $12f_0$; so the partials either coincide or are well separated.

Next to the octave, the fifth is the most nearly consonant interval, traditionally and as shown in Figure 5-5. Figure 5-6 indicates why this is so. In this figure horizontal distance is a measure of frequency separation in octaves; that is, frequency components with ratios of 2 to 1 are one octave apart. The six harmonics of the lower tone are shown as short vertical lines above the axis, labeled 1 through 6. The six harmonic partials of the upper tone, a fifth or .58 octave above the lower tone, are shown as short vertical

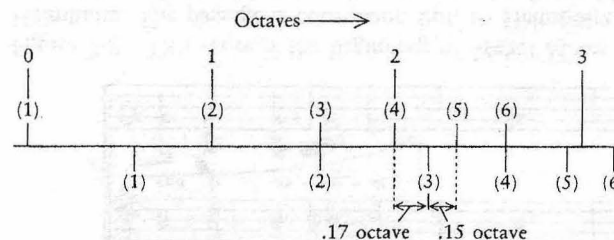


Figure 5-6 The proximity of the partials for the two tones in Figure 5-5, each with six harmonic partials, separated in frequency by the musical interval of a fifth. This interval is consonant because most of the partials of the two tones either coincide or are separated by more than a minor third. The third and sixth partials of the lower tone (shown above the horizontal line) coincide with the second and fourth partials of the upper tone (shown below the horizontal line).

lines below the horizontal line, labeled 1 through 6. Note that two partials of the upper and lower tones coincide. Most others are separated by more than a quarter of an octave. The third partial of the upper tone falls between the fifth partial of the lower tone (.15 of an octave away) and the fourth partial of the lower tone (.17 of an octave away). Thus the smallest separation of any two partials when two tones (each with six harmonic partials) a fifth apart are sounded together is about a whole tone (.17 octave). This is only a semitone less than a minor third (.25 octave).

We may note that the separation between the fifth and sixth partials of each tone is .26 octave, just greater than a critical bandwidth. If more partials are included in an individual tone, the higher partials will be less than a quarter of an octave apart, and the tone will have a sort of internal dissonance. We hear this as the buzzy quality of crude, electronically produced sounds, such as square waves and sawtooth waves. We also hear it in the jangly quality of the harpsichord. The distinction between consonant and dissonant intervals is somewhat muddled for pairs of internally dissonant tones.

We could make diagrams like that in Figure 5-6 for other intervals, such as the minor and major thirds, the fourth, and the sixth. In each case we would find some coinciding partials, and some partials closer (measured in fractions of an octave) than any in the diagram shown. But for dissonant intervals, such as a half tone, a whole tone, a tritone (an augmented fourth, or six semitones), or a seventh, many partials would lie much closer together than a quarter octave; sounded together, notes separated by these intervals sound rough and dissonant.

Perception of Consonance and Dissonance

The foregoing may seem to imply that we have found the root of consonance and harmony in the avoidance of partials that lie too close together in frequency. If partials are too close, they will beat or sound rough. For tones made up of many (say, six) harmonic partials, we can avoid serious beats or roughness by sounding together notes whose fundamental frequencies have integer ratios, such as 3/2 (fifth), 5/3 (sixth), 4/3 (fourth), 5/4 (major third), or 6/5 (minor third). Have we indeed found the unsailable basis of consonance? Perhaps.

For a pair of sine waves, naive listeners will judge *any* interval greater than a critical bandwidth as consonant; this is true even for the seventh, and even for intervals with irrational ratios of frequencies. However, trained musicians will dutifully identify thirds, fourths, fifths, and sixths, and report them as consonant, but will report sevenths as dissonant.

For complex tones made up of many partials, some musicians give strange judgments. I have heard one characterize an isolated deceptive cadence, shown at the right in Figure 5-7, as dissonant. Both final chords in the authentic cadence (left) and the deceptive cadence (right) are consonant, but the final chord in the deceptive cadence sounds unexpected and "wrong," as you can hear by playing it on the piano.

Much simple music closes with a cadence from the dominant seventh to the tonic (see Figure 5-8). The dominant seventh is an acoustically dissonant chord. Is this why the cadence is effective and identifiable? Perhaps historically. Producing sounds with computer, Max Mathews and I carefully deleted all partials closer than a quarter octave in both chords, rendering the dominant seventh chord consonant, and the tonic chord somewhat more consonant than is "natural." A trained musician identified the chords properly, even though the dominant seventh was no longer acoustically dissonant. The clue was most likely the half step from F to E, in addition to the half step from the leading tone (B) to C.



Figure 5-7 An authentic cadence (left) and a deceptive cadence (right), which ends on a chord of the sixth (submediant) instead of on a tonic chord. The final chord of the deceptive cadence is consonant, but the progression sounds odd and not final.

Figure 5-8 A cadence that goes from the dominant seventh, an acoustically dissonant chord, to the tonic, a consonant chord.

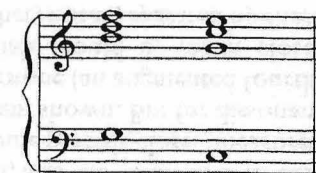


Figure 5-9 This score is the beginning of *Stabat Mater* by Palestrina, cited by Helmholtz. The passage is consonant, but, to Helmholtz, the succession of "chords" seemed odd because they did not clearly establish a key.

7930
86 THE SCIENCE OF MUSICAL SOUND

(repeated phrase)

X half step from F to E, in addition to the half step from the leading tone (B) to C.

Nonetheless, we will see in the next chapter that beats, or lack of beat, does play a crucial role in musical perception. If either the diatonic scale or the harmonic partials essential to it is slightly tampered with, overall musical or consonant effect is destroyed.

How we "hear" musical sounds depends partly on cultural bias and musical training. But it is clear there is a basis for consonance and dissonance in the very nature of human hearing, since the critical bandwidth is related to the way that the auditory nerves transmit messages to the brain. Nevertheless, harmony consists of more than an avoidance of dissonance. We can see this from Helmholtz's remarks about the beginning of Palestrina's eight-part *Stabat Mater* (Figure 5-9). Of this passage, which is certainly consonant, Helmholtz says:

Here, at the commencement of a piece, just where we should require a steady characterization of the key, we find a series of chords in the most varied keys, from A major to F major, apparently thrown together at haphazard, contrary to all rules of modulation. What person ignorant of ecclesiastical modes could guess the tonic of the piece from this commencement? As such we find D at the end of the first strophe, and the sharpening of C to C# in the first chord also points to D. The principal melody, too, which is given in the tenor, shows from the commencement that D is the tonic. But we do not get a minor chord of D till the eighth bar, whereas a modern composer would have been forced to introduce it in the first good place in the first bar.

Helmholtz then observes how different the system of Church modes was from our system of major and minor keys. Someone who is unfamiliar with sixteenth-century polyphony might try to interpret this music as D major or D minor or in some other key.

We will explore the ideas of consonance and harmony further in Chapter 6.

6 Consonance, Dissonance, and Harmony

For many years I was convinced that the work of Helmholtz, as augmented by Plomp, had established the physical and psychophysical basis of both consonance and harmony. Things seemed so simple. Sounds were consonant when no (or few) frequency components (partials) fell within the same critical bandwidth. This concept explains both the traditionally consonant intervals, such as the octave and the fifth, and the traditionally dissonant intervals, such as the second or the seventh.

In conventional counterpoint of any number of voices, a dissonance occurs only when any one voice "offends" against any other. Likewise, in chords consonant by conventional standards of harmony, all the intervals among notes meet Helmholtz's criteria to some degree. Thus I considered harmony an outgrowth of counterpoint.

A Synthetic Scale

Influenced by the ideas of Helmholtz and Plomp, in 1966 I proposed to expand the effects of musical harmony by synthesizing an entirely new scale. This scale consisted of eight equidistant notes spanning the octave, as shown in line B of Figure 6-1. The frequency ratio between any two successive notes in this scale is 1.0905; so the frequency ratio between the first and the third notes of the scale is 1.1892, which is the ratio for an equal-tempered minor third.

In effect, I had used C, Eb, Gb, and A as notes of my scale, and then had added a new note 1.5 semitones above each of these notes.

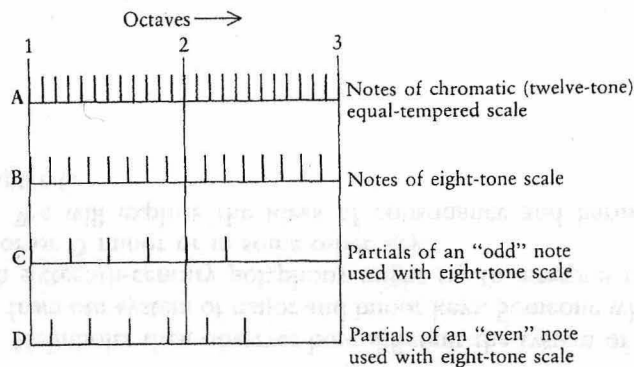


Figure 6-1 Frequencies in this illustration are measured in octaves, not in Hertz. Thus, the frequency of octave 2 is twice the frequency of octave 1, and the frequency of octave 3 is four times that of octave 1. The vertical bars along horizontal line A show the fundamental frequencies (first partials) of the notes of the equal-tempered chromatic (twelve-tone) scale. The vertical bars along line B show the frequencies of the first partials of the eight-tone scale that I used in the *Eight-Tone Canon*. The vertical bars along line C show the frequencies of the first six partials of an “odd” note, the first note in the eight-tone scale. The vertical bars along line D show the first six partials of an “even” note in the eight-tone scale. All partials of an even or of an odd pitch are one-fourth of an octave apart. Partial of odd pitch coincide; partials of even pitch coincide. But the partials of an even note and an odd note are only an eighth of an octave apart, a very dissonant interval.

Not only did I devise a new scale, but for it I synthesized new tones in which all partials except the octave partials were nonharmonic. In fact, the partials of each tone were simply the frequencies of every other higher note in the scale; so all partials were an (equal-tempered) minor third apart, as shown in lines C and D of Figure 6-1. To a musician, the tones and their partials sounded like diminished seventh chords. To me, they sounded rather fruity but not dissonant, in accordance with the theories considered in Chapter 5.

Which of these new tones are consonant when sounded together? Suppose that we distinguish the notes of my scale as *odd* and *even*; that is, the first, third, fifth, and seventh notes are odd, and the second, fourth, sixth, and eighth notes are even. The first six partials of an odd note are shown in line C of Figure 6-1; the first six partials of an even note are shown in line D. Successive odd notes are a quarter of an octave (a minor third) apart; so are successive even notes. This is the same interval as that

between the partials in the tones that I synthesized. But adjacent odd and even notes in the scale are an eighth of an octave apart. An examination of the figure above and a little reflection will show that the laws of consonance for this scale and the tones used with it are very simple. When any two odd notes are sounded together, all partials of the two tones that are sounded coincide with one another, or miss completely — the higher note has no partial that coincides with the lowest partial of the lower note. The same is true for two even notes. Hence any pair of odd notes, or any pair of even notes, will be just as consonant as the tones themselves. However, if we sound an odd note and an even note together, those two notes and some of their partials will lie an eighth of an octave apart. An eighth of an octave is less than a critical bandwidth; so odd notes sounded with even notes should sound very dissonant. Indeed, I found that they did.

Thus, for this eight-tone scale, the laws of consonance and, I thought, of harmony were extremely simple. Any and all odd notes go together; any and all even notes go together; odd and even notes sounded together result in dissonance.

To illustrate all this, I composed a four-voice *Eight-Tone Canon*, which was included in a Decca record, *The Voice of the Computer* (DL 71080), now out of print. In this canon I attempted to use transitions from dissonance to consonance in a “musically effective” way. I thought that the piece sounded pretty good, but in writing contrapuntally I had really evaded a clear test of harmony as we commonly understand it.

Stretched Octaves and Partial

In 1979 I had an opportunity to spend a month in Paris at Pierre Boulez’s IRCAM (Institute for Research and Coordination of Acoustics and Music). Emboldened by the ideas of Helmholtz and my *Eight-Tone Canon*, I resolved to attempt what seemed a crucial experiment, one that Frank H. Slaymaker had performed in part in 1970. This experiment involved the playing of conventional chords and music with “stretched” octaves, or pseudo-octaves, synthesized electronically. This concept of stretching is illustrated in Figure 6-2.

Suppose that, instead of the normal octave, with a frequency ratio of 2, we use a pseudo-octave whose ratio of frequencies is 2.4. Suppose that we also stretch, in a consistent way, the intervals between *all* frequency components of *all* partials of *all* notes that are sounded. A single tone in this system will no longer have harmonic partials (as in part A of the figure). Rather, it will have nonharmonic partials, whose frequencies are not integer multiples of the frequency of the first partial (as in part B). The successive (equal-tempered) semitones of the stretched chromatic scale will

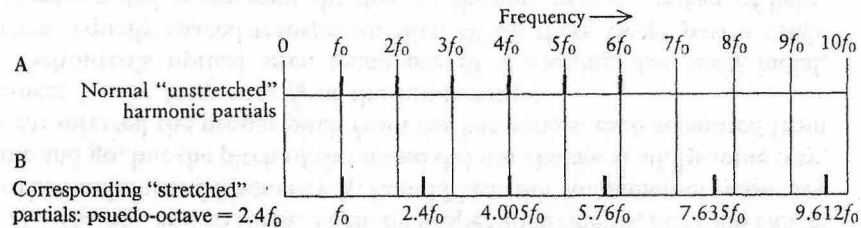


Figure 6-2 Frequency in hertz is measured from left to right. The vertical bars along the upper horizontal line (A) show the frequencies of the harmonic partials of a "normal" tone whose fundamental is f_0 . The vertical bars along the lower horizontal line (B) show the frequencies of the nonharmonic partials of a uniformly "stretched" tone in which the pseudo-octave has a frequency of $2.4 f_0$, rather than the frequency $2 f_0$ of a true octave. Note that the spacing between successive harmonic partials is always the same, but spacing between successive stretched bars increases with increasing frequency of the partials.

have frequency ratios of 1.0757, instead of 1.0594 as in the usual chromatic scale.

The result of this stretching of both the scale and the partials of the tones used with the scale is very simple. Suppose that we play the same music twice, first as *normal* (that is, unstretched and with harmonic partials) and then with stretched and with stretched partials. If any two partials of two notes coincide in frequency in the unstretched version of the music, they will coincide also in the stretched version. Partial that don't coincide in frequency in the normal music will be a little farther apart, and hence a little more consonant, in the stretched version than in the normal version. Therefore, according to Helmholtz and Plomp's theories, any combination of notes that is consonant in the normal music will be consonant, or a little more consonant, in the stretched music.

What did the stretched music sound like? We played a stretched version of *Old Hundredth* in four-part, note-against-note harmony for Pierre Boulez. He said that the only structure he could discern was the two fermatas (somewhat longer notes) at the middle and end of the piece. All familiar harmonic effects had vanished, and he apparently discerned no melody.

Those who heard the unstretched version first (we used both *Old Hundredth* and the *Coventry Carol*) could recognize the melody, and perhaps distinguish the lowest voice, whose first partial was the lowest frequency present. However, it was hard, if not impossible, to follow the inner voices.

One of the most striking harmonic effects is the "finality" or "closing" quality of the *cadence*, a progression from the dominant chord (such

as B-D-G in C major) to the tonic chord (C-E-G), in which we go from the leading tone (the seventh, or B) to the tonic, C. Such a progression sounds like "The End" to anyone familiar with Western music. In the stretched harmony, musicians could not tell a cadence from an "anticadence" that went from the tonic to the dominant! Furthermore, they could not tell an authentic cadence from a deceptive cadence that ended on the chord of the sixth (C-E-A) rather than on the tonic. Neither seemed more final; both seemed equally strange and equally consonant (or dissonant).

Some experiments carried out at about the same time at Stanford by Elizabeth Cohen as a part of her doctoral work shed light on this. Among other things, she asked people to what degree a stretched tone seemed to be fused with its partials. That is, did they hear a stretched tone as a single sound, or simply as a collection of different frequency components? She found that, when the octave was stretched more than about 5 percent, the partials were not heard as a single tone of a particular timbre. Rather, listeners heard the collection of stretched partials as distinguishable tones of different frequencies. The higher-frequency partials were not individually distinct, but stood apart from the lower frequencies as a group.

This was something of a surprise to me, for many sounds with nonharmonic partials, such as the sounds of bells, gongs, drums, or knocking on wood, are heard as distinct, identifiable timbres. In part, such natural sounds are "held together" by a common variation of different frequency components over time, such as a sharp attack and concurrent decay in amplitude, or, in the case of musical sounds, common small variations in the amplitudes or pitches of different frequency components (a common vibrato). The ear does not analyze these sounds into separate components; yet that is just what our hearing did to the computer-produced tones with considerably stretched partials. The ear simply refused to hear such a collection of independently produced, unvarying partials as a distinct tone with a single pitch and a single distinctive quality. Instead, the ear picked these stretched tones to pieces.

This, I think, accounted for the confused impression made by four-part harmony with stretched chords. Because the ear could not identify the tones as single sounds, several tones sounded together were heard, not as a chord, either consonant or dissonant, but rather as a sort of mush of sound. That is why Boulez could hear no structure in a stretched *Old Hundredth*, even though all the mathematical structure present in the normal version was present in the stretched version, and, by the rules of Helmholtz and Plomp, consonance and dissonance were preserved.

In 1987, IPO issued a wonderful disc by Houtsma, Rossing, and Wagenaars entitled *Auditory Demonstrations*. Tracks 58 through 61 illustrate the effects of a moderate stretching (octave of 2.1 rather than 2.0) of

scale frequencies and/or partial spacings. Part of a Bach chorale is played with synthesized tones. When neither scale nor partial frequencies are stretched, we hear the intended harmonic effects. When the scale is unstretched but the partial frequencies are stretched, the music sounds awful. Clearly, melodic intervals in the ratios of small whole numbers are in themselves insufficient to give Western harmonic effects. Further, when the scale is stretched but the partial frequencies are unstretched, again the music sounds awful. The harmonic effect does not lie simply in tones with harmonic partials. But when *both* the scale intervals and the partial frequencies are stretched a little, the music sounds slightly odd but good.

As we have already noted, when both scale intervals and partial spacings are stretched by the same small amount, partials associated with different pitches that fell on or near one another for unstretched intervals will fall on or near one another for stretched intervals and partials. If the stretching is small enough, something of the harmony is preserved.

This demonstration appears to show that the coincidence of near-coincidence of partials we find for normal (harmonic partials) musical sounds and for consonant intervals (with frequency ratios in the ratio of small integers) is a necessary condition for Western harmonic effects—in Bach at least. When this coincidence or near-coincidence is preserved, harmonic effect is preserved, even with small deviations from the ratios of small integers in scale intervals and in frequencies of partials.

Whatever the case may be, for Western ears there appears to be something musically magical about harmonic partials and the diatonic scale.

Virtual Pitch

When I visited the Philips laboratory in Eindhoven, Holland, shortly after World War II, Jan Schouten, an ingenious physicist, showed me a fascinating experiment. He had constructed a sort of optical siren (Figure 6-4) by means of which he could produce sounds with various waveforms. Using this, he produced sounds with harmonically related partials of frequencies $f_0, 2f_0, 3f_0, 4f_0$, and so forth. Then, by proper adjustments, he could cancel out the fundamental frequency f_0 . I could hear this fundamental frequency come and go, but the pitch of the sound did not change at all. In some way, my ear inferred the proper pitch from the harmonics, each separated from the next by the frequency f_0 of the fundamental.

Schouten's optical siren made use of a rotating disc with radial, narrow, equally spaced transparent slits. When these swept past a single stationary radial transparent slit they let through periodic pulses of light. These were translated into electric pulses by a photoelectric cell. The



Figure 6-3 Jan Schouten.

rotating transparent slits also swept past a transparent region shaped so as to produce a continuous sine wave from a second photoelectric cell as the radial transparent slits swept by. The phase of this sine wave could be shifted by rotating the sine-wave-producing region about the axis of rotation. The amplitude of the sine wave could be controlled by a gain control. Thus, the phase and amplitude of the sine wave could be adjusted so as to cancel the fundamental frequency of the sequence of pulses produced as the radial transparent slits swept past the stationary radial transparent slit.

I later found that Schouten had published an account of this work in 1938, and in a paper published in 1940 he gave the name *residue pitch* to the pitch correctly heard in the absence of the fundamental frequency. This pitch has since been called *periodicity pitch* or *virtual pitch*, a term introduced by Ernst Terhardt.

Schouten's observation should not surprise anyone who has listened to a pocket transistor radio. The speaker of such a radio is so small that the fundamental frequencies of all the lower tones are too weak to be audible; yet we hear the proper pitches, however tinny the music may sound. Perhaps the best name for what we hear is simply *musical pitch*.

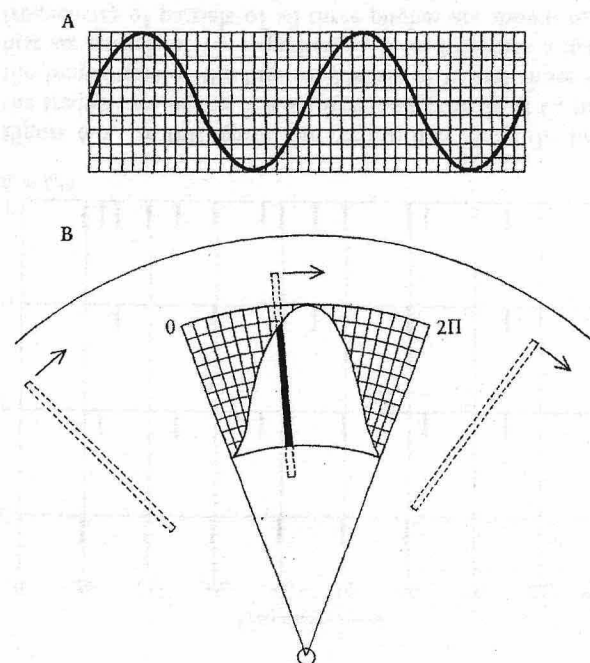


Figure 6-4 Diagram of Schouten's wave siren.

