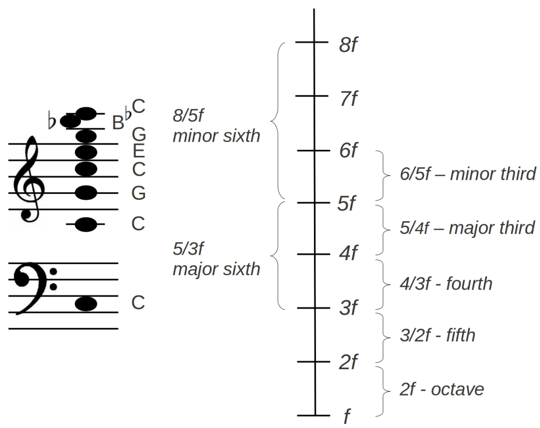
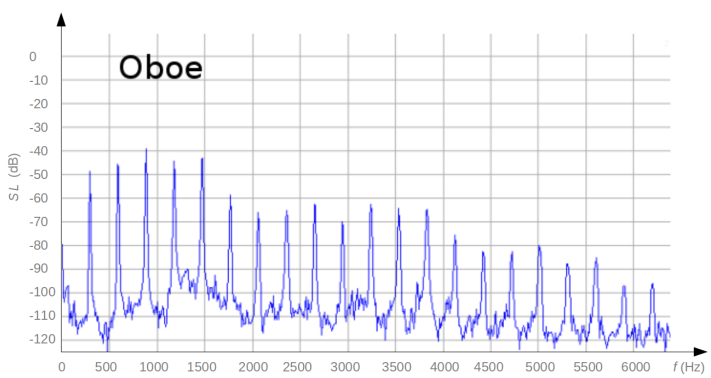
Physics of Music

Science and Art





Warren F. Rogers

Westmont College, Santa Barbara CA

To all my students, through all the years at Westmont College, especially to those who have taken, or are currently taking, the Physics of Music,

this book is dedicated to you.

Physics of Music: Science and Art
Warren F. Rogers, Westmont College

ISBN – not-yet-but-I-can-still-wish
1. Physics of Music, Physics, Music, Sound, Acoustics

Cover design, interior design, and figures by Warren Rogers, unless otherwise noted

Department of Physics Westmont College
955 La Paz Road
Santa Barbara, CA 93108

Printed in the United States of America
First printing January 2012
10 9 8 7 6 5 4 3

Table of Contents

Chapter 0

Preface – Science and Art

My hope is that you find this course to be a rich and rewarding exploration of art and science. The study of music (which, we will learn, is a *physical* phenomenon), using the tools, ideas, and methodology of physics, turns out to be a very fruitful venture, and our understanding and appreciation of each can magnify the other. While physics is not strictly an "art" – we would never characterize it as a fine or performing art – it does nonetheless require the exercise of extraordinary human imagination and creativity. And while music is not normally characterized as a "science," it might be possible, in the hands of a highly abstract and mathematically-minded composer, to see music as something akin to science.

Be that as it may, I hope to convince you (in case you harbor some caution or suspicion on the matter) that the study of music from a scientific standpoint can very much increase and enrich, rather than weaken or spoil your appreciation of music. You will learn how to listen to and understand music in new and exciting ways; you will come to understand many of the underlying scientific principles behind such topics as timbre, resonance, loudness, pitch, response of the human ear, harmony and dissonance, the physical function and operation of various musical instruments, the acoustical enhancement of music in a fine performance hall, the recording, playback, and synthesis of analog and digital music, to name a few. You will learn to *listen* to music differently – there will be subtleties for which you will develop a new sensitivity.

Why Physics and Music?

A good place to start is with some brief definitions. What is, after all, physics? And what is music? Is there a connection between the two? Physics, in short, is the study of the interaction between *matter* and *energy*. Of course, there is much more to it than that, but for the moment this simple definition should suffice. Physics is a "postulate"-based venture, meaning that it starts with some foundational assumptions called postulates, or "givens" that we understand (or assume) to be true about the world, and proceeds to build a model of understanding of how the world works. The language of physics is *mathematics*, and the concepts can be quite abstract and sometimes difficult to grasp with an untrained mind. Physics seeks to find underlying simplicity in seemingly complex systems. When it identifies a pattern that seems to persist regardless of how many times a system is tested or a measurement is made, physicists might elevate it to the status of "law," although, of course, they mean something very different than a law drafted for societal purposes. For example, the "law of gravity" expresses our observation about gravity's consistent behavior when tested over and over again. If the specific behavior of gravity depended on the mood of some Greek god in charge of its operation, we might have little reason to expect it to be consistent. But the clear pattern it exhibits is one of complete consistency – it is a pattern that persists through many and varied observations. There are many other similar laws of physics. Perhaps chief among them is the law of the conservation of energy, which we will get to in due time, and which will be very useful for our study of musical phenomena. Curiously, though, a law of physics is never actually able to be *proven* – in the mathematical sense – having to do with the nature of the scientific enterprise, a discussion for another time perhaps.