Had Helmholtz not created, or perpetuated, a myth that the pitch of a musical sound is conveyed by the presence of the fundamental frequency, Schouten's result would have seemed less surprising. For musical tones of very high pitch, such as those of the glockenspiel, the higher-frequency partials or overtones don't contribute much to loudness; the chief musical impression, timbre as well as pitch, is derived from the fundamental. For tones toward the low end of the piano keyboard, the case is quite different. The omission of the component of fundamental frequency has an inappreciable effect on the pitch or the timbre. The same is true for the second partial, the second harmonic. The third harmonic has some effect on timbre. Clearly, toward the low end of the piano keyboard musical pitch is conveyed by salient harmonics or partials, not by the fundamental.

This is consistent with the procedure of tuning pianos and organs. In the tuning process, the twelve tones of a central octave are tuned by an adjustment of beats between tones of consonant intervals. Then the upper and lower registers are tuned by octaves, that is, by beats between harmonics of the lower tone and the fundamental and harmonics of a tone an

octave higher. In the higher tone there is no frequency component for the fundamental component of the tone an octave lower to beat with.

At the high end of the range of musical pitch as defined by the piano keyboard, the lowest or fundamental frequency component is all-important to musical pitch. At the low end of the keyboard, the fundamental frequency component is unimportant to pitch, and to timbre as well. At the middle of the keyboard (around middle C) the omission of the fundamental changes the timbre (making the tone brighter) but does not change the pitch. In the mid-keyboard range, either the lowest partial, the fundamental, or the higher partials can convey pitch.

Psychoacousticians have studied the residue or virtual pitch, or musical pitch for a wide variety of combinations of sinusoids, and some of the results are curious.

When a few partials have equal spacings but are not exactly harmonics (integer multiples) of any lower frequency, the virtual pitch is some frequency of which the partials are *approximately* harmonics. Thus, the frequencies 820, 1020, and 1220 Hz can be heard as the fourth, fifth, and sixth harmonics of 204 Hz (816, 1020, and 1224 Hz); a pitch of about 204 Hz will be heard when these frequencies are sounded together.

Orderly collections of sinusoids do not always produce a virtual pitch equal to that of the missing fundamental. Sinusoids that are exclusively odd-numbered harmonics of a missing fundamental do not produce the pitch of that fundamental. But the seventh, ninth, and eleventh harmonics of 16 Hz, that is, 112, 144, and 176 Hz, are approximately the fourth, fifth, and sixth harmonics of 28.8 Hz (115.2, 144, and 172.8 Hz). They produce a pitch of about 28.8 Hz when sounded together.

Successive harmonics may or may not be heard as a virtual pitch of the missing fundamental. In generating combinations of sine waves with a Yamaha DX7, I observed that toward the low end of the piano keyboard, three sinusoids with the frequencies of C, E, and G sound like a single tone with a pitch two octaves below C, of which they are the fourth, fifth, and sixth harmonics. But toward the center of the keyboard, C, E, and G sounded together sound like a major triad, with C the lowest tone. However, still near the center of the keyboard, if we add G to C and C', the pitch drops an octave below C, a pitch of which C, G, and C' are the second, third, and fourth harmonics.

Nonetheless, the ear has a strong tendency to ascribe a single pitch to a collection of tones whose frequencies are integer multiples of a common frequency, even though that frequency itself and some of its integer multiples are absent. Furthermore, we can have a sense of pitch and unity of sound even when the frequency intervals between successive higher partials are not exactly equal. This is how we ascribe a pitch to bells. This pitch is

not the frequency of the lowest partial (the *hum tone*), but rather an average of the frequency separations between some higher partials of the sound of the bell.

Rameau and the Fundamental Bass

In his *Traité de l'harmonie* (*Treatise on Harmony*), first published in 1722, Jean-Philippe Rameau attributed the distinct character of a major triad to a fundamental bass (*basse fondamentale*). What is this fundamental bass?

A major triad is a chord made up of a note of fundamental frequency f_0 ; the major third above it, which has a fundamental frequency $(5/4)f_0$; and the fifth above it, which has a frequency $(3/2)f_0$; for example, C-E-G. As Figure 6-5 shows, the harmonics of C have frequencies

$$f_0, 2f_0, 3f_0, 4f_0, 5f_0, 6f_0$$

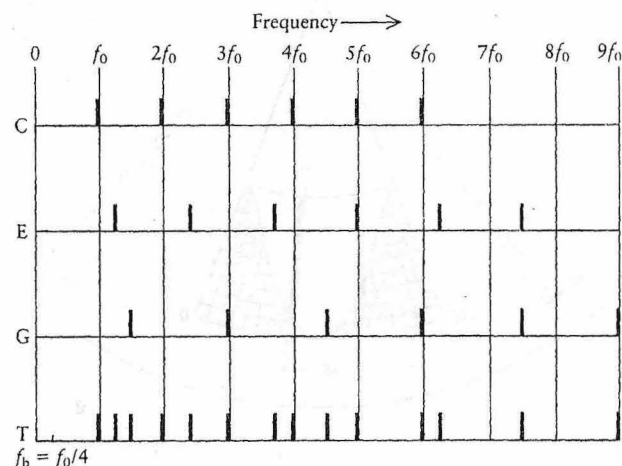


Figure 6-5 In this figure, the vertical bars along the horizontal line C indicate the frequencies of the first six harmonic partials of C; those along line E show the frequencies of the first six partials of E; and those along line G show the first six partials of G. Together, C, E, and G form a major triad. The frequencies of partials of all three pitches are shown together along line T. All the partials of all the pitches of the triad are integer multiples of a frequency $f_b = f_0/4$.

The harmonics of E have frequencies

$$(5/4)f_0, 2(5/4)f_0, 3(5/4)f_0, 4(5/4)f_0, 5(5/4)f_0, 6(5/4)f_0$$

The harmonics of G have frequencies

$$(3/2)f_0, 2(3/2)f_0, 3(3/2)f_0, 4(3/2)f_0, 5(3/2)f_0, 6(3/2)f_0$$

We see that all the partials of the notes present in the chord are integer multiples of the frequency

$$f_b = (1/4)f_0.$$

We can therefore write the harmonics of C as

$$4f_b, 8f_b, 12f_b, 16f_b, 20f_b, 24f_b,$$

the harmonics of E as

$$5f_b, 10f_b, 15f_b, 20f_b, 25f_b, 30f_b,$$

and the harmonics of G as

$$6f_b, 12f_b, 18f_b, 24f_b, 30f_b, 36f_b.$$

In the bottom row of the figure, the vertical bars indicate the frequencies of the first six partials of C, E, and G. Some of these coincide. Furthermore, many successive partials of f_b are present in this triad. The frequencies $4f_b$, $5f_b$, and $6f_b$ are successive harmonics of f_b . So are $15f_b$ (the third harmonic of E) and $16f_b$ (the fourth harmonic of C), as well as $24f_b$ (the sixth harmonic of C) and $25f_b$ (the fifth harmonic of E). In listening to this triad, should we not therefore hear a pitch two octaves below C, the root of the chord, corresponding to $f_b = f_0/4$?

Rameau could have heard this fundamental bass, two octaves below the root of the chord, when he listened to major chords. However, he regarded notes an octave apart as essentially identical. He therefore assumed as the fundamental bass the actual lowest note, or the root, of the

chord (in our case, C) when it is arranged in thirds — rather than the C two octaves below.

Rameau's next, inevitable conclusion was that what we now call various inversions of a chord are the same chord because they have the same fundamental bass. That is, E-G-C and G-C-E are the first and second inversions of C-E-G, and we regard them, as Rameau did, as essentially the same chord. Before Rameau, they were named differently and regarded as distinct chords. Rameau thus drastically reduced the number of "different" chords that a musician had to learn.

If the distinct character of major chords lies in their having a fundamental bass, so that the partials of all notes present in the chord are integer multiples of this fundamental-bass frequency, then we can easily understand the lack of harmonic effect in sufficiently stretched chords. When we stretch a chord, the partials of the stretched tones of which it is composed are not integer multiples of any one frequency. If we stretch the scale and the tones enough, we destroy the fundamental bass. The ear hears nothing to characterize the collection of stretched notes. Indeed, as we have seen, there is nothing that enables the ear to attribute a distinct pitch to the stretched tones with which we try to make up a stretched musical chord.

Although greatly stretched tones seem to be vague collections of sinusoidal components rather than fused sounds with distinct pitches, if we play a stretched scale of stretched tones, we do get a sense of rising pitch. Melody is more rugged than harmony. Melodies played in a stretched scale with stretched partials are easily recognizable. It is only harmony that is confused or lost altogether when the stretching is extreme.

As we have seen, Helmholtz correctly explained the basis of acoustical consonance. Plomp elaborated on his work, discovering that when too many partials lie within a critical bandwidth, a sound will be rough or dissonant, whatever other qualities it may have. This concept adequately explains many things, including the fact that, although melodies played on a carillon sound acceptable, harmony on a carillon sounds discordant. When "consonant" chords are played on a traditional carillon, many nonharmonic partials lie close together in frequency.

In Chapter 5 we noted that single notes as well as chords can have a dissonant character: the jangly sound of a note played on a harpsichord, or the buzzy quality of an electronically generated sawtooth wave or square wave. But there is more to harmony than this sort of consonance and dissonance.

Indeed, there is more to harmony than major triads. Rameau himself had trouble explaining minor triads, which cannot be derived from partials. Before about 1500 the final chord nearly always consisted of the root,

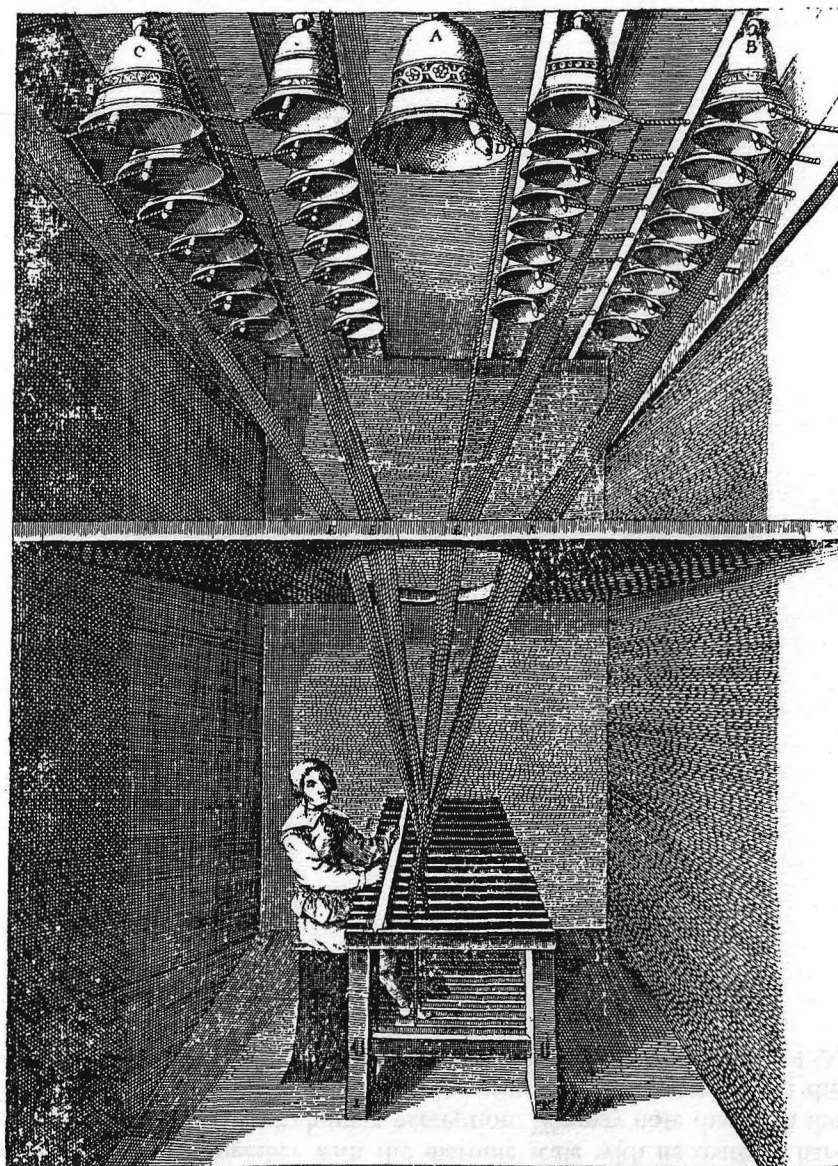


Figure 6-6 Carillon of the church of Notre-Dame at Anvers (Antwerp).

fifth, and octave—without a third, whether major or minor, as they were considered dissonant. Thirds gradually began to be regarded as consonant in the fourteenth century, but not consonant enough to appear in the final chord. Around 1500 the third began to be permitted in the final chord. The major third was preferred because it was regarded as more consonant than the minor third, and so was used in the final chord even when the rest of the piece was in the minor key. In this context, it was known as the “Picardy third.”

The major triad and its fundamental bass, which was first recognized by Rameau, is a sort of scaffolding on which Western European composers built very elaborate structures of harmony. The major triads on the tonic, the dominant, and the subdominant (C, G, and F, for example) contain all the notes of the scale, as discussed in Chapter 4. Sounded in succession, these three chords indicate the key unambiguously.

As the sole resources of harmony, these three chords would be dull. If we add another third, we can get the dominant seventh (G–B–D–F), a most useful chord. The notes of a diminished seventh (B–D–F–Ab) are the same in four different keys; so this chord, which has an ambiguous and unstable character, is useful for modulating from one of these keys to another. A few extra notes will make any major or minor chord sound lush or jazzier, or may lend it an intriguing touch of some other chord or some other key. Likewise, a chord can be rendered provocatively ambiguous by the omission of some of its notes.

We have seen that although conventional harmonic effect is preserved in tones and scales stretched a little, with greater stretching the tones themselves fall to pieces, and harmonic effect is lost. It appears that approximately harmonic partials, and scales embodying frequency ratios that approximate the ratios of small integers, are essential to harmony. Does this rule out scales other than the diatonic scale?

There has been a good deal of experimenting with scales. In his fine piece *Stria* (1977), John Chowning used partial spacings and pseudo-octaves in the ratio of the Golden Mean (approximately .618).

Later, I proposed a scale for use with tones having odd harmonic partials only. A *tritave* with a 3:1 frequency ratio (the ratio of the frequency to the lowest overtone, the third harmonic, to the frequency of the fundamental) is used instead of the octave. The analog of a major tetrad has the frequency ratios 3:5:7:9. A nine-tone scale containing six “major” triads and six “minor” triads can be chosen from a thirteen-tone “chromatic” scale with equal intervals in the ratio of $\sqrt[13]{3}$. The triads do sound chordlike. A few composers have produced attractive pieces using this scale, and singers have learned to sing and transpose in it. It appears to “work” with tones having successive harmonic partials as well as with tones having odd harmonic partials only.

Digital synthesis invites further experimentation with this and with other scales. Yet it remains possible that musical sounds with harmonic or nearly harmonic partials—that is, musical sounds that are periodic or nearly periodic—together with the diatonic scale with its diatonic harmony, are best adapted to human perception. We may note that both the spectrum of a periodic sound with successive harmonic partials and the diatonic scale involve the ratios of small integers, such as 2:1, 3:2, 4:3, 5:4, and 6:5.

7 Ears to Hear With

Imagine standing outdoors on a quiet evening with your eyes closed. What is that faint rustling sound down and to the right? Some small creature must be causing it. The susurrous sound above is the wind whispering among the dry leaves. In the distance we sense the approach of a rickety motorbike along the roughish road. It passes us with a rattle of decrepitude, departing on a slightly falling note. All of this, and more, we sense through two smooth, pencil-sized holes in our heads. And the wonderful sounds of music also reach our brains through the mechanism of the ears and the neural pathways from ear to brain.

In vision, two perspective views of the external world are imaged on the retinas. In hearing there is no such direct physical representation of external space. Yet, somehow, sound can give us a vivid impression of the world outside, of the saw biting through the wood, of the violin bow wakening the string—with a scratching noise if the player is a novice.

*Let us begin our explication of what is known about hearing with ears.

Figure 7-1 shows the outer, middle, and inner ear. The visible outer ear, on the side of the head, is called the *pinna* (and the two ears are the *pinnae*). For many years, students of hearing thought that the pinna wasn't very important. In 1967 an independent scientist named Wayne Batteau showed that it is.

If you fold your ears over, or fill the convolutions with wax or modeling clay, and with your eyes closed listen to a nearby, high-pitched sound, such as the jingling of keys on a ring, you will find that you cannot judge the *height* of the sound, that is, tell whether such a sound is in front

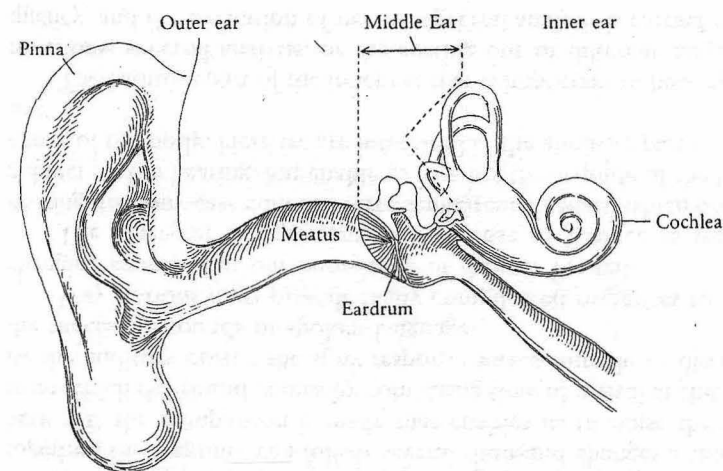


Figure 7-1 The outer ear, the middle ear, and the inner ear. The outer and middle ears are air filled; the inner ear is filled with various fluids. The loops at the upper left of the inner ear are the semicircular canals, which have to do with balance, not hearing. The cochlea is the snail-like organ at the right of the inner ear.

of you, above at an angle, or directly overhead. Because of our pinnae, the sensitivity of the ear to very high-frequency sounds changes markedly with both the direction of the sound source and the frequency of the sound. Somehow, this enables us to judge the height of the sound source.

The pinna channels sound waves to the auditory canal, or *meatus*, which is about 2.7 centimeters (1 inch) long. At the inner end is the ear drum, or *tympanic membrane*. The auditory canal acts as a rather broad-band resonator, with a resonant frequency of about 2,700 Hz. Together with characteristics of the middle and inner ear, this broad resonance helps determine the frequency at which our hearing is most acute, which is about 3,400 Hz.

The eardrum divides the outer ear (the pinna and auditory canal) from the middle ear, which consists of three tiny bones. As the eardrum vibrates in response to sound waves, the bones convey these vibrations to a third portion of the ear, the inner ear. Figure 7-2 shows the bones of the middle ear. Their names are the hammer (*malleus*, which looks more like a club), the anvil (*incus*, which looks rather like a tooth), and the stirrup (*stapes*, which really does look like a stirrup). These bones, which are flexibly connected together, convey the eardrum's vibration to a membrane cover-

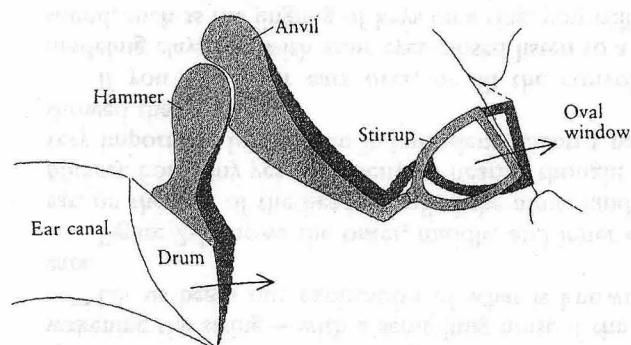


Figure 7-2 The three tiny bones of the middle ear convey vibrations from the eardrum to a membrane covering the oval window of the inner ear.

ing an opening called the *oval window*, which separates the air-filled middle ear from the fluid-filled inner ear.

In the inner ear the vibrations of sounds are sorted into overlapping ranges of frequencies. Then vibrations corresponding to different ranges of frequencies are converted into electric impulses that travel along different nerve fibers to the brain by means of a complicated, interlinked system of nerve pathways (see Figure 7-3). A good deal is known about the anatomy of these neural pathways to the brain and the “way stations” at which they are interconnected. Through animal experimentation, something is known of their function.

Because the vibrations of sound are sorted out into different, overlapping ranges of frequency, a sine wave, or a sinusoidal component of a complicated sound, will send nerve impulses toward the brain over a particular nerve pathway. Along the way, the response to one sound may interact with responses to others. But a distinct *tonotopic* mapping (*ton* for tone, *topic* for place or locality) of excitation of nerves, in response to the frequency of the sound heard, is preserved all the way up to the auditory cortex of the brain.

At each way station a band of firing neurons depicts a spectral analysis of the sound that caused the firing. If an animal hears a sequence of sine waves of increasing frequencies, the maximum responses at the way stations will shift in position in an orderly way, and across the surface of the auditory cortex also.

As we travel up the nerve pathways and to the auditory cortex, the tonotopic response will follow changes in the sounds heard more and more sluggishly, until at the cortex itself the response changes roughly at the rate

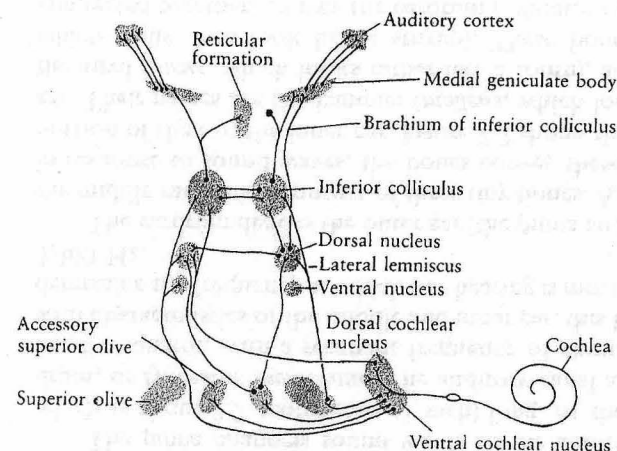


Figure 7-3 Neural pathways from the hair cells in the cochlea to the cerebral cortex, a part of the surface of the brain. Pathways from both left and right ears are interconnected at various “way stations.” A good deal is known about the pathways and their interconnections, but less about the functioning of this complex system.

of the successive sounds of speech or the successive tones of music. Presumably, we make various *uses* of the spectral features, the frequency patterns of sounds, and these uses, quick or tardy, are made at different levels in the auditory system.

At the very lowest level, where the nerves from the two ears first come together, the response can follow several thousand changes a second. It is here that the comparison is made that enables us to sense the direction from which the sound comes by comparing time of arrival at the two ears. At the auditory cortex the slow response seems suitable to disentangling the successive sounds of spoken language.

Way stations must provide many complicated responses to complex, changing sounds, but our knowledge of these is limited.

The inner ear is important to our sense of balance as well as our hearing. Its bony case contains three semicircular canals, which do not take a direct part in hearing, but enable us to sense the attitude of the head, and hence of the body. Here we are interested in the auditory part of the inner ear.

The auditory part of the inner ear that is important to hearing has two functions: spectral analysis, or the sorting out of different ranges of frequency, and the excitation of nerves. Spectral analysis is carried out in the

cochlea, a spiral, tapered tube like the inside of a snail shell. The function of the cochlea illustrates things that we have already learned in connection with resonance, filters, and spectral analysis.

The cochlea is shown on the right in Figure 7-1. The spiral tube of the cochlea makes two and a half turns and is about 3 cm (1.2 inches) long. It is about 0.9 cm (0.4 inch) in diameter at the beginning (basal) end, and about 0.3 cm (0.2 inch) in diameter at the far or apical end.

In order to be more easily understood, the cochlea is customarily drawn as a straight, tapered tube, that is, unrolled (as in Figure 7-4). The schematic cross section of the cochlea is shown at the right in the figure. W indicates the width of the basilar membrane. This springy membrane is shown arched downward, as it would be if the pressure in the fluid above was greater than the pressure in the fluid below.

The width, W , and the stiffness of the basilar membrane vary from the left (basal) end to the right (apical) end. The membrane is widest and laxest at the right end, narrowest and stiffest at the left end.

The longitudinal, or end-to-end, section of the unrolled cochlea is shown at the left in the figure. The basilar membrane does not extend all the way to the closed, apical end (on the right). Fluid above the basilar membrane can flow into the space under the membrane through an opening called the *helicotrema*.

At the basal end (at the left) are two windows in the bone that surrounds the tube of the cochlea. These are covered with thin, flexible membranes. The upper opening is called the *oval window*. The stirrup (see Figure 7-2) is connected to the membrane covering the oval window. When a sound wave causes the stirrup to move in, it pushes the membrane in and causes the fluid in the upper part of the cochlea to move to the right; when the sound wave causes the stirrup to move out, it pulls the

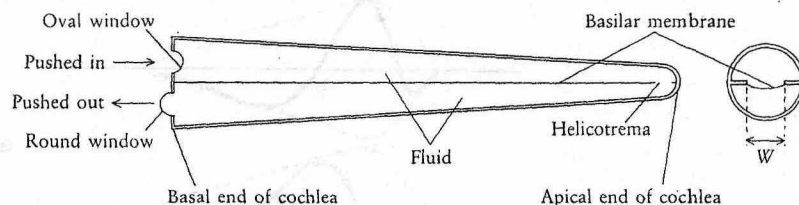


Figure 7-4 The cochlea “unrolled” and shown as a straight tapered tube. The cross section (at the right) is simplified to a circle with rigid inner projections that support the springy basilar membrane, whose width increases along the length of the cochlea, from the basal to the apical end. The longitudinal (end to end) section of the unrolled cochlea is shown at the left.

membrane out and causes the fluid in the upper part of the cochlea to move left.

The membrane that covers the round window below the basilar membrane isn’t connected to anything. This membrane flexes in and out in accord with the pressure of the fluid under the basilar membrane.

If we pushed the oval window in slowly, fluid above the basilar membrane would flow to the right, pass through the helicotrema, flow to the left under the basilar membrane, and push the membrane over the round window out. Figure 7-4 shows the membrane over the oval window pushed in (by the stirrup) and the membrane over the round window pushed out by the motion of the fluid through the helicotrema. Actually, a sound wave causes the stirrup to move rapidly right and left, thus increasing and decreasing the pressure in the fluid above the basilar membrane. This causes a wave to travel along the membrane, through the cochlea from left to right. We can best visualize this wave as a traveling, up-and-down motion of the basilar membrane.

Let us assume that the stirrup moves left and right sinusoidally with a particular frequency f in response to some sinusoidal sound or sinusoidal component of a sound. The speed with which the wave travels to the right along the basilar membrane depends on the frequency f . It also depends on the mass per unit length of the fluid above and below the membrane, on the mass per unit length of the membrane, and on the stiffness of the membrane. The cross section of the cochlea and the width, mass, and stiffness vary along the cochlea. A mathematical analysis shows how the speed of travel of a wave along the cochlea varies with distance. How the speed varies with distance depends on the frequency of the wave.

For any particular frequency, the speed of travel of the wave decreases in going from the left, basal end to the right, apical end. For each particular frequency the speed falls to zero at some *place* along the cochlea. Just to the left of that place the oscillation caused by the wave is greatest.

The motion of a wave along the cochlea from left to right is indicated in the upper part of Figure 7-5. Parts A, B, and C show the same wave as it moves to the right. The dashed lines show the envelope of the wave, that is, the greatest movement of the basilar membrane up or down at each point as the wave travels past that point.

Parts D through G of Figure 7-5 show the envelopes of four different waves along the basilar membrane for four different frequencies. We see that for low frequencies the place of greatest motion is closer to the apical end of the cochlea; for high frequencies the greatest motion is closer to the basal end. That is, the greatest vibration of the basilar membrane in response to a sinusoidal sound occurs at a particular place along it that depends on the frequency of that sound. The vibrations of the basilar

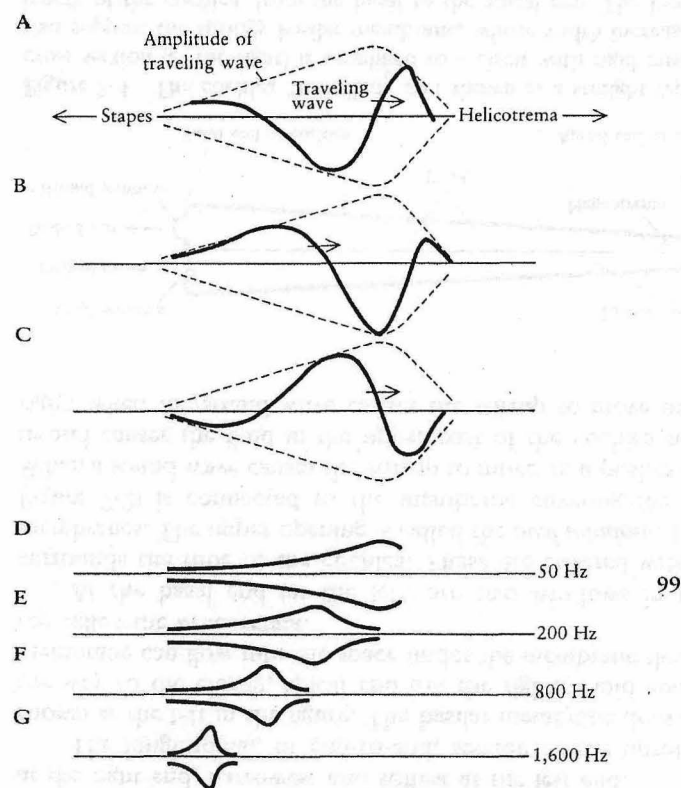


Figure 7-5 Waves traveling along the basilar membrane from left to right: A, B, and C show the same wave as it moves toward the right; D, E, F, and G show the envelopes for waves of four different frequencies. The drawings in this figure are adapted from the work of Nobel laureate Georg von Békésy, who made measurements on the cochleas of cadavers. Subsequent work on living animals has shown that the response of the basilar membrane is narrower and falls off very sharply to the right (toward the apical end). Békésy's drawings are used today chiefly because no other simple, cogent illustrations are available.

membrane excite electric pulses in nerve fibers that end on hair cells at this place on the basilar membrane (see Figure 7-6).

The reader may have seen by now that the frequency-dependent place at which the motion of the basilar membrane is greatest provides a means for sensing the frequency content of a sound wave, and hence, the pitch

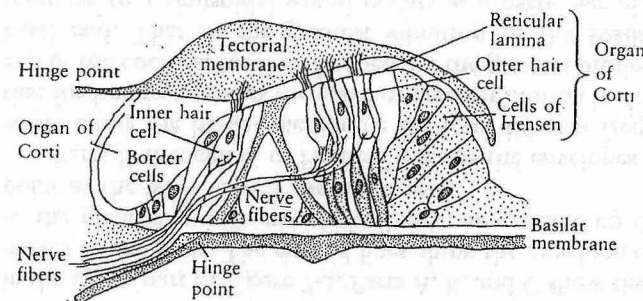


Figure 7-6 The basilar membrane and the structures surrounding it. As the basilar membrane moves, the hair cells send electric impulses along nerve fibers to the brain.

(and the timbre as well). Sine waves of different frequencies send messages to the brain along different nerve fibers; so the brain may judge the pitch of a wave by “knowing” which particular fibers carry the message to the brain. This is called the *place theory* of pitch perception, first put forth by Hermann von Helmholtz.

There is something more, however. A sharp pulse (a click) sounds just the same whether the pulse is a brief increase in pressure that pushes the eardrum in, or a brief decrease in pressure that pulls the eardrum out. When pulses or clicks follow one another in rapid succession, the ear can sense the rate at which the pulses arrive (if it is not too fast) and can infer a pitch from this rate.

As James L. Flanagan and Newman Guttman showed in 1960, sequences of pulses can be constructed for which the clue to pitch conveyed by the pulse rate conflicts with the spectral information we would get through a frequency analysis of the pulse trains. If the pulse rate predominates we infer one pitch; if the frequency analysis by the ear predominates, we infer a different pitch. If both pulse rate and frequency analysis are effective, there can be a pitch ambiguity, or there may be no clear pitch.

Consider the two pulse trains shown in Figure 7-7. The pulse rate is the same for both pulse sequence A and pulse sequence B. In A all pulses are positive, and the period (of repetition) is the time interval between pulses. In B the pulses are alternately positive and negative, and the period of repetition is twice the time interval between pulses. This means that the fundamental frequency of A, which is shown as a sine wave, is twice the fundamental frequency of B, also shown as a sine wave.

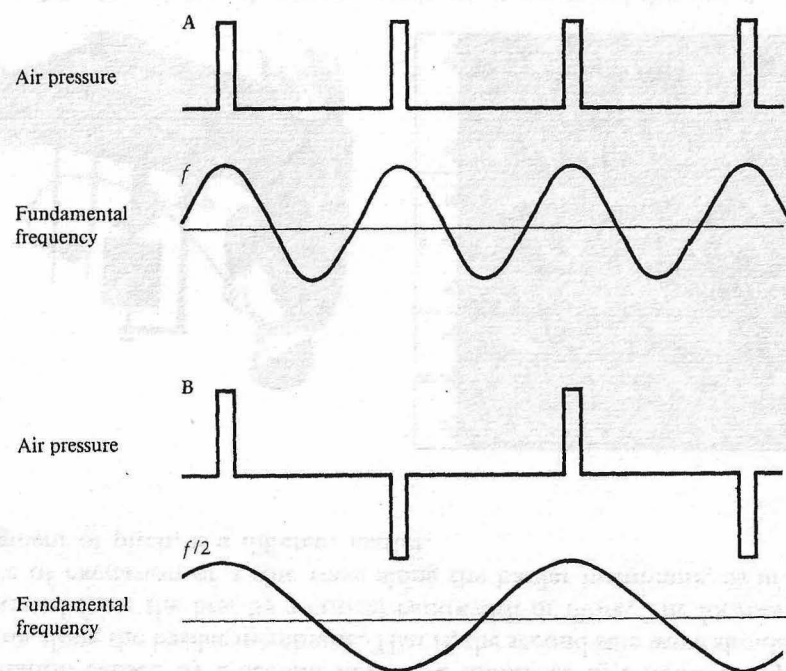


Figure 7-7 Two different pulse trains of sound. In part A, successive pulses represent sharp increases of air pressure. In part B, alternate pulses are positive and negative, that is, represent sharp increases and decreases of air pressure. At 100 pulses per second, both pulse trains sound the same pitch, but at 200 pulses per second, B sounds an octave lower than pulse train A.

What do we hear when we compare two such pulse trains? When the pulse rate is low enough, say, 100 pulses a second, pulse train B will sound like pulse train A for the same number of pulses per second, despite the fact that the fundamental frequency of B is only half that of A. But if the pulse rate of A is higher, say, 200 pulses a second, pulse train B sounds most like pulse train A when the two fundamental frequencies are the same, that is, when pulse train B has 400 pulses per second. The “match” is clearly of fundamental frequencies rather than of the number of pulses per second. At intermediate pulse rates time clues and frequency cues conflict, and pulse train B doesn’t sound like A for any pulse rate.

How is it that at 100 pulses per second the ear doesn’t make any effective frequency analysis of the pulse sequences? The ear can’t. Whatever happens in the mechanism of hearing as a result of one pulse has died

away before the next pulse comes along. At low rates the ear has no way of comparing the signs of successive pulses, and hence, no way of deducing either the fundamental frequency of the train of pulses or what harmonics of the fundamental frequency are present.

The response of the mechanism of hearing at low frequencies is more sluggish, and dies out more slowly, than the response to higher frequencies. Suppose that we filter out the lower harmonics of pulse patterns A and B, that is, those salient partials that most strongly convey the sense of pitch of musical tones. If we do this, the response to pulse rate rather than fundamental frequency can be pushed up to higher rates. Through such filtering the pulses are turned into short bursts of tone, of sinelike oscillation.

When pitch can be inferred from either repetition rate or spectrum, the clue given by spectrum tends to dominate. It seems that the first six or so partials give a very clear sense of musical pitch. When only high harmonics are present the tone tends to be buzzy and unmusical, and the sensation of musical pitch can be weak.

We have seen that there can be a conflict between time information and spectral information, with spectral information “winning” at comparatively moderate rates. But other experiments show that time information is preserved in nerve pathways at high rates. Suppose that we send a sine wave of constant frequency to one ear, and send the other ear a sine wave of the same frequency but of different phase. The sound that we hear will seem to be inside the head, but closer to the ear in which the sine wave peaks first. This effect persists for frequencies as high as 1,000 to 1,500 Hz. When a sine wave of one frequency goes to one ear and a sine wave of a slightly different frequency goes to the other ear, most of us hear *binaural beats*. These persist up to frequencies of 1,300 to 1,500 Hz. Animal experiments in which electric nerve impulses are observed by means of very fine electrodes inserted into individual nerve fibers show that time information is present for frequencies up to 4,000 to 5,000 Hz, but it is not clear that animals actually use such high-frequency time information.

It is clear that both time and place (of maximum motion of the basilar membrane) are available to us in judging periodic sounds. For nonperiodic sounds, we must rely on place information only. When we hiss by blowing air past our tongues, we can produce a sound of lower or higher pitch by moving the tongue back or forward. A hissing sound is not periodic, and any sense of pitchiness or brightness of a nonperiodic hiss must be judged by the place or places of vibration along the basilar membrane.

The place mechanism is also important in perceiving the timbre or quality of sounds. Suppose that in singing the same pitch we sing successively the vowel *u* as in *blue* and the vowel *e* as in *he*. We hear the pitch of the two vowels as identical, and this pitch agrees with the frequency of the

first partial of the note sung. However, *e* sounds shriller than *u*. And we can distinguish vowels in pitchless, whispered speech.

We distinguish between vowel sounds by means of certain *formants*, or frequency regions of high energy with particularly strong overtones. These formants are centered on the resonant frequencies of the vocal tract and are changed by changes in the shape of the tract, depending on what vowels we utter. The frequencies of the formants all lie above the first partial of the sound we utter. The formants do not change frequency as we change the pitch with which we speak or sing a vowel. (This isn't quite true for sopranos; the timbre, or quality, of the sound changes when they sing high pitches, because they are singing above the formant range for a given vowel.) We must perceive the formant frequencies, and hence distinguish among vowels, by means of the place mechanism of the ear.

It is by means of the place mechanism that we distinguish the various frequency components in a sound. The critical bandwidth discussed in Chapter 6 should correspond to the "frequency range" along the basilar membrane that a single sine wave excites strongly. If a second sine wave of different frequency is to be perceived distinctly and independently, the excitation caused by a second sine wave should lie in a nonoverlapping region along the basilar membrane. That is, the second sine wave should be separated from the first by a critical bandwidth or more. The *location* of place of excitation of a sine wave along the basilar membrane, as in the judgment of pitch, is a different matter.

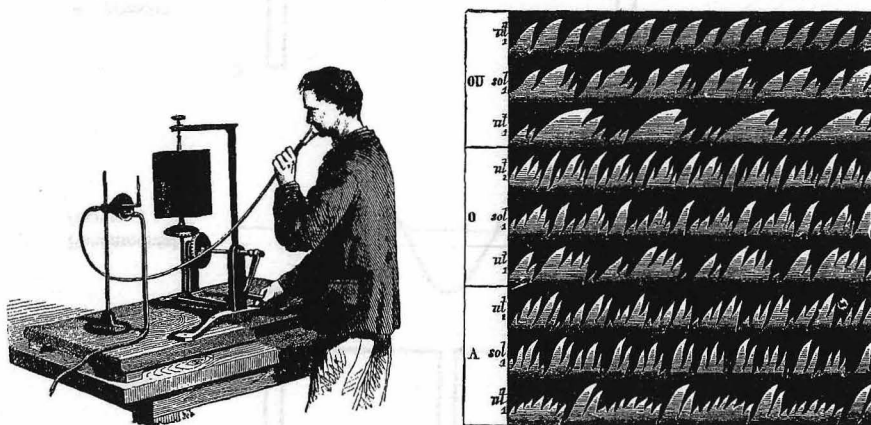


Figure 7-8 In a nineteenth-century experiment, it was found that vowel sounds would consistently impose the same pattern on the flame seen in the revolving mirrored drum regardless of the pitch level at which they were sung.

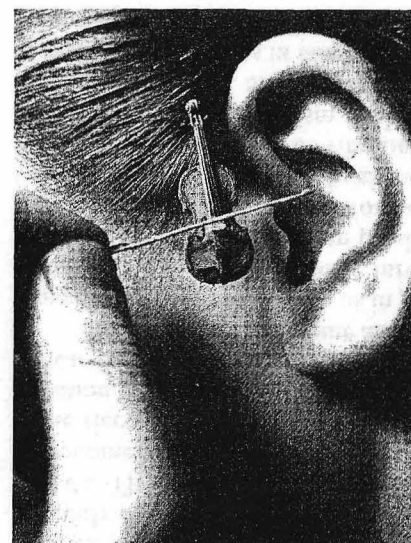


Figure 7-9 The world's smallest violin is quite shrill.

When we turn a sine wave on suddenly, at first the excitation along the basilar membrane is only roughly around the place corresponding to the wave's frequency. After some tens of cycles, a stable pattern of excitation develops. The amplitude of this stable excitation pattern falls very rapidly to zero as we move toward the apical end of the membrane. For musical tones made up of many harmonic partials, this should happen for the first six partials. The establishment of such sharp discontinuities might be the means through which the ear "measures" the places of the lower, salient harmonic partials that are essential to the sensation of musical pitch.

Georg von Békésy, who received a Nobel Prize for his work on the ear and hearing, made experiments with cochleas excised from human cadavers. He found the envelopes of the displacement of the basilar membrane in response to a sine wave to be rather broad, as shown in Figure 7-5. Later experiments he made with living animals indicated narrower envelopes, which broaden quickly after the animal dies. It appears that in the living cochlea some mechanism adds coherent energy to the oscillations set up on the basilar membrane and makes the frequency analysis sharper than it would otherwise be.