And what about a definition of music? Consider the raw sounds of nature – the call of a bird, the sound of a rushing river, the whistling wind through the trees – do these constitute what we would call music? I like what a very famous composer of the 20^{th} century, Igor Stravinsky, said in his lectures entitled "Poetics of Music."

> I shall take the most banal example: that of the pleasure we experience on hearing the murmur of the breeze in the trees, the rippling of the brook, the song of a bird. All this pleases us, diverts us, delights us. We may even say: What lovely music! Naturally, we are speaking only in terms of comparison. But then, comparison is not reason. These natural sounds suggest music to us, but are not yet themselves music. If we take pleasure in these sounds by imagining that on being exposed to them we become musicians and even, momentarily, creative musicians, we must admit that we are fooling ourselves. They are promises of music; it takes a human being to keep them: a human being who is sensitive to nature's many voices, of course, but who in addition feels the need of putting them in order and who is gifted for that task with a very special aptitude. In his or her hands all that I have considered as not being music will become music. From this I conclude that tonal elements become music only by virtue of their being organized, and that such organization presupposes a conscious human act. [1]

At its core the ideas offered by Stravinsky can help us to understand music as a *pattern of sounds* arranged in a particular way by an artistically sensitive human being. While nature can provide the raw material of music in the form of various beautiful sounds, it does not produce music on its own. It takes a person, a human being, to organize tones in an artistically meaningful fashion. The art of music is an inextricably human enterprise and can't be articulated fully in scientific terms. However, the *phenomenon of sound* can be studied scientifically, and herein lies the connection between physics and music. In order for there to be music, there must be *sound*, and it turns out that in order for there to be sound, there must be 1) physical motion, 2) a material medium to support the propagation of the motion's effect, and 3) an auditory system to detect the motion. You can think of these three essentials as the three P's: Production, Propagation, and Perception. In slightly more technical detail, the sound of music begins with the formation of *vibrations* that are produced by physical means, which cause a disturbance to propagate through the air as *compression waves*. Eventually, these waves reach the ear, at which point they set into vibration the eardrum, the three bones in the middle ear, and the fluid of the inner ear. At this point nerves send impulses to the brain which translates the phenomenon into the sensation of music. Physical vibrations, waves, motion ... these concepts belong to the domain of physics, and are at the same time the "stuff" of music. Therefore, music is special among the arts in that it is so deeply rooted in the interaction of matter, energy, time, and space. It is impossible to experience music without the active participation of each of these physical phenomena (unless you refer to music that you hear "in your head," which is not music at all, but merely the *memory* of music).

So, in conclusion, the study of physics and music together provide a wonderful pair for the fulfillment of a general education physical science course requirement since:

> 1) music is based on physical phenomena described and understood by the principles of physics, and
> 2) the understanding and appreciation of music can be made much richer by learning the physical principles at work behind it.

Place in a Liberal Arts Education

At Westmont we embrace a Liberal Arts model of education. This course fits very nicely into our mission of providing a broad education. For one thing, this course has a strong multi-disciplinary character, bringing together an art and a science under one roof. It helps to compare briefly the motivations behind the study of each. A metaphor that I have greatly enjoyed that seems to capture one of the essential aspects of a liberal arts education is that of *seeing*. Both science and art provide us unique ways of "seeing" aspects of what it means to be human. On the one hand, science provides us with unique eyes for seeing, studying, and understanding the physical world, just as our physical eye together with the optic nerve and the brain, enable us to "see" the world around us, of which we are a part. While we are clearly using the word "seeing" metaphorically here, there is a very real way in which the apparatus of scientific experimentation, as well as the apparatus of our scientific ideas, models, and theories of nature, coupled with the results of careful experimentation, really do enable us to see the world through new eyes. Likewise, the arts enable us to better understand, to "see," or better appreciate the creative impulse of humanity, and that which drives us to the passion and ability to create expressions of beauty.