What we do know is that, in judging the periodicity and quality of sounds, the ear supplies the brain with two types of information. One type is information about time of occurrence and rate of repetition. Small

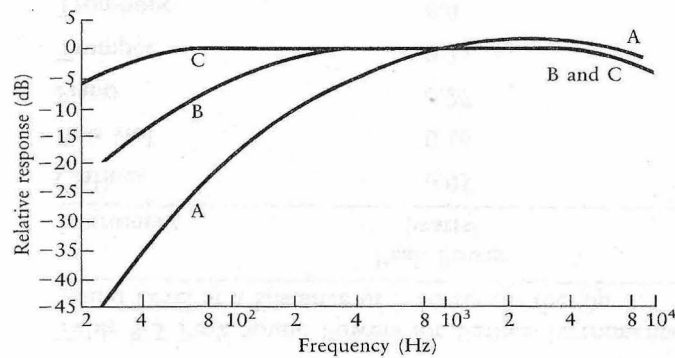


Figure 8-1 The intensity of sound or noise is measured with a sound-level meter (*bottom*). The meter does not ordinarily add the power of all frequency components of the sound and measure the sum. Rather, it “weights” the powers of different frequency components before adding. Weighting curve A (*top*) is commonly used, because the reading so obtained is closely related to loudness to the human ear for sounds of moderate intensity. Sound levels measured using the A setting (and weighting) of a sound-level meter are often quoted as dBA, as, for example, a noise level of 40 dBA. To measure the actual power of a sound, we would use the C weighting. The B weighting lies between A and C and would be appropriate for quite intense sounds.

Table 8-2 Noise Levels for Various Sources and Locations.

Source or Description of Noise	Noise Level (dBA)
Threshold of pain	130
Hammer blows on a steel plate (2 ft)	114
Riveter (35 ft)	97
Factories and shops	50–75
Busy street traffic	68
Ordinary conversation (3 ft)	65
Railroad station	55–65
Airport terminal	55–65
Stadiums	55
Large office	60–65
Factory office	60–63
Large store	50–60
Medium store	45–60
Restaurant and dining rooms	45–55
Medium office	45–55
Automobile at 50 mph	45–50
Garage	55
Small store	45–55
Hotel	42
Apartment	42
Home in large city	40
Home in the country	30
Motion picture theater, empty	25–35
Auditorium, empty	25–35
Concert hall, empty	25–35
Church, empty	30
Classroom, empty	30
Broadcast studio, no audience	20–25

For full, add
5 to 15 dB

(Continued)

Table 8-2 Noise Levels for Various Sources and Locations. (*Continued*)

Source or Description of Noise	Noise Level (dBA)
Television studio, no audience	25–35
Television, studio, audience	30–40
Sound motion-picture stage	20–35
Recording studio	20–30
Average whisper	15–20
Quiet whisper (3 ft)	10–15
Threshold of hearing	0–5

comparison, the sound level of conversational speech ranges from 40 to 70 dB (a range of 30 dB, or a thousand times).

Table 8-3 gives the peak sound powers in watts for various musical instruments. It also gives the sound level at 3 meters from each instrument in open space, calculated by assuming that the sound travels equally in all directions, with no reflections. In an enclosed room, the sound level at a distance of 3 meters would be appreciably higher because of reflections from the walls.

If the sound of a musical instrument is not reflected or interrupted, its intensity drops 6 dB (that is, to a fourth of its value) every time we double

Table 8-3 Peak Sound Powers for Various Instruments, and Their Sound Level at a Distance of 3 Meters in the Open.

Instrument	Peak Power (watts)	Decibels Above Reference
Clarinet	0.05	86
Bass viol	0.16	92
Piano	0.27	94
Trumpet	0.31	94
Trombone	6.0	107
Bass drum	25.0	113

the distance from it. If the sound level is 90 dB at 3 meters from the instrument, it will be 84 dB at 6 meters, 78 dB at 12 meters, and 72 dB at 24 meters.

Sound Intensity and Loudness Level

As long as we stick to sound power or intensity, things are straightforward. We merely have to remember that sound level expresses the measured intensity relative to a reference intensity that is near the threshold of audibility, and that the ratio of intensities is expressed in decibels rather than as “times.”

We have all observed that people differ in their hearing, as in other matters. Some have very acute hearing; some are quite deaf. The United States Public Health Service made an extensive survey of the *threshold of hearing* of many individual people; that is, of the level of sound at which they could just hear sinusoidal sounds of various frequencies. The results are shown in Figure 8-2.

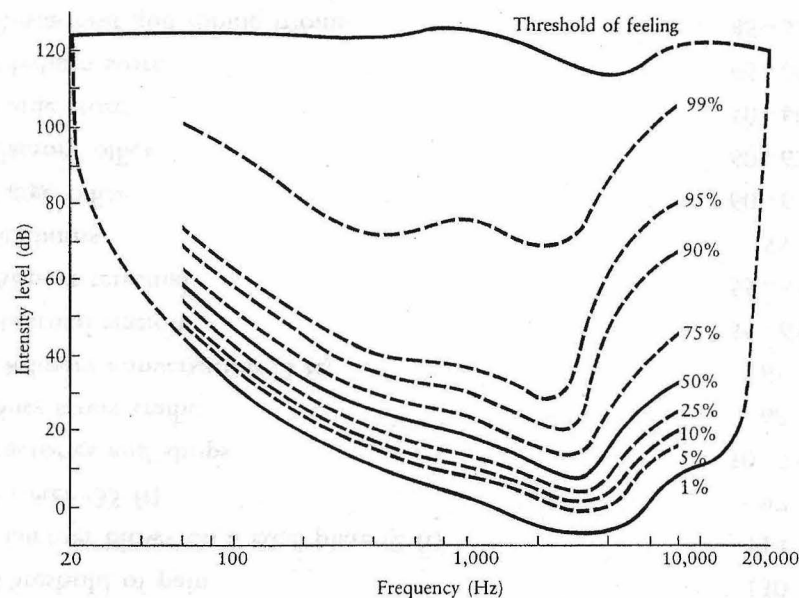


Figure 8-2 Curves showing threshold of hearing at various frequencies for a group of Americans: 1 percent of the group can hear any sound with an intensity above the 1 percent curve; 5 percent of the group can hear any sound with an intensity above the 5 percent curve; and so on.

In this figure, frequency of a steady tone is plotted from left to right, and sound intensity in decibels is plotted upward. The flat top curve shows the level of feeling, the intensity level above which sound hurts. We see that this is about 120 dB for all frequencies.

Each lower curve is designated by a percentage: 1 percent, 50 percent, ninety-nine percent. The 1 percent curve means that 1 percent of the subjects tested could hear a sound of a particular frequency whose intensity lay above the curve. For example, at 1,000 Hz, 1 percent of all individuals in the group tested could hear a pure tone whose intensity level was above 3 dB.

Clearly, how loud a sound is depends on who is listening. A sound that is inaudible to one person may be 10 or 20 dB above the lowest audible level for another.

Figure 8-2 tells us something else. Our ability to detect a sinusoidal sound, and sometimes the loudness of the sound, depends on its frequency, as well as on its intensity (power density).

All of us judge a particular sound as being louder if its intensity is increased. However, we can't discuss quantitatively the loudness of sounds for all people; so we must single out a group of persons, those with good or *acute* hearing, who are a minority among us.

For this group, loudness judged for sine waves is related to intensity and frequency by the curves shown in Figure 8-3. The curve at the bottom is for the threshold of hearing, the intensity level at which a sound of a particular frequency can just be heard. The other curves are *equal loudness curves*. That is, if two sinusoidal sounds of different frequencies have intensity levels lying on the same curve, they will sound equally loud.

We should note at once that the upper curves dip down less than the lower curves. For very loud sounds, the intensity level required to produce a given loudness doesn't change much with frequency. For very weak sounds it changes a lot. Equivalently, the loudness of a weak sound of a given intensity changes greatly as we vary its frequency; the loudness of a strong sound doesn't change much as we vary its frequency.

A consequence of this fact is that, as we turn the volume control of a stereo, the *relative* loudness of sounds of various frequencies changes. Some stereos have internal networks provided to compensate for this; mine has a switch labeled *loudness contour*. This is to be used when the sound produced by the speakers is less intense than the sound of the original music. When the loudness contour switch is on, the amplification is boosted at low and high frequencies to compensate for the dip in the low-intensity constant-loudness curves shown in the lower figure on the preceding page.

The equal-loudness curves in the figure are labeled in numbers of phons—20 phons, 40 phons, and so on. These curves can also be called

Fletcher-Munson

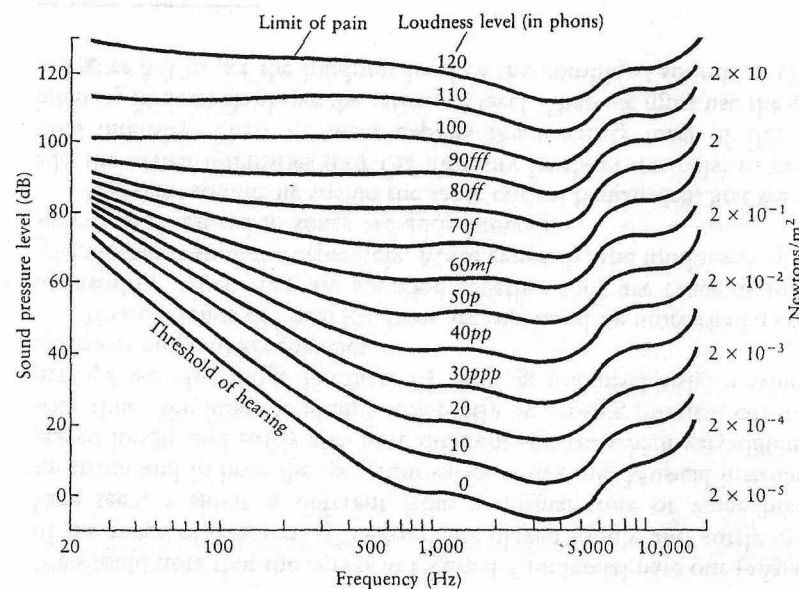


Figure 8-3 Constant-loudness curves for persons with acute hearing. All sinusoidal sounds whose levels lie on a single curve (an *isophon*) are equally loud. A particular loudness-level curve is designated as a loudness level of some number of *phons*. The number of phons is equal to the number of decibels only at the frequency 1,000 Hz.

isophons. A phon measures *loudness level*. The number of phons (of any particular equal-loudness contour) is merely the sound pressure level in decibels of an equally loud tone at 1,000 Hz. In other words, only at the frequency of 1,000 Hz does the number of phons equal the number of decibels—30 phons = 30 dB, 60 phons = 60 dB, and so on. For all other frequencies, the relationship of phons to decibels has to be arrived at by experiment.

Loudness

Loudness itself is measured in *sones*. A sound with a loudness of 20 sones sounds twice as loud as a sound with a loudness of 10 sones. A sound with a loudness of 50 sones sounds twice as loud as a sound with a loudness of 25 sones. To measure phons relative to sones, we ask a person to turn up the intensity level of a sinusoidal sound until he or she hears the sound as being twice as loud. Surprisingly, we get consistent results.

made possible more quantitative and subtler experiments. In our own time the computer and other digital technology have made easy what was difficult or impossible with earlier electronic means. They have done something more: We can now transcend all the limitations of early sound sources. We can imitate the sounds of fine instruments. We can go beyond them. Our understanding of musical effects has increased, and so has our ability to produce them.

As I noted in the first chapter of this book, an increased capability to generate, experiment with, and understand sounds, new and old, has led some enterprising composers to pay more attention to the subtle qualities of the sounds used in their compositions. This seems to me to be a healthy alternative to excessive concern with formal structure, or to a search for "spontaneity" based on some form of improvisation.

Both rules and spontaneity have their place in music, but deep understanding and careful work are essential, too. However much we may wish that we could have heard the improvisations of Bach, Mozart, or Debussy, it seems likely that their best music is the music that they left us.

The genius of the past can be a weight upon the present. Perhaps by using new resources and new understanding, music can avoid being crushed by it. Scientists are not crushed by Newton and Einstein, for they have experimental resources, knowledge, and understanding that Newton and Einstein lacked. Whatever its "absolute" value, new science is new and worthy when it succeeds in going successfully beyond the old.

May it not be so with new music? But to succeed, new music must really be heard in the sense that its composers intended, must be understood, must hold the interest of and move the listener. Here an understanding and exploration of the science of musical sounds can help. The rest only talent or genius can supply.

Appendix A Terminology

When scientists and engineers deal with well-defined, measurable physical properties, they use only unambiguous, well-defined terms, such as *time*, measured with a watch, *mass*, measured with a balance, or *length*, measured with a meterstick. This helps us grasp their intended meaning.

Many legitimate physical terms aren't nearly so simple. They can't be explained in a few words. Understanding them and using them properly results only from long exposure to experience or experiment, until the terms and their place in physics become familiar and, indeed, commonplace.

Efforts to "define" words briefly in terms of other words aren't very helpful. In both everyday life and science (if not in philosophy), we learn to use words understandably by protracted experience with things and by communicating with others. In this book I have tried to use words understandably, but it is difficult to eliminate every vestige of ambiguity.

In music we deal with many difficult qualities. They would not be made less difficult if I were to depart from the common words used by musicians, to invent a jargon, or to import one from psychology. I believe that the chief difficulties lie in the facts and experiences themselves, not in the words we use about them. I believe that the best "definitions" of the words that I use are in the text, explicitly or implicitly. Nonetheless, this brief discussion of terminology may prove useful to the reader.

Strictly, a *sound* is what we hear when a *sound wave* going through the air strikes our ears. A sound wave acts as what psychologists call a *stimulus*. Our *response* to the stimulus is the "sound" that we hear. By this definition, if no one is listening (or if only deaf persons are present), then there is no sound, but only a sound wave in the air.

The word *note* can designate *either* a mark on a musical staff *or* the sound produced, which we hear when someone "plays a note." Some try to avoid confusion by using the word *tone* for the sound produced when someone "plays a note."

A *tone* is a *musical* sound, one that can be heard as having a *pitch*. We can apply "tone" to the sound of a bell, but not to the sound of a drum. Musical sound waves are periodic fluctuations of air pressure. A *pure tone* is a sinusoidal sound wave. (One can also speak of the *good tone* of a violin, violinist, or pianist, but in this book I try not to use *tone* in that sense.)

Pitch is a quality that we hear in some sounds. Happily, for real periodic sounds the pitch that we hear is tied firmly to the periodicity, or frequency, of the sound wave. At concert pitch, A above middle C has a frequency of 440 Hz (vibrations per second). I therefore think it proper to designate pitch quantitatively by specifying frequency.

Loudness is how low a sound sounds. It is related in a complicated way to the *intensity* of a sound. Intensity is measured in watts per square meter, a good, solid physical quantity.

Timbre is a quality that a sound has in addition to pitch and loudness. Sounds that don't have a clear pitch, such as those of drums and blocks, can differ in timbre. We can use many common words to distinguish timbres: *shrill, warm, harsh, dull, percussive*. Such words describe real, consistent differences in our responses to musical sounds and in the sound waves, but it is no simple matter to pin these differences down.

Is it differences in timbre that distinguish good from bad violin playing? Physicists rightly assert that a piano emits the same sound whether an expert player strikes a key or a weight falls on it; yet some pianists obviously have a "good tone" and others don't. I don't know how to explain this difference, and you won't necessarily get help by asking a pianist who has a "good tone." She (or he) can produce the effect, but will probably not be able to tell you in words how she does it. "Make it sing," Claude Shannon's clarinet teacher told him. Shannon knew what was wanted, if not how to do it.

Appendix B Mathematical Notation

I have tried either to avoid mathematics in the text or to make it as simple as possible. Some mathematical expressions are necessary in conveying quantitative relations. For example,

$$ml$$

means m times l . Clearly, this won't work for numbers, for we would not know whether

$$27$$

was two times seven or twenty-seven. Hence, when we want to indicate the multiplication of numbers, we enclose them in parentheses. Thus

$$(2)(7)$$

is 2 multiplied by 7. We can do this, if we wish, in multiplying quantities represented by letters,

$$(m)(l) = ml$$

but nothing is gained by including the parentheses.

The expression

$$t^3$$

is the *third power* of t (or the *cube* of t). The meaning is

$$t^3 = ttt = (t)(t)(t)$$

The *exponent*, 3, tells how many t 's to multiply together. A negative exponent indicates division rather than multiplication. Thus,

$$t^{-3} = 1/t^3 = (1/t)(1/t)(1/t)$$

Let's consider a numerical example of powers. In an equal-tempered scale the frequency ratio of a semitone is

$$(2)^{1/12} = 1.059468$$

Sometimes a very large number is written in the following way:

$$\begin{aligned} 5.4 \times 10^5 &= (5.4)(10)(10)(10)(10)(10) \\ &= 540,000 \end{aligned}$$

We could write 5.4×10^5 as

$$(5.4)(10^5)$$

but for some reason we don't. A very small number can likewise be written as

$$\begin{aligned} 6.2 \times 10^{-4} &= 6.2(1/10)(1/10)(1/10)(1/10) \\ &= 0.00062 \end{aligned}$$

As a review,

$$ml/t^2$$

is the product of m and l divided by t^2 .

The square root of a quantity, say, x , is such that

$$(\sqrt{x})(\sqrt{x}) = x$$

Thus,

$$\sqrt{x} = 2$$

and

$$(\sqrt{4})(\sqrt{4}) = (2)(2) = 4$$

We encounter decibels, abbreviated, dB, in several chapters. Decibels provide a way of expressing ratios of powers. If P_1 and P_2 are two powers (commonly measured in watts), P_2 is greater than P_1 by

$$10 \log_{10} (P_2/P_1) \text{ dB}$$

The logarithm to the base 10 of a number can be looked up in a table or obtained by using a "mathematical" hand calculator. Table B-1 gives the flavor of logarithms.

Table B-1 Logarithms and decibels.

Power ratio R	Amplitude ratio (\sqrt{R})	$10 \log_{10} R$
.0001	.01	-40 dB
.001	.0316	-30 dB
.01	.1	-20 dB
.1	.316	-10 dB
1	1	0 dB
10	3.16	10 dB
100	10	20 dB
1,000	31.6	30 dB
10,000	100	40 dB
2	1.4	3 dB
1/2	.71	-3 dB

Appendix C Physical Quantities and Units

The MKS, or meter-kilogram-second, system of units is used in this book.

The units in which quantities are measured are:

mass, m	kilogram
distance, l	meters
time, t	seconds
force, F	newtons
power, P	watts
intensity, I	watts per square meter
energy, E	joules

The pull (or "acceleration") of gravity at the earth's surface, which is 9.80 meters per second per second, produces a force of 9.80 newtons on a mass of one kilogram. Force is mass times acceleration.

A force of one newton pushing something a distance of one meter requires an expenditure of one joule of energy.

The number of watts is equal to the number of joules of energy expended per second.

In MKS units, pressure is measured in newtons per square meter. One newton per square meter is called a *pascal*.

Appendix D Mathematics and Waves

Chapter 2 correctly describes the propagation of waves along a string or through the air as a traveling disturbance. Such a disturbance involves continual changes in momentum (mass times velocity) caused by a force. The force may be associated with the bending of a stretched string or the compression of air. Continual changes in the force occur because the stretched string is bent as the wave travels along, or because the air is compressed when the velocity associated with the wave is lower ahead than it is behind.

The propagation and properties of both transverse and longitudinal waves can be demonstrated by simple but rather tedious mathematical analysis (reasoning). One result of such mathematics is the discovery that waves behave in a very simple fashion only for small amplitudes; that is, when the stretched string along which a wave travels isn't bent too sharply or when a sound wave traveling through the air raises or lowers the pressure only by a small fraction. The behavior of such small-amplitude waves is *linear*. In essence, this means that when two waves are present in the same medium (string, air), they don't interact with one another. Each goes its own way as if the other weren't present. The total motion (displacement, velocity, or pressure) is simply the sum of the motions associated with the two (or more) waves.

We will not attempt a conventional mathematical analysis of waves here. Instead, we will assume that we are dealing only with small-amplitude linear waves. We will then ask, How can the velocity of the waves, the power carried by them, and other properties be expressed in terms of various physical properties? We do this by a seeming magic called *dimensional analysis*.

All physical quantities—including force, velocity, and momentum—have a dimension that is expressed in terms of the dimensions of time, mass, and length. Actual time measured in seconds is designated here by the italic letter t , whereas the dimension of time is designated by the boldface letter t , and similarly for mass, length, and any other physical quantity and its dimension. The three fundamental physical quantities and their dimensions are given in Table D-1. Let us illustrate the dimensions of some common physical quantities. What is the dimension f of frequency or periodicity in time? Frequency is the number of something per second. Number is dimensionless, and so the dimension of frequency is simply

$$t = 1/t \quad (\text{D-1})$$

What about velocity? Velocity is distance per unit time. The dimension of distance is l , and the dimension of time is t , so p , the dimension of velocity, is given by

$$p = l/t \quad (\text{D-2})$$

Table D-1 Symbols for Physical Quantities and Their Dimensions

Physical Quantity	Symbol for Physical Quantity	Symbol for Dimension of Physical Quantity
Time	t	t
Mass	m	m
Length	l	l

Acceleration is change in velocity with time. Numerically, it is the amount that velocity changes in a unit of time. Dimensionally, the dimension a of acceleration is given by

$$d = v/t = l/t^2 \quad (\text{D-3})$$

We now come to a physical law, not a matter of definition. This law was first stated by Newton. It is that, numerically, force is equal to mass times acceleration. Thus, the dimension F of force is

$$F = ml/t^2 \quad (\text{D-4})$$

Energy (or work) can be defined as force times distance. Thus, the dimension of energy, E , is

$$E = ml^2/t^2 \quad (\text{D-5})$$

Let us take a look at equation D-5. We note from equation D-2 that the dimension of velocity is l/t . Thus, the dimension of energy can also be written

$$E = mv^2 \quad (\text{D-6})$$

Dimensionally, this is correct. But the numerically correct expression for kinetic energy, or the energy of mass in motion, is

$$E = (1/2)mv^2 \quad (\text{D-7})$$

Here E is actual energy, not the dimension of energy, m is actual mass, and v is actual velocity. Dimensionally, equations D-6 and D-7 are in accord. The numerical factor $(1/2)$ is a number, and has no dimension. Finding the dimension of energy by dimensional analysis has led us to an expression for kinetic energy that is correct in everything except a multiplying numerical factor.

Let us keep what we have learned in mind, and see how we can use dimensional analysis in connection with waves.

As an example, what is the expression for the velocity v of a transverse wave traveling along a string of mass M kilograms per meter, a string that is stretched with a tension or force of T newtons?

First, what are the dimensions of M and T ? M is mass per unit length; so the dimension of M is

$$M = m/l \quad (D-8)$$

The tension T is simply a force, which has dimensions ml/t^2 , so the dimension of T is

$$T = ml/t^2 \quad (D-9)$$

The dimension of the velocity v of the wave must be l/t . We can see that this will be true if the expression for the velocity is

$$v = \sqrt{T/M} \quad (D-10)$$

We verify this by writing for the dimensions of $\sqrt{T/M}$,

$$\begin{aligned} \sqrt{T/M} &= \sqrt{(ml/t^2)/(m/l)} \\ &= \sqrt{l^2/t^2} \\ &= l/t = v \end{aligned} \quad (D-11)$$

Actually, equation D-10 is the numerically correct expression for the velocity of a wave traveling along a stretched string. The numerical factor turns out to be unity, but we couldn't know this from dimensional analysis.

Let us turn to something that is very important about plane waves traveling through air, that is, the intensity I , which is the power density measured in watts per square meter.

We noted earlier that energy or work can be defined as force times distance. Power is energy per unit time; so power will have the dimensions of force times (l/t) , or force times velocity. Intensity I is power per square meter; so intensity will have the dimensions of force times velocity divided by l^2 . Hence, the dimension I of intensity is given by

$$I = F(l/t)/l^2 = F/lt = m/t^3 \quad (D-12)$$

But in plane waves in air we are concerned not with force F , but with force per square meter, or pressure, p . The dimension p of pressure is

$$p = F/l^2 \quad (D-13)$$

From these last two relations we see that the dimension I of intensity is

$$I = p(l/t) \quad (D-14)$$

Here p is the dimension of pressure, and l/t is the dimension of velocity. We may easily conclude that if p is the fluctuating pressure associated with a sound wave and u is the fluctuating velocity associated with a sound wave, the intensity of the wave will be

$$I = pu \quad (D-15)$$

This is not only dimensionally correct, but also numerically correct, and we could have arrived at it more directly.

In a small-amplitude linear sound wave, the pressure p will be some constant, which we will call K , times the velocity fluctuation u :

$$p = Ku \quad (D-16)$$

Thus, by using equation D-16, we can express intensity I in terms of either p or u :

$$I = Ku^2 \quad (D-17)$$

$$I = (1/K)p^2 \quad (D-18)$$

K is called the *characteristic impedance* or *wave impedance* for a plane sound wave. But how are we to find an expression for K ? The dimension of K is

$$K = I/v^2 = l/(t^3) \quad (D-19)$$

Hence, from equations D-19 and D-12,

$$K = m/l^3t = (m/l^3)(l/t) \quad (D-20)$$

The first factor on the far right of equation D-19 has the dimensions of mass density, which we will call D . The second factor has the dimensions of velocity. Can it be that

$$K = Dv \quad (D-21)$$

in which D is the density of the air and v is the velocity of sound? It can be and it is, equation D-21 is numerically correct. This is very plausible, for from equations D-21 and D-17 we can write

$$I = (Du^2)v \quad (D-22)$$

Du^2 is proportional to the kinetic energy per cubic meter of the air moving at a velocity u , and in some sense this energy is transported through the air at a rate v . The kinetic energy is only half the energy transported; there is an equal amount of potential energy associated with the compression of the air by the sound wave.

We can also express the intensity of the sound wave in terms of the pressure p as

$$I = p^2/Dv \quad (D-23)$$

For air at 20°C,

$$D = 1.2174 \text{ kilograms per square meter}$$

$$v = 344 \text{ meters per second} \quad (D-24)$$

$$I = 0.002388 p^2$$

We should note that for a fluctuating pressure, the average intensity is given by the average value of p^2 divided by Dv . For a sinusoidal variation

of pressure with time, the average value of p^2 is half the square of the peak pressure (That is, the "top" of the sine wave).

In Chapter 7 the reference level of intensity is given as 10^{-12} watt per square meter. The usual reference level is a pressure of 0.00002 pascal (a pascal is a pressure of one newton per square meter). If we calculate I for this pressure using (D24), we get

$$I = 0.955 \times 10^{-12} \text{ watt per square meter}$$

This is so close to the 10^{-12} watt per square meter (only 0.2 dB different) that I chose to use the round number 10^{-12} watt per square meter in Chapter 8.

In Chapter 8 I commented on the sensitivity of the ear, and said that "theoretically" we should just be able to hear a 1-watt, 3,500-Hz sound source at a distance of 564 kilometers (352 miles). If a sound has a power of W watts and travels out equally in all directions, so that at a distance L the power passes evenly through a sphere of area $4\pi L^2$, the intensity I at a distance L must be

$$I = W/4\pi L^2 \quad (\text{D-25})$$

If we let $W = 1$ and $L = 564,000$ meters, we get from equation D-22 just about 10^{-12} watt per square meter, which is about the threshold of hearing.

Let us turn to the velocity with which a sound wave travels through air. This velocity does not vary with pressure, but it does vary with temperature. The pressure of air is caused by the velocity of air molecules. The square of this velocity is proportional to the temperature in kelvins, that is, in degrees measured with respect to absolute zero; that is, 0 kelvin is -273 degrees Celsius (centigrade).

It is dimensionally plausible that the velocity of a sound wave should be proportional to the velocity of the molecules of the air through which the wave travels, and this turns out to be so. The velocity of a sound wave at a temperature T can thus be expressed as

$$v = v_k \sqrt{T/T_k} \quad (\text{D-26})$$

in which v_k is the velocity of the sound wave at a temperature of T_k kelvins. If we take the velocity of the sound wave to be 344 meters per second at a temperature of 20 degrees Celsius (centigrade), then at a temperature T kelvins,

$$\begin{aligned} v &= 344 \sqrt{T/293} \\ v &= 20.1 \sqrt{T} \end{aligned} \quad (\text{D-27})$$

The velocity of sound varies with humidity as well as with temperature. The velocity of the molecules of a gas varies with the mass of the molecules as well as with the temperature: the lighter the molecules, the greater their velocity. Molecules of water vapor are less massive than molecules of dry air, and so the velocity of sound increases with increasing humidity.

Wind instruments are provided with tuning adjustments to compensate for the effects of temperature and humidity on pitch. The pitch of a pipe organ inexorably changes with temperature and humidity, and other instruments must accommodate this.

Appendix E Reflection of Waves

In order to understand the reflection of sound waves, we must take into account the fact that such waves consist both of an increase or decrease in the pressure of the air, designated by p , and of a forward or backward velocity of the air, designated by u . As noted in Appendix D, these two components of the wave go hand in hand. In a wave that travels from left to right,

$$p = Ku$$

in which K is a constant.

Figure E-1 shows graphically the pressure p and the velocity u of a "square" wave traveling to the right. Part A depicts a wave in which the pressure p is positive. The associated velocity u is positive; that is, it represents motion of the air to the right, in the direction in which the

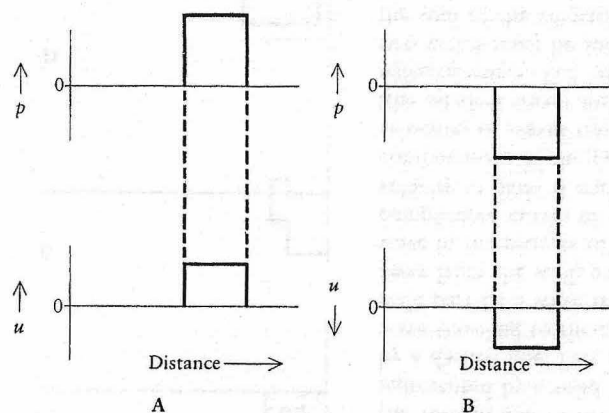


Figure E-1 For a "square" sound wave traveling to the right, the pressure p and the velocity of the air u are plotted against distance. When the pressure is greater than average, the pressure difference p is positive and lies above the axis, the horizontal line marked 0. When u lies above the axis, the velocity is to the right; when u is negative (plotted below the axis), the velocity is to the left. For a wave traveling to the right, the pressure p and the velocity u have the same sign; that is, if one lies above the axis, both do. When p is positive, u is positive, as shown in part A. When p is negative, u is negative, as shown in part B.

wave travels. In part B, the pressure p is negative, and is represented as lying below the axis, the horizontal line marked 0. This merely means that the total pressure of the air is less than the average air pressure. The velocity u is also negative (motion of the air to the left), and this is shown as a velocity lying below the axis.

Figure E-2 shows p and u for a square wave traveling to the left. Why do the graphs of waves traveling to the right look so different from those of waves traveling to the left? A wave can travel to either right or left. Shouldn't the picture be much the same in either case? The difference arises because in both cases we measure a velocity as positive (above the horizontal axis) if it is directed toward the right, regardless of whether the wave actually travels to the left or to the right. In any sound wave, if the pressure p is positive, the velocity of the air, u is in the direction in which the wave travels; if the pressure p is negative, the velocity of the air is directed opposite to the direction of travel of the wave. Thus, in the graphs of a wave traveling to the left, the pressure p is positive and the velocity u is negative, that is, to the left, the direction in which the wave travels.

Graphs such as these are very useful for understanding the reflection of waves. Figure E-3 illustrates successive stages in the reflection of a wave

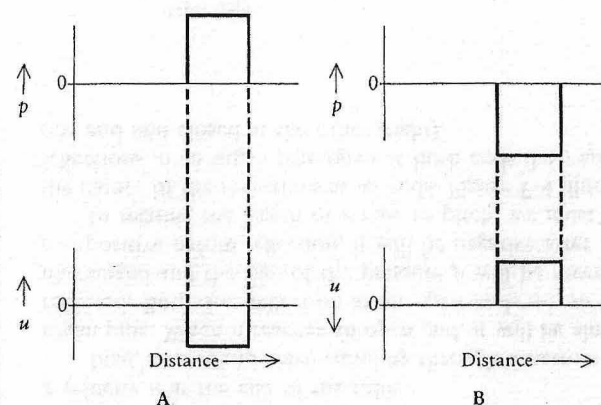


Figure E-2 For a wave traveling to the left, the pressure p and the velocity u have opposite signs. When the pressure is positive (as in part A), the velocity is negative. When the pressure is negative (as in part B), the velocity is positive. This is so because a positive velocity is generally depicted as a velocity to the right. Thus in part A the pressure is positive, and the velocity is in the same direction as that in which the wave travels, that is, to the left, shown as a negative velocity.

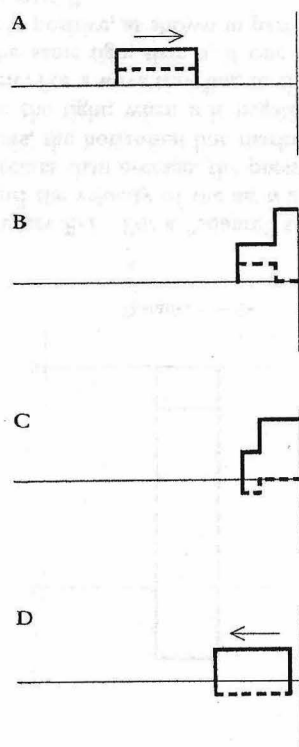


Figure E-3 The reflection of a sound wave from a solid wall, represented by the vertical line to the right. Pressure is represented by a solid line, and velocity by a dashed line. Part A represents a wave traveling to the right, toward the wall; part D, a wave traveling to the left, away from the wall; parts B and C, the wave in the process of reflection. The complicated curves of pressure and velocity in parts B and C are simply combinations of the pressures and velocities of waves traveling to the right (the *incident* wave) and to the left (the *reflected* wave). The velocities of these two waves must be such that at the wall the sum of the velocities is zero, for the air can't move at the wall. This simple condition determines the velocity, and hence the pressure, of the reflected wave.

from a wall, represented by the vertical line at the right. In this figure, pressure p is represented by a solid line and velocity u by a dashed line. In part A, we see the pressure and velocity of an *incident* wave approaching from the left. Both are positive, as is proper for a wave traveling to the right. In part D, the wave has been *reflected* from the wall. The pressure is positive and the velocity is negative, as is proper for a wave traveling to the left.

What of the pressure and velocity during the process of reflection, as shown in parts B and C? Here the pressure near the wall is twice as great as that of the incoming wave, and the velocity near the wall is zero. Can we explain this complex behavior?

The behavior is actually very simple. During the process of reflection we have overlapping waves, one traveling to the right and the other to the left. The pressure and velocities shown in parts C and D are the sums of

the pressures and velocities of these two waves: The pressures add, but the velocities, being in opposite directions and therefore having opposite signs, cancel each other out.

We see that, in reflection from a solid obstacle, the pressure of a sound wave is reversed. When a sound wave travels through a tube, we can have another sort of reflection. If the end of the tube is open, there can be no (or very little) pressure at the end of the tube, but we can have a velocity u at the end of the tube.

Imagine a sound wave traveling through a narrow tube, such as an organ pipe. When it reaches an open end, it will be almost completely reflected. But, after reflection at an open end, the air velocity u will be unchanged and the sign of the pressure p will be reversed. If the pressure p is positive before reflection, it will be negative after reflection.

In relating the length of a pipe to pitch, we must take into account the nature of the reflections at its ends. Figure E-4 illustrates successive reflections in an organ pipe open at both ends (left) and in a pipe open at one end and closed at the other (right).

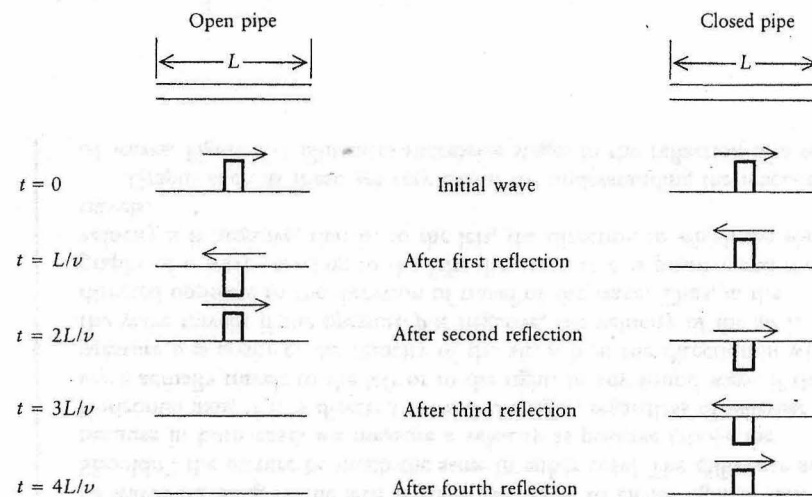


Figure E-4 The pressure of a wave after successive reflections at the ends of an open organ pipe (open at both ends, left) and a closed organ pipe (open at one end, closed at the other, right). In the open pipe, after two reflections the pressure is back where it started, and the pitch frequency is $\nu/2L$ Hz. In the closed pipe, the pressure is back where it started only after four reflections, and the pitch frequency is $\nu/4L$ Hz. For the same pitch, open organ pipes must be twice as long as a closed pipe.

In reflection from an open end, the pressure p changes sign on reflection, but, in reflection from a closed end, the sign of the pressure p remains the same after reflection. For an open pipe, we see that after a time,

$$2L/v$$

the wave has undergone two reflections and is the same as at the start. However, for a closed pipe the wave must undergo *four* reflections before it is the same as at the beginning, and this takes time.

$$4L/v$$

Thus the pitch frequency f of an open organ pipe is

$$f = v/2L \text{ Hz} \quad (\text{E-1})$$

whereas the pitch frequency f of a pipe closed at one end and open at the other is

$$f = v/4L \text{ Hz} \quad (\text{E-2})$$

Here v is the velocity of sound and L is the length of the pipe.

In stringed musical instruments, both ends of the string are rendered immovable. Hence, at the ends of the string, the transverse velocity is always zero. All reflections are the same, and the pitch frequency is always

$$f = v/2L \text{ Hz} \quad (\text{E-3})$$

But, in this case, v is the velocity with which a transverse wave travels along the string. As noted in Appendix D, this velocity is higher the greater the tension, and lower the greater the mass of the string.

Appendix F Digital Generation of Sound

Initially, the digital generation of musical sound was accomplished on general-purpose computers through special software. The generation of one second of sound commonly took several tens of seconds of computer time.

Today a host of digital synthesizers that run in real time are available. In their organization and function these draw on techniques first developed in software, such as Max Mathews's Music V, described in more detail in his 1969 book, *The Technology of Computer Music*. Music V and other software programs derived from it are still used in digital sound synthesis and in the processing of natural sounds. It seems appropriate to use Music V in explaining the capabilities and problems of digital synthesis.

How is it possible for a computer to generate sounds? The *sampling theorem* gives us a clue. Consider any waveform made up of frequency components whose frequencies are less than B . That is, consider any sound wave those frequency components lie in the bandwidth between zero and B . Any such waveform or sound wave can be represented *exactly* by the amplitudes of $2B$ samples per second. These samples are merely the amplitudes of the waveform at sampling times spaced $1/2B$ apart in time. Part A of Figure F-2 represents a waveform. In part B, the amplitudes at times $1/2B$ are shown as vertical bars drawn from the horizontal axis (which represents zero amplitude) to the curve of these bars. The successive heights can be represented by $2B$ numbers each second. These numbers describe the required samples from which the waveform can be reconstructed. Part C is an exact replica of the waveform in part A. This replica can be obtained by passing short electric pulses of the heights given in part B and described by the numbers that express these heights, through a low-pass filter of bandwidth B .

In high-quality computer-generated sound, it is customary to use 44,100 samples per second to represent a waveform. This allows us to produce frequencies up to 22,050 Hz. Because of technological limitations, the frequency range of bandwidth actually attained is 20,000 Hz.

The samples in part B of Figure F-1 are shown as lines of various heights, and the succession of sample heights stands for a succession of numbers. A computer can't produce every exact number, for most numbers can be represented only by an infinite number of digits to the right of the decimal point. A computer can produce a set of numbers such as 00 (0), 01 (1), 05 (5), 27, 44, 99—all of which are examples of the 100 possible two-digit numbers starting with 00 and ending with 99.

The internal organization of computers is such that they use only two *binary digits*, 0 and 1, and represent numbers in terms of them.

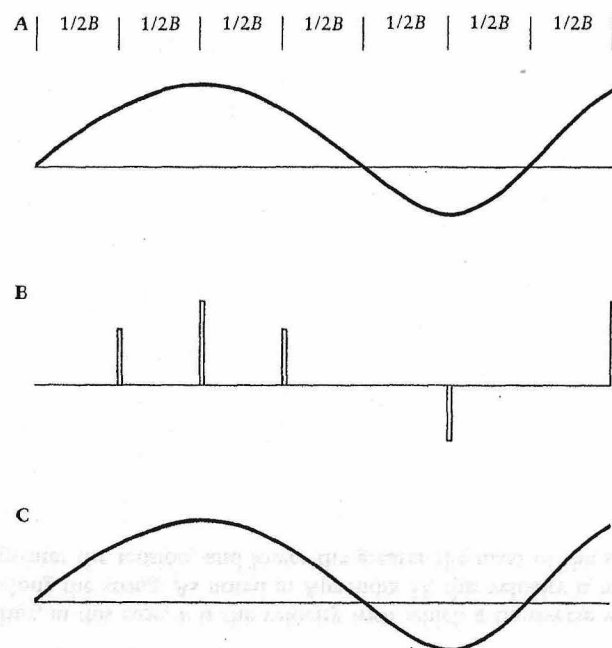


Figure F-1 A waveform of bandwidth B can be represented *exactly* by $2B$ samples a second, taken at time intervals $(1/2B)$. In part A, the samples are represented by the lengths of vertical lines drawn from the horizontal line to the curve representing the waveform. The sample amplitudes may be described by numbers. In part B, the sample amplitudes are represented by very short electric pulses whose heights correspond to the sample amplitudes. In part C, the original waveform is recovered by low-pass filtering of the sequence of pulses shown in part B.

The successive digits of common or *decimal* numbers are interpreted by means of powers of 10. Thus, 257 means

$$7 \times 10^0 + 5 \times 10^1 + 2 \times 10^2 \\ = 7 \times 1 + 5 \times 10 + 2 \times 100$$

The binary number 1001 means

$$1 \times 2^0 + 0 \times 2^1 + 0 \times 2^2 + 1 \times 2^3 \\ = 1 \times 1 + 0 \times 2 + 0 \times 4 + 1 \times 8$$

In decimal notation, the binary number 1001 is 9.

In good computer-produced sounds, 16 binary digits are used to represent the samples. This enables us to represent 65,536 different sample amplitudes. In effect, half of these binary numbers are used to represent positive sample amplitudes, and half are used to represent negative sample amplitudes.

If we use 16 binary digits to represent the largest possible sine wave, the signal-to-noise ratio of the sine wave so produced will be about 98 dB.

Suppose that we choose 44,100 sample amplitudes a second. We then produce a sequence of 44,100 short electric pulses a second, such that the amplitude of each is equal to the sample amplitude. We then pass this train of pulses through a low-pass filter to eliminate any frequency components above 22,050 Hz. Because we are free to choose the numbers that give the sample amplitudes *in any way that we wish*, we can by this means produce *any possible* sound wave whose bandwidth is 22,050 Hz or less, at least with the accuracy indicated. This is good enough for high-quality musical sounds.

This is too much freedom of choice to be of any use. In Music V, Mathews found a way to program a computer to produce a wide variety of musically useful sounds. Indeed, subsequent experience seems to indicate that Music V can be used to produce sounds as elaborate as we wish, including the sound of a singing voice and even speech sounds.