If you're not accustomed to the study of physical science and are a bit tentative about engaging a topic that is perhaps outside your comfort zone, you might find another metaphor to provide some helpful encouragement. To quote a very musical Mary Poppins, "a spoonful of sugar helps the medicine go down." The sweet art of music (*i.e.* the sugar), for which nearly everyone has a deep appreciation at some level, provides a wonderful vehicle for the study of physical science (*i.e.* the "medicine" – though in truth physics is already a sweet deal if you ask me - but then I'm biased!). But rather than take my word for it, you now have the chance to decide for yourself. My hope is that you enjoy this course, that you come away with a deeper understanding of physical science, and that you find your appreciation for and understanding of music richer and more meaningful for the journey.

References

[1] Stravinsky, *The Poetics of Music*, pg. 34 (1974)

Chapter 1

Music, Time and Periodicity

1.1 Time and Music

Music is unique in the arts, in that it is "bound to time." One cannot experience a musical piece, whether it derives from a live performance or a recorded copy, without spending the time it requires for us to pause and listen. You are "held captive" for the time it takes the piece to unfold, much like a flower, as it reveals in sequence the elements of its musical narrative. You can't change a piece, either by speeding it up or slowing it down, without substantially altering its artistry. While you can play fragments of a whole and enjoy the sound of them individually, they are nevertheless divorced from the organic whole and thus aren't experienced in the larger narrative context crafted by the composer. And the aspect of time does not merely apply to the piece's duration. As you drill down into the detail of its musical structure, the sequencing and execution of the individual notes and phrases (including their dynamic expression, changes in speed, pauses, etc.) rely importantly on their expression in time.

While different performances of the same piece might vary in their overall tempo (*i.e.* speed) and in the details of their expressive dynamic throughout, the quality of the performance relies on the performers' sensitivities to this important aspect of time. We can study a musical score at leisure, apart from actually listening to the piece, which gives us more of a "bird's eye view" of the piece, but this pastime will never replace the critical experience of enjoying and absorbing the actual performance, in time, on its own terms. This feature of its strict reliance on time differentiates music from the visual arts. A painting, sculpture, or architectural masterpiece can be viewed at one's leisure. You can stand and enjoy it anytime you like, for as long or short as you care to, and the art can still effectively communicate its essence. The enjoyment of a musical piece on the other hand requires that we be attentive and available at a particular time, for a specific duration, and that we direct our attention from beginning to end on each moment that passes, in order that the piece can effectively communicate its narrative and tell its full story.

1.1.1 Musical Respiration and Levels of Time

Life happens in time. It has an ebb and a flow to it that brings variation and season. There are changes to our routine, to our emotions, to our weather, and to our health.

As with life, music too breathes with an ebb and a flow, a dynamic that unfolds through passages of similarity and contrast, tension and resolution, with passages filled with uncertainty and others offering resolution and comfort. Perhaps it is from our experience of time, and change, and seasons in life that music derives part of its deeply moving and transformative power. Music is dependent on time in three ways worth noting.

Macroscopic On the macroscopic or large-scale level, musical expression has a *duration*, that is, a single musical piece tells its narrative over the course of a pretty well determined time period. The musical narrative unfolds in time, much like a story, with sections of similarity and of contrast, sections that develop tension and sections that resolve it, sections that explore unknown territory and generate uncertainty, and other sections that bring the listener home to the familiar, all of which are part of a woven narrative that is unique and internally self-consistent, much like a well-written story. All of this happens in time, and while it is not strictly *periodic*, it does unfold in a way that relies heavily on time.

On a related note . . .

Musical Tempo Indicators

The Italian word "Andante" is one of several expressions used in Western Music to indicate the speed at which a piece of music is intended to be played. Andante connotes something akin to "walking" speed, typical of a casual stroll, whose frequency is, incidentally, close to that of the beating heart. The terms range from slow or ultra-slow, such as *Lento*, *Largo*, or *Grave,* to the fast *Allegro* or *Vivace*, to the ultra-fast *Prestissimo*. While these indicators are only guides, it is up to the musician to determine how that indication will inform the interpretation of the piece. Some composers in the past have included more specific metronome markers at the beginning of their pieces (i.e. how many beats per second) to provide a more precise indicator of their intentions (Ludwig van Beethoven and many others have done this in some of their pieces), and some have provided no indicators whatsoever (it was generally not practiced in the Baroque era; Johann Sebastian Bach used no such numerical or Italian phrases in most of his music).

As we become absorbed in listening to a piece, our very sense of the passage of time itself may become altered. In musical performance, fine artistry consists of the shaping of music both in time and in dynamic (or "loudness") so that the music tells its story with an organic wholeness. There are many other elements that contribute to a fine musical performance, but these two elements in particular are straightforward, spoken about in physical terms, and we shall be doing that shortly.