How is this done in Music V? We have all heard of the idea of a computer *simulating* some sort of mechanical or electrical device. The easiest way for noncomputer people to understand Music V is to think of it as making the computer simulate the operation of several fundamental electronic devices that are connected together in various ways. These fundamental devices are shown in Figure F-2.

One important device is the *oscillator*, represented in part A. This has an output and two inputs. The number going into input I1 specifies the amplitude of the output wave; the number going into I2 specifies the frequency of the output wave. Each oscillator is programmed to produce a particular waveform F_n , which may be a sine wave, a square wave, or some other wave.

The *adder* shown in part B is essential. The output of the adder is the sum of the two inputs, I1 and I2. We may use the adder to add the outputs of two sinusoidal oscillators to get the sum of two partials. Or we may use the adder to add a small sinusoidal vibrato to the number that specifies the average frequency of the oscillator.

The output of the *multiplier* shown in part C is the product of the two input numbers I1 and I2. The multiplier may be used in several ways: to multiply by some chosen number the amplitudes of all outputs produced by an oscillator, so that we can use one number as a volume control; or to change the frequency of an oscillator by a constant factor, thus transposing any notes played.

The final device (part D) is an output device that stores the sequence of numbers that represent the samples of the waveform produced. This storage may be accomplished in the computer memory, on disk, or on tape.

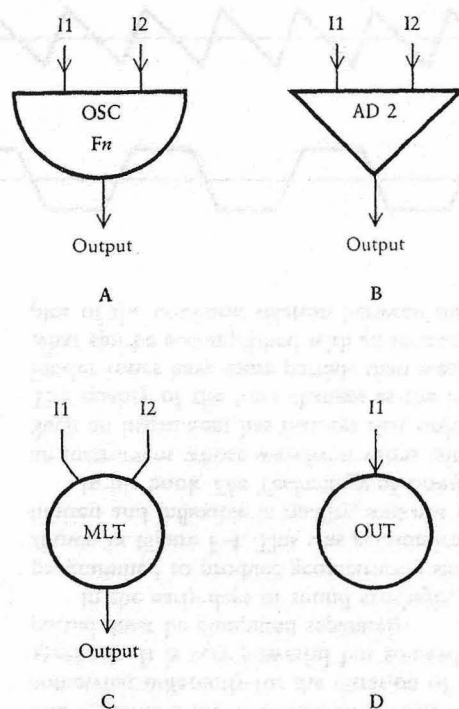


Figure F-2 The four devices that Music V “simulates.” Part A represents an oscillator F_n whose frequency is controlled by input number 12 and whose output amplitude is controlled by input number 11; part B, an adder whose output number is the sum of the two input numbers 11 and 12; part C, a multiplier whose output number is the sum of the input numbers 11 and 12; and part D, an output device to store a sequence of numbers that represent sample amplitudes.

Figure F-3 shows how to program an *instrument* and play two notes. This instrument makes use of two oscillators. Oscillator F2 produces the “squarish” waveform shown in part C. Oscillator F1 produces a single time-varying output that rises from zero and falls to zero again, as shown in part B. This output determines how the amplitude of the output of oscillator F2 rises and falls with time. Thus, the amplitude and duration of the note produced by F2 are controlled by the input numbers P5 and P6. The frequency of the note is controlled by input P7.

The program for “creating” the instrument and for playing the two notes shown in part D is given in lines 1 through 10 of the figure. In this program, lines 1 through 5 define the instrument shown in part A. Line 6

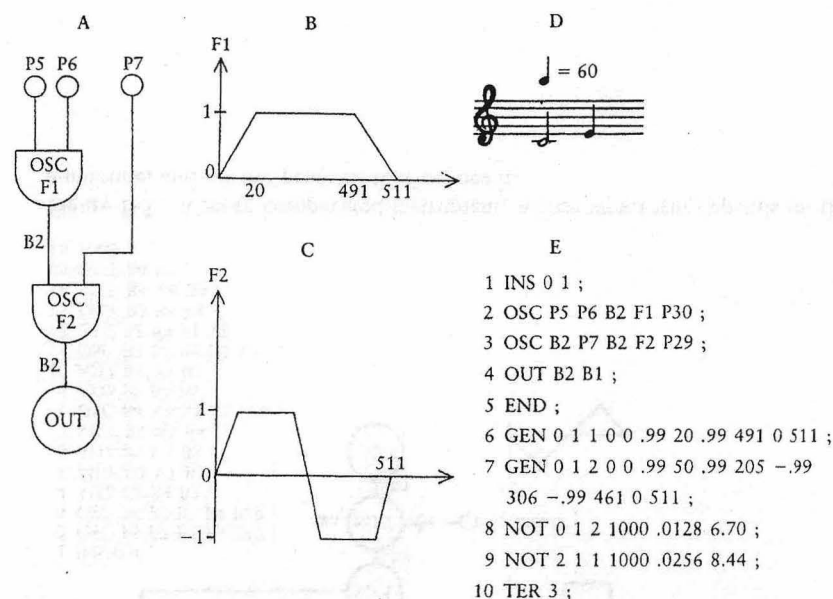


Figure F-3 A simple instrument, and the program that causes it to play the two notes shown in part D. The instrument consists of an oscillator, F1, that produces the envelope (part B) of the output of oscillator F2. The waveform of oscillator F2 is shown in part C. Lines 1 through 5 describe the instrument; lines 6 and 7 describe the envelope and waveform; and lines 8 through 10 play the notes.

defines the time function (part B) that oscillator F1 produces. Line 7 defines the waveform (part C) that oscillator F2 produces. Lines 8 and 9 cause the instrument to play the two notes shown in part D.

In line 8, the 0 following NOT says to start this note at a time zero. The 1 that follows says that the note is to be played by instrument number 1 (defined in lines 1 through 4). The 1000 that follows specifies the amplitude of the output. The following number, .0128, is, in fact, the input number P6 of part A. When this number times the number of successive samples is equal to 511, the output of oscillator F1 will have risen from zero and fallen to zero again, as shown in part B and line 6. The first note is two seconds long, and it is assumed in the example that there are 20,000 samples per second. The entry for P6 in line 8 is .0128. We note that

$$(.0128)(2)(20,000) = 512$$

which is close enough.

The last number in line 8 is the input P7, and this specifies frequency. According to part C and line 7, one complete cycle happens when P7 times the number of samples equals 511. Hence, the number of cycles per second, or Hz, is P7 times 20,000, so that the frequency is

$$(6.70)(20,000)/(511) = 262 \text{ Hz}$$

This is indeed the frequency of middle C.

The NOT lines of the figure are a very primitive way of playing a very primitive instrument. The next-to-last number, P6, depends on note duration. Surely the computer can compute this for us! Indeed, it can and does. If we wish, we can enter the frequency instead of P7, and the computer can compute the required P7. Or we can enter the name of the note and its octave number, or the octave number and a number of semitones. We can have much more complicated instruments that make it possible to have vibrato, to sweep the frequency smoothly with time, to do a host of things.

Let us just take it for granted that we can do anything we want with Music V. What do we *want* to do?

There are several methods of sound synthesis. The most direct is to add together a lot of sinusoidal partials whose amplitudes rise and fall somewhat differently for the duration of the note. This is called *additive synthesis*. It is very powerful but somewhat slow, because each sinusoidal partial must be computed separately.

In the early days of sound synthesis, the oscillators of Music V were programmed to produce geometrically simple waveforms, such as those shown in Figure F-4. This was economical, but the sounds produced were limited and inflexible in quality, and not very good.

In his book *The Technology of Computer Music*, Mathews describes an instrument whose waveform varies with amplitude (see Figure F-5). Such an instrument has features that ordinary musical instruments have: The quality of the tone changes as the intensity of the tone is increased; louder tones have more partials than weaker tones. Figure F-6 illustrates what can be accomplished with an instrument of this sort. Curve A is a plot of the nonlinear relation between the output amplitude and the input

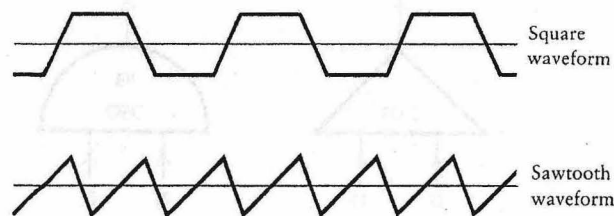


Figure F-4 Two simple waveforms that can be produced by using Music V.

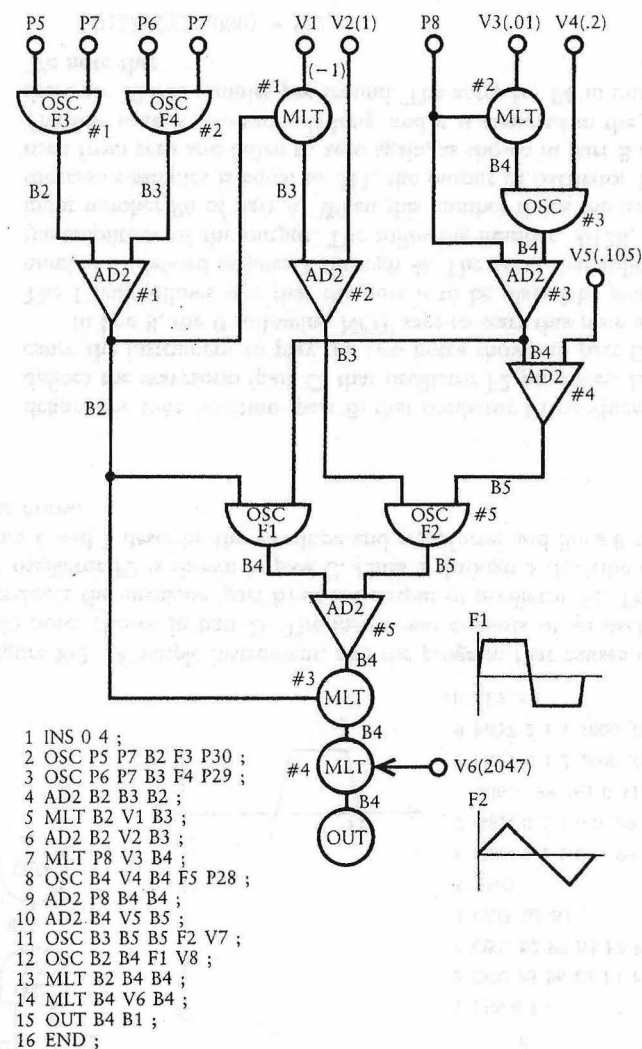


Figure F-5 A more complicated instrument, whose waveform depends on the amplitude, and on the program that defines it.

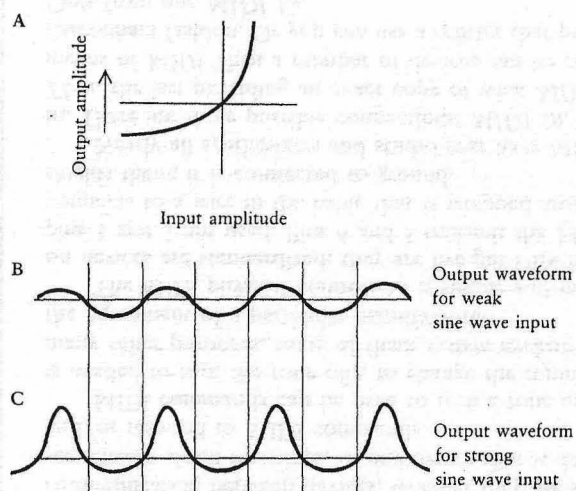


Figure F-6 A waveform that changes with amplitude can be produced by using a sine wave as an input to a device with the input-output characteristic shown in part A. If the input is a sine wave of small amplitude, the output will be almost sinusoidal, as shown in part B. If the input is a sine wave of larger amplitude, the output will be peaked waveform, as shown in part C; this has many harmonic partials.

amplitude. Waveform B shows the output waveform when the input is a weak sine wave; the output waveform is almost sinusoidal and has chiefly one partial. As shown in part C, the output for a strong sinusoidal input is peaked, and has many harmonic sinusoidal partials. The stronger the input, the stronger the higher partials. Composers Marc Lebrun and Daniel Lafitte have used this means to obtain rich tones whose timbres change with amplitude.

John Chowning invented another method, which easily produces rich, brassy sounds, called *fm* (frequency modulation) *synthesis* (see Chapter 13). It is done, in effect, by putting a sinusoidal vibrato on a sinusoidal oscillator, but the vibrato has the *same frequency as the oscillator*. In part A of Figure F-7, the amplitude of the output is plotted against a sort of pseudotime t' , which governs the rate of generation of samples. Curve B shows the effect of the vibrato on the rate of change of t' with actual time t . At points p , t' changes rapidly with t ; at points v , t' changes slowly with t . The resulting output waveform is shown in part C. The peaks rise and fall rapidly; the valleys fall and rise slowly. The resulting waveform has many high-frequency harmonics. Moreover, the intensities of these high-frequency partials can be increased by increasing the amount of frequency

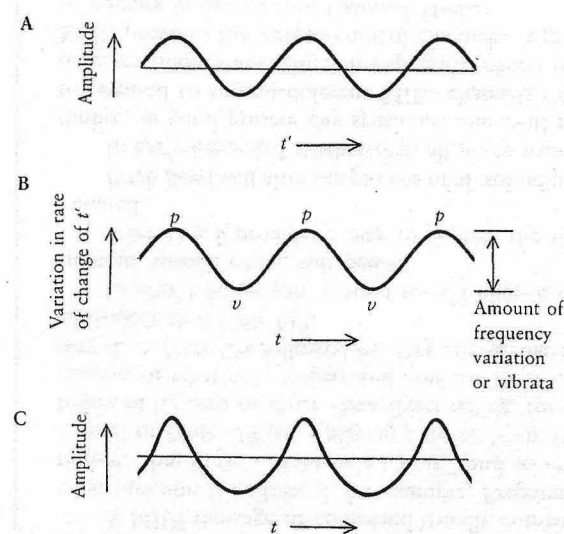


Figure F-7 Chowning's fm synthesis. In part A is a sine wave whose amplitude is a function of a variable t' . Part B shows the rate at which t' changes with time, t . This rate has a sinusoidal variation with time, a sort of vibrato. In part C, the waveform of part A is plotted, not against t' , but against actual time t . The waveform is peaked and has many harmonic partials. By gradually increasing the amount by which t' varies with time, we can gradually increase the strengths of the higher partials.

modulation (vibrato). If we gradually increase the amount of frequency modulation during the playing of a note, the intensities of the higher partials gradually increase. This is characteristic of trumpet tones. Of course, the amplitude of the sine wave that is frequency modulated in this way should rise and fall during the playing of the note.

Other interesting effects can be obtained by frequency modulating with a frequency that is an integer multiple of the oscillator frequency, or with a frequency that is not quite equal to the frequency of the oscillator.

Appendix G The MIDI Standard

In the early days of computer music, a few centers exchanged software and modified it. Music V and its derivatives offered a standard of sorts, but actual programs were different for different machines. So were formats for inputting data, and means for converting digits into sound.

These software synthesis programs enabled the user to generate new instruments with new sound qualities and new control parameters. A strong emphasis on sound itself was built into such programs. In contrast, MIDI (Music Instrument Digital Interface) is fundamentally keyboard oriented. It provides a way to play notes, and it provides ways to operate whatever controls a synthesizer may have.

As various keyboard synthesizers began to appear, users wanted them to be connected and communicate with each other. There was strong motivation for an industry-wide standard.

MIDI was the response to this need, and it is the most important standard in computer music today. It came about through initial discussions among several synthesizer manufacturers in 1981. In 1983 an official description, called the MIDI 1.0 "Detailed Specification" was released.

The design of MIDI reflects its early concept as a keyboard standard. But MIDI proved to be adaptable to many functions, to the control of a host of synthesizer functions, and to the control of synthesizers by means other than keyboard: by simulated wind and string instruments, by sequencers or computers, or by devices such as Max Mathews's Radio Baton.

In essence, MIDI is both a hardware standard and a protocol for communication between devices, whether they are synthesizers, computers, sequencers, drum machines, or any device that is designed to receive and/or respond to MIDI commands.

MIDI commands can be used to turn a tone on (a separate command is needed to turn the tone off), to change the tuning or timbre, and for many other purposes, some of them *system exclusive*, that is, pertaining to the equipment of a particular manufacturer.

The MIDI physical connection is simple and inexpensive. Connectors on devices are standardized; they are five-pin DIN female connectors, with pins 1 and 3 not used. Pins 4 and 5 transmit the MIDI signal. Pin 2 connects to a wire in the cable that is wrapped around the others and shields them; it is connected to ground.

Nearly all synthesizers and studio gear have MIDI connections built in. There are three possible connections: *MIDI In*, *MIDI Out*, and *MIDI Thru*, the last providing an exact copy of what *MIDI In* receives. By means of MIDI Thru a number of devices can be connected together in a daisy-chain fashion. Or you can use a splitter that provides several *MIDI Outs* from one *MIDI In*.

It is not so easy to control a MIDI device from more than one MIDI source—for example, to control a synthesizer from both a sequencer and a keyboard. A special "merge box" can be used that will preserve the integrity of each MIDI source. Many computer-based sequencers and many control devices have a merge function that will combine MIDI information from sequencer and keyboard so that a merge box is not needed.

Computers may or may not have MIDI connections built in, but the MIDI interfaces built for personal computers are quite inexpensive. They plug into the serial port(s) of the computer.

The MIDI signal is bit serial, a 5-milliampere current loop, opto-isolated, running at 31.25 kBaud (31.25 thousand bits a second). It is asynchronous, that is, a byte or burst of digits can be sent at any time, but always at the 31.25 kBaud rate.

Given the hardware, let us examine the protocol for MIDI commands.

Each MIDI command consists of a *Start* bit (to alert the receiver), eight data bits, and a *Stop* bit, a total of ten bits per serial byte; so the transmission time per byte is 0.00032 seconds. The device to be controlled discards the *Start* and *Stop* bits and interprets the remaining eight-bit message byte. There are two kinds of bytes:

Status Bytes—any byte starting with a 1

Data Bytes—any byte starting with a 0

This leaves seven bits per byte for *Status* or *Data*.

A MIDI message or command usually consists of a *Status Byte*, to tell what function is addressed, for example, *Program Change* (choose a timbre. "Big Brass," "Heavenly Choir," and so on), *Note On* (start playing a note) or *Note Off* (stop playing a note). Such *Status Bytes* will be followed by zero or more *Data Bytes* telling, for example, which timbre to choose, or what note to play and how loudly (with what *Key Velocity*) to play it. A *Note On* followed by data saying that the key velocity is zero is equivalent to a *Note Off*.

Control Change can be used to add nuance to a sound, as vibrato, tremolo, sustain pedal, soft pedal.

After Touch provides a way to control the sound after it has been initiated.

Pitch Bend will alter the pitches of all sounding notes on a given channel.

In early keyboard synthesizers, all notes were played with the same timbre. In good present-day synthesis, one synthesizer can be configured to respond to several different MIDI channels simultaneously, with different timbres and different expressive effects on different channels. MIDI provides for sixteen control channels. A given synthesizer can be set to operate in one of four *Channel Modes*:

MODE 1: Omni On, Poly. Receive on all channels, can play more than one note at a time. "Goof proof" for beginners.

MODE 2: Omni On, Mono. Receive on all channels, but play only one note at a time, as with the older monophonic synthesizers.

MODE 3: Omni Off, Poly. Receive only on specified MIDI channels, each of which can play a different timbre and more than one note at a time. Typical for use with sequencers. This has become the most common mode.

MODE 4: Omni Off, Mono. Receive only on specific channel. Typically, you would use this mode with a guitar controller; you would have one MIDI channel for each string of the guitar. This gives the ability to add individual expression in each channel.

When the *Mode* has been chosen, you must still play notes. *Channel Voice Messages* consist of a status byte followed by one or more data bytes. They affect only those (*Omni Off*) synthesizers that are set to receive on specified channel(s) by a *Channel Mode* message. *Channel Voice* messages are basically everything you play on the keyboard, for example, which key is pressed and which key is released, which are *Note On* and *Note Off*, respectively. Some *Channel Voice Messages* encode the note number as well as channel, and some are "channel-wide" (these do not encode note number).

System Messages (not encoded with channel number) can be received regardless of MIDI *Channel Mode* set. These commands control the whole MIDI system, for example, to synchronizing a sequencer and a drum machine, start and stop performance, and other uses.

These are some examples of *System Message* commands:

System Common—Status byte followed by zero or more data bytes. Examples are: *MIDI Song Position Pointer* (location in a sequence) and *MIDI Song Selection* (sequence number).

System Real Time—A single status byte. These commands are used to start, stop, continue, and synchronize sequencers or drum machines.

System Exclusive—A status byte, followed by *Manufacturer's ID*, followed by any number of data bytes, followed by *EOX* (End of Exclusive status) byte.

System Exclusive commands are specific to each machine. The "trapdoor" of MIDI, *System Exclusive* commands let you do "bulk dumps" of synthesizer parameters, or even of samples of digital audio. (It is not immediately obvious that MIDI can be used to transfer sounds!)

Several important additions have been made to the original MIDI standard. First, the *Sample Dump Standard* (SDS), adopted in 1986, allows a more convenient way to upload or download sound files from

computers. This is a non-real-time process, but still useful for creating, preparing, and editing sounds on a host computer to be used in a sampler.

Another supplement to MIDI is *MIDI Time Code* (MTC). This adds an *absolute* time reference to the standard. The basic idea is to encode *SMPTE* code (Society of Motion Picture and Television Engineers) synchronization standard) over MIDI. Synchronization has become extremely important in video post-production now that using a sequencer or computer to control synthesizers by MIDI has become a major activity.