Mesoscopic On the mesoscopic, or medium-scale level, music has meter and rhythm. *Meter* corresponds to the basic grouping of time units that structures the piece. It specifies how many beats correspond to one measure of the music. While more periodic than the macroscopic structure, there is also typically variation in this time structure throughout the piece. Rhythm relies on the complex interplay between strong and weak beats, providing much of the emotion-setting tone of the piece.

Microscopic And then there's the microscopic, or the small-scale time structure of music. We are referring here to the rapid periodic vibrations that give rise to sound waves and ultimately to our sense of pitch. Variation in the speed of vibration yields variations in pitch, thus allowing musical expression in terms of *melody*. Similarly, the timbre, or "personality" of an instrument (the tone color that gives rise to its unique and recognizable voice) is based on the specific collection of simultaneous vibrational frequencies it emits.

1.2 Periodic Motion

In order to understand better how the dimension of time is woven into and foundational to the phenomenon of music, we need to better understand the important role of *repetition*. What I'm *not* referring to is repetitive music! There is plenty of that around, and we'll leave it to someone else to speculate on the artistic merits of excessive repetition. The kind of repetition we want to consider is *periodic motion*. A cycle of motion that repeats itself within a characteristic time is called periodic. Examples of periodic motion could be the flapping of a bird's wings, or the volley of a tennis ball back and forth across the net. The amount of time required to complete one full cycle of motion is defined as the *period*. Figure 1.1 shows a graphical representation of a particularly simple type of periodic motion – sinusoidal. This graph might refer, for example, to the motion of a fishing pole tip being swished back and forth. The x-axis corresponds to time, and the y axis to the distance that the tip of the pole is from its resting position.

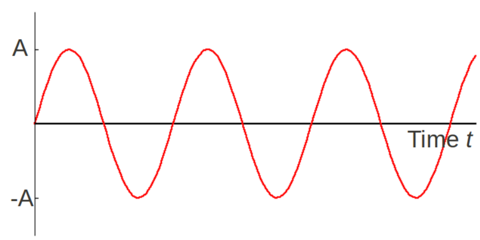


Figure 1.1: Graphical representation of the simplest kind of periodic motion (sinusoidal), that of a simple harmonic oscillator.

On a related note . . .

The Human Heart

One of the most important timepieces we have available to us is our heart. There is good reason to believe that the rate at which the human heart beats provides a "standard" rate around which musical structure is organized. The healthy heart beats periodically at around 60 to 80 beats/minute. The rhythm of many musical pieces unfolds at rates related to this frequency. A typical waltz (whose meter is usually 3/4, giving the familiar "*1-2-3, 1-2-3*" dance rhythm) completes one 3-beat cycle – its "respiration rate" – in approximately one heartbeat. Music that respirates at faster than 5 times the human heart rate can be too fast to follow and create excessive tension. Music that respirates at a much slower pace of one cycle per 10 heartbeats is much less likely to capture the interest of the listener.

The world is full of periodic motion: the ocean tides, the phases of the moon, a duck bobbing on the waves, the wings of a bird in flight, tree branches blowing in a strong wind, a child on a swing, or the rising and setting of the sun (more accurately, the rotation of the earth!). While there are many examples of periodic motion, only some of them are *simple harmonic* in character. Simple harmonic motion is the simplest kind of periodic motion, expressible in very simple mathematical form - that of a sinusoidal wave (as depicted in figure 1.1). Of the examples mentioned above, the child on a swing probably comes closest to this ideal. It turns out that simple harmonic motion is absolutely foundational to music, and so we will consider it in much greater detail in the coming chapters.

1.3 Chapter Summary

Key Notes

- Music is "bound to time" in that it must be experienced over a specific period of time to be realized, unlike the majority of other art forms.
- Music has "respiration" – it embodies variety and differentiation in time.
- Temporal variation in music occurs on three levels, the *macroscopic*, corresponding to the unfolding of music's story with the ebb and flow of variation in its expression, the *mesoscopic*, corresponding to its expression in rhythmic terms, and the *microscopic*, corresponding to the periodicity of vibration of the individual tones that gives rise to pitch, and therefore melody.
- A cycle of motion that repeats itself with a characteristic time is called periodic. Periodic motion is central to music.

♪ ♪ ♪

Chapter 2

Basic Physical Quantities

2.1 Length, Mass, Time

Progress in physics (and more generally in science) is based on *experimentation.* On the basis of observations derived from experiments, scientists develop theories to help us assemble an increasingly rich understanding of how nature works. These theories can then make predictions about additional behaviors or phenomena not yet seen, and can be tested by experiment. If the results of the experiment are consistent with the prediction, the theory is advanced; if not, the theory is discarded or modified. In this way theory and experiment work hand-in-hand to advance our knowledge about the world.