It is important to realize that MIDI commands have no sense of time at all. When a MIDI command goes over a wire, it is executed as soon as it arrives at its destination. There is no intrinsic way to "save" a MIDI message, to be executed later. Saving MIDI messages is done by sequencers or computers that form a part of the system and schedule the MIDI events.

Although the abbreviation *MIDI* implies some form of digital music, it is significant that no mention is made in the specifications of what will actually *make* the sound. It could be a MIDI-enabled acoustic piano, or a MIDI-controlled carillon, calliope, drum set, merry-go-round, or rocket launcher. That is to say, the sound-producing apparatus can just as easily be mechanical as electronic. A near-ideal application of MIDI is the "MIDI piano" that has appeared recently, such as the *Disklavier* from Yamaha. In this context, it appears that it is possible both to capture and regenerate (or analyze) full concert performance via MIDI. Of course, this is really a modern version of the piano roll, but the information is easily edited, copied, and stored as MIDI data in a personal computer.

Phenomenally successful, MIDI has been built into the whole digital sound industry. Yet it has its limitations, and these have been criticized.

Researchers in the field of computer music initially tended to see MIDI as a "toy interface standard." Comparing the bandwidth of MIDI to, for example, DMA (Direct Memory Access) between a computer and dedicated peripherals is like comparing a garden hose to the Columbia River. But the years have shown that MIDI has a great deal of potential. Ultimately, the question about what we are giving up is less obvious than one might think.

Stored in the MIDI format, a fifteen-minute piece of music with sixteen voices will easily fit on an 800-K disk, tremendously reducing storage problems compared to digital audio—by around a thousand to one. Obviously, we have traded something here: We are limited to the sounds that the synthesizer or other devices we are controlling can make. Questions still remain as to whether MIDI is fast enough to capture every nuance, and whether two-way communication is possible? Though a great deal of thought was put into the standard initially, it was impossible to foresee just how far this standard would be pushed.

Appendix H MAX

MAX is a visual programming language designed for music composition and performance. The musician who learns MAX can develop control software for MIDI-compatible synthesizers and related equipment. This software can be suited to a musician's individual needs or styles of working. Originally developed at the *Institut de Recherche et de Coordination Acoustique/Musique* (IRCAM) in Paris by Miller Puckette, MAX has been in use at computer music research facilities since 1988 and became a commercial product for the Apple Macintosh computer with the additional development of David Zicarelli in 1990. It is distributed by Opcode Systems, Inc., of Menlo Park, California.

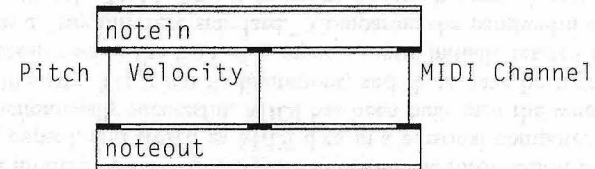
If you connected a MIDI controller such as a keyboard directly to a sound-generating device, you would normally get a single sound for each note you play. However, by inserting a computer between the controller and the sound generator, it is possible to extend your playing to produce "more" than you are actually doing. The computer effectively reinterprets the meaning of the musical gestures produced by the musician. Simple examples of this reinterpretation might include sounding a chord when a single note is played, starting or stopping a melodic sequence, or transposing a melody to the key which corresponds to the note most recently played by a performer.

Because the possibilities of extending musical performance with a computer are so vast, the best approach to developing a piece that uses such techniques is an experimental and interactive one. Unfortunately, writing a computer program to send and receive MIDI commands is not a task which can be easily accomplished by most musicians. Even if one were to write such a program, it would be necessary to use a compiled language such as C or Pascal because the computer will be responding to the performer's gestures in real time and must execute rapidly. The need to simplify the process of developing MIDI control software motivated the development of MAX as a special-purpose language. MAX simplifies the task of MIDI software development in three ways.

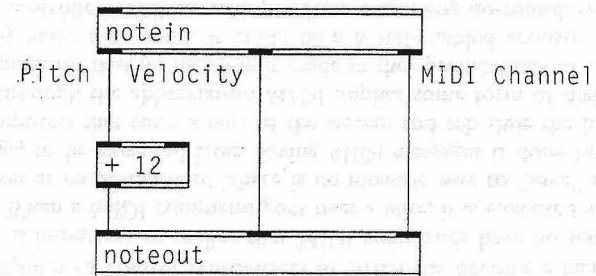
First, since it is a visual language, people with little or no programming experience can learn the basic principles without months of study. MAX programs are called *patches* because the user "patches together" boxes (called *objects*) that represent input, output, arithmetic operators, or control functions. On the computer screen, the programs look like boxes connected by wires, so you can write programs just as you might configure a stereo system or use a modular synthesizer. Conceptually, one thinks of data as "flowing through" the wires from one box to another.

Second, MAX eliminates the need to deal with MIDI codes directly. One does not need to worry about the exact format of a MIDI message or how to send it to the serial port of the computer or, a task which is

even more difficult, interpreting the type and nature of incoming MIDI messages arriving from a MIDI controller such as a keyboard. Instead, there are MAX objects for turning each part of a MIDI message into a *number* which is sent down a wire. For example, suppose our task is to write a MAX program that sends a MIDI note one octave above that which the computer receives. The MIDI note is received in an object called *notein*, which has three outputs, one for the pitch of the note, one for the velocity of the note (how hard it was played), and one for the MIDI channel (1 to 16). For the computer to merely repeat the note to the output, one just connects these outputs at the bottom of a *notein* object to the corresponding inputs at the top of a *noteout* object, as shown below. The *noteout* object takes care of putting the numbers back together into a MIDI message which is then sent out the serial port of the computer to a synthesizer.



Now, to add an octave (12 steps) to the pitch value, we merely add 12 to the pitch value, leaving the other two numbers alone.



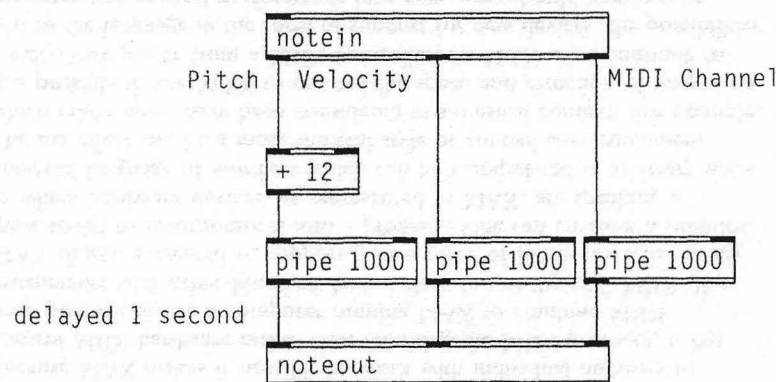
We could expand upon this example to modify the note's velocity or MIDI channel. In order to do this, we would deal with each of these numbers as an individual entity, because the MAX *notein* and *noteout* objects allow us to do so.

Finally, MAX contains a scheduler, a special facility for delaying an action until a specified time. This is especially useful in MIDI applications, because the MIDI specification says that you are responsible for turning

254 THE SCIENCE OF MUSICAL SOUND

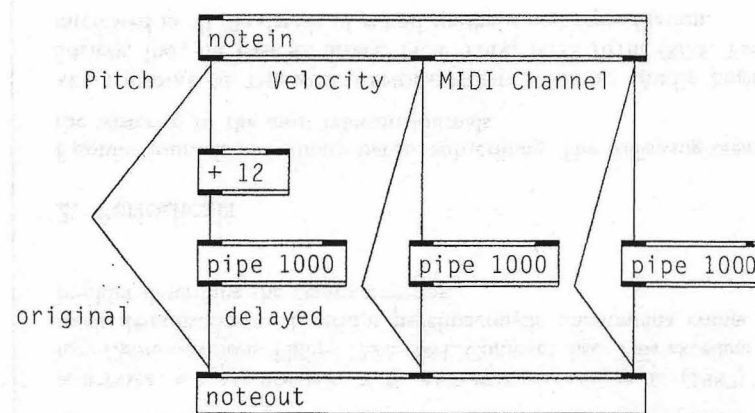
off every note you turn on. Without a scheduler, after you turn a note on, your computer program would have to tie up the computer to wait until the specified time occurs to turn it off. With a scheduler, the programmer can specify when the note is to be turned off and the computer will regularly check the clock and execute the note-turning-off task at the specified time. Meanwhile, the computer can be used for other things, including turning on and off other events. In MAX, there are a number of different objects that make use of the scheduler. One of them is called *pipe*, so named because you can send a number into the pipe and it will come out after a certain delay. (In MAX, time is expressed in milliseconds, usually accurate enough for MIDI control tasks, though not appropriate for the actual synthesis of sound.)

Returning to our simple example above, we will modify the MAX program to delay an incoming note by one second by inserting a *pipe* between the input (*notein*) and the output (*noteout*). We need a separate *pipe* for each of the three numbers.

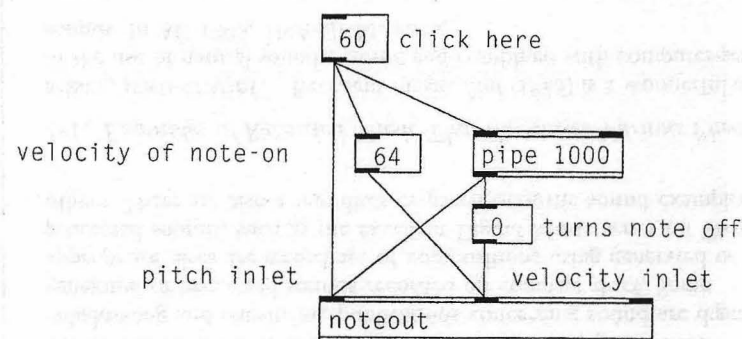


As an improvement, we can make two notes sound for each note played with a simple step. We merely reestablish the direct connection between the *notein* and *noteout* objects. This causes each number to go to two different places. The program shown at the top of page 255 will repeat the original note played on the keyboard, then, a second later, play another note an octave above the original. Note that we can make changes to the representation of the program on the computer screen without any intervening steps such as compilation. Thus, MAX lends itself to the experimentation that is so desirable when working with MIDI equipment.

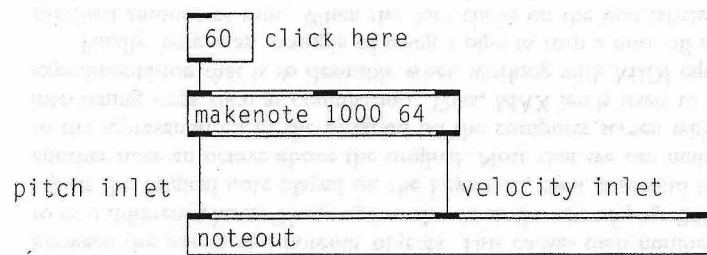
Finally, here is an example of using a *pipe* to turn a note off after a specified amount of time. When the user clicks on the box labeled 60 (middle C), the number is sent immediately to three places (the order is



determined by the right-to-left position of the destination boxes on the screen). First, the number is sent to a pipe, which will delay it for 1000 milliseconds (one second). Second, it is sent to trigger another number, 64, which is in turn sent to the velocity inlet of the *noteout* object. Finally, 60 is sent directly to the pitch inlet of *noteout*, which sends the MIDI message out the serial port. A second later, the number 60 will come shooting out of the pipe, first sending a 0 to the velocity inlet of the *noteout*. When the piped 60 arrives at the *noteout* pitch inlet, the MIDI message to turn the note off is sent out. A MIDI note message with a velocity of 0 is interpreted to mean "turn the note at the specified pitch off."



It turns out that most of the complexity of the example above has been incorporated into a single MAX object called *makenote*. At the top of page 256 is the same example rewritten to use this object, which is quite a bit simpler.



Conclusion

Because MAX makes it possible to work with individual numbers to control MIDI hardware rather than requiring the MIDI protocol, it has been possible to use a computer running MAX to combine MIDI instruments with other kinds of devices that do not "speak" MIDI. If a MAX object is created to support a new piece of hardware, it is usually quite trivial to incorporate it into a program. One can envision a situation in which hardware devices, as represented in MAX, are speaking a universal language of *numbers* which can be manipulated in arbitrary ways. The net effect can be a more musical style of control over equipment which might never have been considered in a musical context. For example, it is possible to use MAX to control the speed and direction of images on a video disk player from a MIDI controller. As MAX users continue to add to the language in the form of support for new devices, the possibilities for extending musical performance into new areas should continue to expand over the next several years.

Appendix I Bibliography

1. Compact Discs

Ultimately, sound is to hear, not to read about. Today the most enlightening and convincing publications concerning sound are digitally generated or processed sounds recorded on compact discs. Some appropriate discs are recordings of compositions using generated or processed sounds, such as the excellent Digital Music series on Wergo and others. There are also a few discs of psychoacoustic sound examples.

1-1. Examples of Recorded Music That Illustrates Various Phenomena

RISSET, JEAN-CLAUDE. Excellent music. *Sud* (1985) is a wonderful example of the use of natural sounds, edited and combined with computer-generated sounds. In AC 1003, INA-GRM, Paris.

CHOWNING, JOHN. Each of Chowning's compositions embodies something distinct and different. *Turenas* (1972) illustrates excellently the motion of sound through a room. Wergo, WER 2012-50. Schallplatten GmbH, Mainz, Germany.

1-2. Psychoacoustic Examples

MATHEWS, MAX V. (ED.) *Sound Examples, Current Directions in Computer Music*, MIT Press, 1989. Compact disc. Distributed separately from the book *Current Directions in Computer Music* (Max V. Mathews and John R. Pierce, eds.), this compact disc contains a wide range of synthesized sounds illustrating various points. John R. Pierce wishes to apologize for the fact that one of the six examples he provided (75 through 81) is erroneous (number 76).

HOOTSMA, A. J. M., ROSSING, T. D., AND WAGENAARS, W. M. (1987). *Auditory Demonstrations*. Philips 1126-061. Compact disc. This excellent eighty-track demonstration of various psychoacoustic phenomena comes with a booklet describing the demonstrations.

2. Periodicals

Examine journals in a library before subscribing. The following seem to the writer to be the most relevant journals.

AES, JOURNAL OF THE AUDIO ENGINEERING SOCIETY. Audio Engineering Society, Inc., 60 East 42 Street, New York, N.Y. 10165-0075. For those interested in all the details of sound synthesis and reproduction.

MUSIC PERCEPTION. University of California Press, Department of Psychology, University of California, San Diego, La Jolla, Calif. 92093. Covers a wide range of musical phenomena.

COMPUTER MUSIC JOURNAL. MIT Press, 55 Hayward St., Cambridge, Mass. 02142. Quite wide coverage, with a good deal of the nitty-gritty of synthesis.

JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA. Acoustical Society of America, 500 Sunnyside Boulevard, Woodbury, N.Y. 11797. This weighty journal covers all of acoustics. The papers on music and musical instruments, psychological acoustics, physiological acoustics, speech production, and speech perception will perhaps be of most interest.

3. Books

Books are cited in the hope that they can serve as a bridge to past publications and into ongoing activities. Some important papers published in various journals can be very useful. It isn't easy to select accessible and reliable books. Some get bogged down in details. Some cite out-of-date results or ideas. I advise looking at books in a library before purchasing them.

RANDEL, DON MICHAEL, ED. *The New Harvard Dictionary of Music*. Cambridge, Mass., The Belknap Press of Harvard University Press, 1986. A very useful volume.

BÉKÉSY, GEORG VON. *Experiments in Hearing*. Huntington, N.Y.: Robert E. Krieger Pub. Co. 1960, reprinted 1980. The work of a wise Nobel laureate, described by himself.

BREGMAN, ALBERT S. *Auditory Scene Analysis*. Cambridge, Mass: MIT Press, 1990. Our interpretation of sounds from the external world.

DEUTSCH, DIANA (ED.) *The Psychology of Music*. San Diego: Academic Press, 1982. Excellent. Includes chapters by bright people who are deep in computer music.

DODGE, CHARLES, AND THOMAS, JERSE. *Computer Music, Synthesis, Composition, and Performance*. New York: Schirmer Books, 1985. An excellent guide from excellent musicians.

FLETCHER, HARVEY. *Speech and Hearing in Communication*. New York: Van Nostrand, 1953. Still an interesting and useful book by a true pioneer.

GALILEO GALILEI. *Dialogues Concerning Two New Sciences*. Translated by Henry Crew and Alfonso de Salvio. New York: Dover, 1954. Pages 95–108 give Galileo's ideas concerning the relation of pitch (and harmony) to rate of vibration, including the way in which rate varies with the length, tension, and mass of a string.

HARRIS, CYRIL. *Handbook on Noise Control*. 2nd edition. New York: McGraw-Hill, 1979. By a truly reliable expert in architectural acoustics.

HELMHOLTZ, HERMANN L. F. *On the Sensations of Tone*. New York: Dover reprint, 1954. A must by the founder of modern psychoacoustics—and much more.

HUNT, FREDERICK VINTON. *Origins in Acoustics: The Science of Sound from Antiquity to the Age of Newton*. New Haven, Conn.: Yale University Press, 1978. The background that some unfortunately forget.

MATHEWS, M. V. *The Technology of Computer Music*. Cambridge, Mass.: MIT Press, 1969. Chiefly about software synthesis, but a fair amount about sound.

MATHEWS, MAX V., AND PIERCE, JOHN R. (EDS.). *Current Directions in Computer Music Research*. Cambridge, Mass.: MIT Press, 1990. Twenty-one chapters by people active in computer music research.

MOORE, F. R. *Elements of Computer Music*. Englewood Cliffs, N.J.: Prentice-Hall, 1990. The author has been deep into digital sound generation and processing from almost the beginning.

PLOMP, R. *Aspects of Tone Sensation*. New York: Academic Press, 1976. A very informative book.

ROADS, CURTIS, AND STRAWN, JOHN (EDS.). "Foundations of Computer Music." Cambridge, Mass.: MIT Press, 1985. Revised and updated articles published from 1977 to 1979 in *Computer Music Journal* on digital sound synthesis, synthesis hardware, music software, and psychoacoustics and signal processing.

ROSSING, THOMAS D. *The Science of Sound*. 2nd edition. Reading, Mass.: Addison-Wesley, 1989. A compendium of all aspects of acoustics. The information on musical instruments is particularly helpful.

SABINE, WALLACE CLEMENT. *Collected Papers*. New York: Dover, 1964. How architectural acoustics come to be.

SCHUBERT, E. D. (ED.). *Psychological Acoustics*. Stroudsburg, Pa.: Dowden, Hutchinson, and Ross, 1979. The original papers reprinted here, which date from 1876 to 1970, give a wonderful sense of how the discoveries in acoustics were really made, and the wise introductory comments to the six sections give references up to 1977.

SUNDBERG, J. *The Science of the Singing Voice*. DeKalb, Ill.: Northwestern Illinois University Press, 1987. This careful student of singing tells a lot that is worth knowing.

Sources of Illustrations

page 3

Courtesy of the Conservatoire Royal de Musique, Brussels.

page 4

Courtesy of King Musical Instruments, Inc., Eastlake, Ohio.

pages 5 and 7 (top)

Music Division, The New York Public Library at Lincoln Center; Astor, Lenox, and Tilden Foundations.

page 6

© 1963 by Hermann Moeck Verlag, Celle. Used by permission of European American Music Distributors Corporation, sole U.S. agent for Hermann Moeck Verlag.

page 7 (bottom)

Courtesy of Lejarin Hiller.

page 9

Courtesy of Moog Music, Buffalo, New York.

page 10

Courtesy of James A. Moorier.

page 15

From *The Science of Sound*, by John Tyndall, Philosophical Library, New York, 1964.

pages 18-19

Adapted from a drawing in *Musical Acoustics: An Introduction*, by Donald E. Hall, Wadsworth, 1980.

page 20

Courtesy of the Metropolitan Museum of Art, Fletcher Fund, 1956.

page 23

From *Harmonie universelle*, by Marin Mersenne/courtesy of Martinus Nijhoff Publishers.

page 24

From *Anecdotal History of the Science of Sound*, by Dayton C. Miller. Copyright 1935 by Macmillan Publishers Co., Inc., renewed in 1963 by The Cleveland Trust Company, Executor.

page 25

Photography by Peter Simon/Black Star.

page 31

Photography by Saxon Donnelly/courtesy of the University of California, Berkeley.

page 33

Photography by Philip L. Molten.

page 43

Courtesy of Elizabeth A. Cohen, Andrew Schloss, and Eric Schoen.

page 45
Courtesy of the American Institute of Physics, Niels Bohr Library.

page 46
From *On the Sensations of Tone as a Physiological Basis for the Theory of Music*, by Hermann von Helmholtz.

page 48 (top)
Photography by Constantine Manos/Magnum.

page 48 (bottom)
Photography by William P. Gottlieb.

page 49
Photography by Bob Shamis.

page 50 (bottom)
Adapted from a graph in "The Acoustics of Violin Plates," by Carleen Maley Hutchins. Copyright © 1981 by Scientific American, Inc. All rights reserved.

page 54
By A. M. S. Quinn/courtesy of Max V. Mathews.

pages 57 and 58
Courtesy of Elizabeth A. Cohen.

pages 65 and 66
Courtesy of Elizabeth A. Cohen.

page 77
Copyright © 1947 by Associated Music Publishers, Inc. All rights reserved. Used by permission.

pages 80 and 82
Adapted from graphs in *Experiments in Tone Perception*, by R. Plomp, Institute for Perception RVO-TNO, National Defense Research Organization TNO, Soesterberg, The Netherlands, 1966.

page 81
Adapted from a graph in "Critical Band Width in Loudness Summation," by E. Zwicker, E. G. Flottorp, and S. S. Stevens, *The Journal of the Acoustical Society of America* 29(1957):548.

pages 93 and 94
Courtesy of the Institut voor Perceptie Onderzoek, Eindhoven, The Netherlands.

page 99
From *Harmonie universelle*, by Marin

Mersenne/courtesy of Martinus Nijhoff Publishers.

pages 103 and 104
Adapted from drawings in *Waves and the Ear*, by Willem van Berjeick, John R. Pierce, and Edward E. David, Jr., copyright © 1960 by Anchor Books. Reprinted by permission of Doubleday & Company, Inc.

page 105
Adapted from a drawing in "Neuroanatomy of the Auditory System," by R. R. Gacek in *Foundations of Modern Auditory Theory*, volume 2, J. V. Tobias, editor, Academic Press, 1972.

pages 108 and 109
Adapted from drawings in *Waves and the Ear*, by Willem van Berjeick, John R. Pierce, and Edward E. David, Jr., copyright © 1960 by Anchor Books. Reprinted by permission of Doubleday & Company, Inc.

page 112
From *The Wonders of Acoustics*, by Rodolphe Radau, Scribner's 1886.

page 113
Photography by Nobu Arakawa/Image Bank.

page 118 (top)
Adapted from a graph in *Modern Sound Reproduction*, by Harry F. Olson, Robert E. Krieger, 1978.

pages 121 and 123
Adapted from drawings in *Speech and Hearing in Communication*, by Harvey Fletcher, D. van Nostrand, © 1953 D. van Nostrand.

page 124
Reproduced from *The Acoustical Foundations of Music*, by John Backus, by permission of W. W. Norton & Company, Inc. Copyright © 1969 by W. W. Norton & Company, Inc., and John Murray, Ltd., London.

page 126
Adapted from a graph in "Musical Dynamics," by Blake R. Patterson. Copyright © 1974 by Scientific American, Inc. All rights reserved.

page 132
Adapted from graphs in *Speech and Hearing in Communication*, by Harvey Fletcher, D. van Nostrand, © 1953 D. van Nostrand.

page 135
Adapted from graphs in "On the Masking of a Simple Auditory Stimulus," by James P. Egan and Harold W. Hake, *The Journal of the Acoustical Society of America* 22(1950):622-630.

page 137
Adapted from a graph in "The Acoustics of Singing," by Johann Sundberg. Copyright © 1977 by Scientific American, Inc. All rights reserved.

page 148 (bottom)
Adapted from an illustration in *Music, Acoustics, and Architecture*, by Leo Beranek, copyright by John Wiley & Sons, Inc.

page 151
Adapted from a drawing in *The Collected Papers on Acoustics*, by Wallace Sabine, Harvard University Press, 1922.

page 152
Adapted from a graph in *Architectural Acoustics*, second edition, by K. D. Ginn, Bruel and Kjaer, 1967.

page 153
Photography by James R. Holland/Black Star.

page 155
Photography by Ezra Stoller, © ESTO.

pages 156-158
Adapted from graphs in "Acoustical Measurements in Philharmonic Hall," by M. R. Schroeder, B. S. Atal, G. M. Seeler, and J. West, *The Journal of the Acoustical Society of America* 40 (1966):434-440.

page 159
Photography by Susanne Faulkner Stevens.

page 161
Adapted from a graph in "Acoustical Measurements in Philharmonic Hall," by M. R. Schroeder, B. S. Atal, G. M. Sessler, and J. West. *The Journal of the*

Acoustical Society of America 40 (1966):434-440.

page 163
Adapted from an illustration in "Construction of a Dummy Head after New Measurements of the Threshold of Hearing," by V. Mellert, *The Journal of the Acoustical Society of America* 51 (1960):1359-1361.

page 165
Courtesy of Manfred Schroeder.

page 166
Adapted from a drawing in "Binaural Dissimilarity and Optimum Ceilings for Concert Halls: More Lateral Sound Diffusion," by M. R. Schroeder, *The Journal of the Acoustical Society of America* 65(1979):958-963.

page 169
Adapted from illustrations in "Symposium on Wire Transmission of Symphonic Music and Its Reproduction in Auditory Perspective: System Adaptation," by E. H. Bedell and Iden Kemey, *Bell System Technical Journal* 13 (1934):301-308, © 1934 AT&T.

page 170 (top)
Courtesy of DAR Magazine.

page 170 (bottom)
Courtesy of Max V. Mathews.

page 176
Photography by Jacob/Black Star.

page 181
Courtesy of Ludwig Industries, Chicago, Illinois.

page 190
Adapted from a graph in "Electronic Simulation of Violin Resonances," by M. V. Mathews and J. Kohut, *The Journal of the Acoustical Society of America* 53 (1973):1620-1626.

page 192
Courtesy of Elizabeth A. Cohen.

page 193
Reproduced with permission from "Phonetics" in *Encyclopaedia Britannica*, fifteenth edition, © 1974 by Encyclopaedia Britannica, Inc.

264 SOURCES OF ILLUSTRATIONS

page 198

Adapted from a drawing in "Scaling the Musical Timbre," by J. M. Gray, *The Journal of the Acoustical Society of America* 61(1977):1270-1277.

pages 200, 202, 203

Photographs courtesy of Ludwig Industries, Chicago, Illinois.

page 204

Photography by Michael Samson/courtesy of Amy Malina.

page 206

The Metropolitan Museum of Art, gift of Mrs. Alice Lewisohn Crowley, 1946.

page 216

©BEELDRECHT, Amsterdam/V.A.G.A., New York, Collection Haags Gemeentemuseum—The Hague, 1981.

pages 242 and 243

Adapted from illustrations in *The Technology of Computer Music*, by Max V. Mathews, MIT Press, 1969.

Index

Page numbers in *italics> indicate illustrations.*

- Absorption coefficients, 162
- Acoustics, 1, 2. *See also* Architectural acoustics
- Amplifiers, 41
 - nonlinear distortion, 177-179
- Amplitude of sine waves, 39-42, 50, 51
- Analog instruments, 8
- Architectural acoustics, 150-163
- Atal, Bishnu, 162
- Audience noise, 136, 137
- Auditoriums. *See also* Concert halls
 - echoes, 146
- Avery Fisher Hall, New York, 145-146
- Babbitt, Milton, 5, 7
- Bach, Johann Sebastian, 73-74, 212
- Baryton, 47
- Bass, fundamental, 96-100
- Batteau, Wayne, 102
- Beats, 64-75
 - and consonance, 78-79, 84
- Békésy, Georg von, 108, 113
- Bells
 - pitch of, 95-96
 - sounds of, 37
- Beranek, Leo L., 155
- Borish, Jeffrey, 164
- Boston Symphony Hall, 153
- Boulez, Pierre, 90
- Brain, and musical ability, 208-209
- Brass instruments, 2. *See also* specific brasses by name
 - resonance in, 49
- Bregman, Albert S., 211
- Brightness of sound, 36, 37, 58, 59
- Cadences
 - and consonance, 84, 85
 - in stretched tones, 90-91
- Cagniard de la Tour, Charles, 15-16
- Cannon shot simulation, 76
- Carillon, 99, 202-203, 205
 - harmony, 98
- Cent (interval), 72-73
- Chavez, Carlos, 14
- Chimes, 200-201
- Chords. *See also* Cadences
 - and consonance, 78, 84-86
 - inversions, 98
 - Picardy third, 100
 - of stretched tones, 98
 - triad, 96, 98-100
- Chowning, John, 12 100, 144, 182, 195, 212, 217-218, 246
- Clappers, 14
- Cohen, Elizabeth, 91, 207
- Compass, musical, 16

- Composing, and computers, 8-9
 Computer
 and music, 2
 sound production, 219, 222
 Computer music, 5-13, 56
 Concert halls. *See also* Architectural acoustics; individual concert halls design of, 154
 Concert pitch, 15, 16
 Concerts, and computer music, 9-11
 Consonance, 76-101
 Consonants, and vocal tract, 196
 Constant-loudness curves, 122-123
 Constitution Hall, Washington, D.C., 168, 169
 Cook, Perry, 184, 195-196
 Cosine wave, 61-63
 Counterpoint
 first-species, 5-7
 and harmony, 87
 Critical bandwidth, 79-82, 87
 and hearing, 112
 and masking, 134
 Damaske, P., 162
 Dandrieu, François, 76
 Davis, Miles, 48
 Decibels, 116, 117, 128
 Deutsch, Diana, 209-211
 Diplacusis binauralis, 40
 Dissonance, 76-101
 Distortion, nonlinear, 177-179
 Dolby stereo, 174
 Dolson, Mark, 190
 Drums, pitch of, 205
 Dullness of sound, 58, 59
 Ear, 102-114. *See also* Hearing
 Echoes, 146-149. *See also* Reverberation
 and frequency and pitch, 31-33
 reflected waves, 29-30, 34
 Electronic music, 3-4, 8
 Equal temperament, 68, 70-74
 Eyring, Carl F., 160
 Filters, in sound analysis, 50, 53-55
 Flanagan, James L., 109
 Fletcher, Harvey, 168, 171, 179, 185, 187
 Folk music, 75
 Formants, 49, 50, 112, 191-195. *See also* Voice, human
 Fourier, François Marie Charles, 41-42
 Fourier analysis, 50, 185
 Fourier integrals, 44-45
 Frequency, 185-187
 analysis of, 49-56
 and consonance, 79-82
 of echoes, 31-33
 of formants, 49, 191-196
 fundamental, 109-111
 just noticeable difference, 140-143
 levels of and hearing, 112, 116, 125
 and masking, 131-136, 138
 and musical illusions, 214, 215
 and pitch, 17-23, 34, 37, 95
 ratios, 67-70, 73
 and reflected sounds, 147-149
 resonant, 45, 46
 of sine waves, 39-44, 61-63
 of sirens, 15-17
 and temporal acuity, 149
 Fundamental, 36, 46
 Fux, Johann Joseph, 5
 Gamelan, 47, 48
 Gilbert, Edgar N., 161
 Glockenspiel, 201
 Gong, 199
 Gottlob, D., 162-163, 165
 Green, David M., 149
 Grey, John M., 197, 211, 220
 Guttman, Newman, 109
 Haas effect. *See* Precedence effect
 Halet, Pierre, 215
 Harmonics. *See also* Beats; Partial
 and frequency components, 42-45, 52
 and human voice, 53
 Harmony, 76-78, 84, 87-101
 Harpsichord, 188, 189
 Harris, Cyril, 159
 Headphones, 167-168, 177
 Hearing, 102-114. *See also* Ear;
 Illusions, musical; Psychoacoustics
 and critical bandwidth, 79, 86
 and frequencies, 130
 just noticeable difference, 140-143
 precedence effect, 143-146
 and pure tones, 40-41
 stretched tones, 91
 temporal acuity, 149
 threshold of, 121-122
 Helmholtz, Hermann von, 1, 45-46, 78, 85-87, 98, 109, 184-185, 221
 Hemony, Frans and Pieter, 202
 Hemony bell, 202, 204
 Henry, Joseph, 144
 High-fidelity, 179
 Hiller, Lejaren A., Jr., 5, 7, 8
 Hooke, Robert, 23
Iliac Suite, 7
 Illusions, musical, 209-217
 of moving sounds, 217-219
 Institute for Perception Research, Eindhoven, 203
 Institute for Research and Coordination of Acoustics and Music (IRCAM), Paris, 179
 Instruments, musical. *See also* specific instruments and groups of instruments by name
 electronic, 180
 improvements in, 2
 resonators, 47, 49
 simulation of, 184
 sound levels of, 120, 124, 126
 sounds of, 197-199
 Intensity level of sound, 116-118, 121-123, 125-126, 223
 just noticeable difference, 140-142
 and masking, 130, 131, 134, 136, 137
 precedence effect, 143, 144
 Intervals, musical, 14, 70-71, 75
 cent, 72-73
 and consonance, 81-84
 consonant, 67, 87
 dissonant, 87
 ideal, 71-72, 74
 and string length, 17, 20
 Ionisation, 14, 15
 Isaacson, L. M., 5
 Isophons, 122-123
 Ives, Charles, 76, 77
 Just noticeable difference (jnd), 140-143
 Just temperament, 70, 72-74
 Kepler, Johannes, 24
 Kettledrum, 205, 206
 Keyboard instruments. *See also* specific instruments by name
 digital, 11, 12
 Knowlton, Kenneth, 216
 Krumphorn, 3
 Lafitte, Daniel, 246
 Lane, C. E., 131
 Lansky, Paul, 183
 Lebrun, Marc, 246
 Ligeti, György, 4
 Limen. *See* Just noticeable difference
 Line spectrum, 50-52
 Listening, 138
 Loudness, 115-129, 223
 perception of, 79
 Loudspeakers, 147, 176
 stereo systems, 172-173
 Louise M. Davies Symphony Hall, San Francisco, 165-166
 McGee, Howard, 48
 Marimba, 47, 201, 202
 Masking, acoustic, 79, 130-139
 Mathematical notation, 225-226
 Mathematics, and waves, 228-233
 Mathews, Max V., 5, 11, 30, 56, 74, 84, 182, 190-191, 239, 244
 Mayer, Alfred M., 130
 Mel scale, 37
 Mellert, V., 162
 Melody, and stretched tones, 98
 Mersenne, Marin, 20-22, 24, 30
 Microphone, 147-148
 and recording, 175, 176
 MIDI (Music Instrument Digital Interface), 248-251
 Mode, musical, 41
 Church, 74-75, 86
 Monochord, 23
 Moog, Robert, 8
 Moog synthesizer, 8
 Moorer, Andy, 9-11
 Moorman, Charlotte, 8
 Morrill, Dexter, 197
 Motion, and pitch, 17
 Mozart, Wolfgang Amadeus, 212, 213
 Music, electronic reproduction of, 167-179
 Music V, 239, 248
Musique concrète, 8
 Noise
 cocktail-party effect, 138
 in concert halls, 154
 levels of, 116-117, 119
 masking by, 134-136
 masking of, by tones, 136-138
 and pitch, 56-60

- Octaves
 sense of, 36
 stretched, 89-92
 tuning of, 64-66
- Olson, Harry F., 179
- Ondes Martenot, 8
- Orchestras
 intensity levels of, 137-138
 sound of, 2-4
 sound level of, 117, 124-125
- Organ, 35, 233
 sound of, 191
- Outdoor performances, 154
- Overtones. *See* Harmonics; Partials
- Partch, Harry, 8, 72-73
- Partials, 36, 37, 96-97, 100-101
 bells and gongs, 199-200, 202-205
 chimes, 200-201
 and consonance, 82-84, 87, 90
 and frequency components, 42-44, 52
 glockenspiel, 201
 hearing of, 113
 human voice, 53
 and musical illusions, 214-217
 sound intensity level, 127
 and sound quality, 186-189
 and stretched tones, 89-92
 stringed instruments, 47
 and time variations, 196-197
 violin, 49
- Pendericki, Krzysztof, 3, 6
- Pepys, Samuel, 23
- Perception
 categorical, 220-221
 of musical illusions, 209-217
- Percussion instruments, 181. *See also*
 specific instruments by name
- Periodicity, 15-17, 55
 and pitch, 15, 36-37
 and resonators, 45
 of sine waves, 39-40, 42, 44
- Phase of sine waves, 39, 40
- Phase vocoder, 55
- Philadelphia Academy of Music, 168
- Philharmonic Hall, Lincoln Center,
 New York, 154-159
- Phon, 122-124
- Phonemes, 220
- Piano
 partials of, 187-188
 pitch of, 94, 95
 strings of, 34
 string vibration, 46, 55
 tuning of, 64-68, 94
- Pitch, 14-15, 95, 223
 of drums, 205
 and echoes, 31-33
 and frequency, 17-23, 34, 37
 just noticeable difference, 140-142
 and masking, 130
 and musical illusions, 212, 214
 and noise, 56-60
 perception of, 109-111, 114, 221
 periodicity and, 17, 19, 36-37, 115
 of piano, 94-95
 place theory of perception, 109, 111
 residue, 93-95, 201-202
 of sine waves, 36-37, 41
- Plomp, Rainier, 78, 80, 81, 98
- Precedence (Haas) effect, 143-146
- Predictive vocoder, 183
- Primitive music, 14
- Psychoacoustics, 140, 217. *See also*
 Hearing
 Pure tones. *See* Sine waves
- Pythagoras, 17
- Pythagorean tuning, 71-73
- Quadrasonic sound, 173-174
- Radio Baton, 11-12
- Rameau, Jean Philippe, 96-100
- Ray tracing, 160-161
- Recording techniques, 167, 175. *See also*
 Sound, reproduction of
- Reflection of sounds. *See* Reverberation
- Resonance, 45-46. *See also* Formants
- Resonators, 45-47, 49, 51
- Reverberation, 144-146, 234-238
 artificial, 176
 concert halls, 154, 160
 echoes, 146-148
- Reverberation time, 146, 151-152,
 160-162, 165
- Rhythm, 14
- Risset, Jean-Claude, 190, 196, 212,
 214-217
- Rock music, 175
- Rodet, Xavier, 183
- Roughness of sound. *See also* Beats
 and consonance, 78-79, 84
- Royal Eijshouts Bell Foundry, the
 Netherlands, 203
- Royal Festival Hall, London, 166
- Sabine, Wallace Clement, 150-154,
 160
- Scales, 64-75
 and consonance, 88-89
 diatonic, 68-70, 72-75, 100, 101
 synthetic, 87-89
 tonal centers of, 74-75
- Schaeffer, Pierre, 8
- Schouten, Jan, 78, 92-94
- Schroeder, Manfred, 159-163, 165,
 171, 218, 219
- Schuster, K., 160
- SCORE program, 8
- Sensation level, and masking, 131-133
- Serra, Xavier, 206
- Shaw, Woody, 49
- Shepard, Roger, 69, 214-215
- Shouting, 115
- Shrillness, 36
- Siebrasse, K. F., 162-163, 165
- Sine waves, 36-46, 52, 61-63,
 216-217, 223. *See also* Periodicity
 and beats, 64-66, 78-79
 and consonance, 79-81, 84
 and hearing, 109, 112-113
 just noticeable difference, 141-142
 and loudness, 122, 127
 and masking, 134-137
 phase of, 39, 40
 and pitch, 95
- Singing. *See also* Voice, human
 formant, 191
 intensity level of, 137-138
- Siren, 15-17, 92-93
- Slaymaker, Frank H., 89
- Smith, Leland, 8
- Sone, 123-124
- Sonograph, 52
- Sound. *See also* Hearing; Sine waves;
 Synthesis of sound; Waves, sound
 absorption of, 144-145
 analysis of, 205-207
 decay of, 161, 188-189
 digital generation of, 239-247
 electronic, 3-4
 frequency analysis, 49-56
 illusions of, 217-219
 perception of and reverberation time,
 151
 power and loudness of, 115-129
 reproduction of, 167-179
 reverberant, 176-177
- Spectrum analyzer, digital, 50, 51
- Speech. *See also* Voice, human;
 Vowels; Whispering
 reverberation time, 144-146, 151
 synthesizing, 196
- Sperry, Roger, 209
- Square waves, 42, 43
- Stereo systems, reproducing sound,
 167, 168, 171-177
- Stockhausen, Karlheinz, 8
- Stokowski, Leopold, 168
- Stretched intervals and tones, 98, 100
- Stretched octaves, 89-92
- String length, and musical intervals, 17
- String vibration, 26-29, 33-34
 echoes, 30
 frequency of, 17-22
 and resonance, 46-47
- Stringed instruments. *See also* specific
 instruments by name
 vibrations in, 33-35
- Studio d'Essai, 8
- Sundberg, Johann, 137, 191
- Sympathetic strings, 47
- Synthesis of sound, 195-197, 205-207
 analog, 8-9
 digital, 9, 11-13, 182-187, 219-
 220, 222, 239-247
- Tape music, 3
- Tension, in stringed instruments, 34
- Terhardt, Ernst, 93
- Tetrachords, 74
- Theremin, 8
- Timbre, 223
 and musical illusions, 211-212
 perception of, 111-112, 221
 physical aspects of, 196-199
 and resonance, 49
- Timpani, 205, 206
- Tom-tom, 52, 53
- Tone, 223
 and masking, 136-138
- Trumpet, 196-197
- Tuning
 equal-tempered, 68, 70-74
 just temperament, 70, 72-74
 use of beats, 64-66
- University of California at Berkeley, 31
- Van Eyck, Jacob, 202
- Varèse, Edgard, 2, 3, 5, 14, 15
- Vibraphone, 201
- Vibrations, and sine waves, 41

- Vibrato, 60, 190, 196
- Viola bastarda, 47
- Viola d'amore, 47
- Violin. *See also* Stringed instruments
 - bowing, 189
 - resonance of, 50
 - sound of, 189–191
 - timbre of, 49
- Vocal tract
 - formants, 49, 52–53, 191
 - resonance of, 56
 - and waves, 195–196
- Voice, human, 191–196. *See also*
 - Speech; Vocal tract
 - sonograms of, 52–55
- Vowels
 - and formants, 49, 50, 112
 - and human voice, 191–196
 - and resonance, 56
- Waetzmann, E., 160
- Waves, sound, 16, 24–30, 58–63, 185–187, 223. *See also* Sine waves; Square waves
 - analysis of, 50, 53
 - and hearing, 130
 - intensity level, 116
 - and mathematics, 228–233
 - power of, 115–116
 - reflected, 147–148, 234–238
- Wegel, R. L., 131
- Wessel, David L., 211, 212
- Whispering, 56, 191
- White noise, 56, 136
- Wind instruments. *See* specific instruments by name
- Wire. *See* Strong vibration
- Woodwinds, 2, 49. *See also* specific instruments by name
- Xenakis, Iannis, 3
- Xylophone, 47, 201
- Young, Lamont, 8