We live in a universe characterized by three *spatial dimensions* and one *time dimension.* Furthermore the volume of the universe is populated by bodies with *mass.* The three foundational quantities with which we will be primarily concerned are therefore *length*, *mass*, and *time.* Each of these can be expressed using different *units.* For example, length can be expressed in miles, inches, light-years, microns, meters, yards, *etc.* We tend to use different units for different circumstances. For example, in America we measure our own height in feet and inches, but our road trips in miles; we measure our weight in pounds but our trucks in tons; we measure the volume of our rooms in cubic feet, but our recipes in cups and tablespoons.

It will help us to have a *standard* set of units with which to work. In physics it is standard to work in the internationally agreed upon Système Internationale (SI). We will be adopting that standard throughout this course. We will refer to it as the “MKS” system because the standard units for length, mass and time are the *meter* (m), the *kilogram* (kg), and the *second* (s).

Whenever you solve a problem in this course, it is highly recommended that you first convert all of the given quantities to their MKS unit representations, so that all resulting quantities also end up in their appropriate MKS representations. For example, we will soon see that the MKS unit for *force* is the Newton (N), for energy is the Joule (J), and for power is the Watt (W). These units can each be expressed as products and quotients of meters, kilograms, and seconds. For example, in terms of these basic quantities, one Newton can be expressed as follows:

$$1\text{ N} = 1\,\frac{\text{kg}\cdot\text{m}}{\text{s}^2}$$

When calculating force, as long as we use all of the basic quantities in their MKS representations and we perform our calculations correctly, then we are guaranteed to end up with a result that is in N.

2.2 Matter and Energy

The study of physics basically boils down to the study of the interaction between mass and energy. There are the two main constituents of our corner of the universe, and the interaction between them forms all of the dynamics of nature that surround us. Scientists have recently discovered that there is a lot more "stuff" that makes up our universe, and that the mass and energy with which we are familiar (and of which we are made) forms a very small fraction of this universe. The other two recently discovered quantities, about which we still know very little, are "dark matter" and "dark energy." We won't be concerned with these quantities in this course, but it is helpful to point out that we live in a very interesting universe about which there is still much to discover.

You might be wondering how the concepts of physics apply to the study of sound and music. As we will see very shortly, the source of sound is *vibrations*, which are mechanical in nature and involve mass and energy. Sound propagates through the air as *traveling waves*, consisting of vibrations that pass from one air molecule to the next. The perception of sound involves the transmission of these vibrations from the air to the eardrum and then to the middle and inner ear where they are then translated into nerve signals and sent to the brain for processing. All musical instruments are made of mass, and it is only through the input of energy by the performer that they can make any sound at all.

2.3 Mathematical Considerations

2.3.1 Dimensions

All physical quantities have "dimension." The simplest of these are *length*, *mass*, and *time*. All related physical quantities (energy, force, *etc.*) we will be encountering in this course can always be expressed as products and/or quotients of these basic dimensions.

It is important to recognize that dimensions need to be treated just as algebraic variables when solving problems. Additionally, the dimensions of all quantities being added or subtracted must have the same dimensions. It makes no sense to add 5 kg to 16 m since the first is a mass and the second is a length, but it makes perfect sense to add 5.3 m and 2.9 m since both are lengths. Likewise, the units on both sides of an equation must be the same.

2.3.2 Notation for Differences

In this course we will be dealing with *differences* in some quantities. For example, if a sound wave moves from one position to another, starting at position x_1 at time t_1 and arriving at position x_2 at time t_2, the total distance the sound wave has moved, denoted by use of the "Δ" symbol, is

$$\Delta x = x_2 - x_1$$

and the time it took to move that far is

$$\Delta t = t_2 - t_1.$$

2.3.3 Scientific Notation

Some of the numbers we will be dealing with in this course will be very large or very small. We will therefore express some of these numbers using exponential notation. Table 2.1 lists some numbers expressed more conveniently using scientific (exponential) notation.

Example 2.1

Multiplication of Numbers with Exponents *Multiplying numbers with exponents involves adding the exponents. Multiply* 5.10×10^{-6} *by* 23.5×10^{18} *by* 1.61×10^{-19}.

Solution: We need to multiply the three numbers and add the powers of ten:

$$(5.1 \times 10^{-6}) \cdot (23.5 \times 10^{18}) \cdot (1.6 \times 10^{-19})$$
$$= (5.1 \times 23.5 \times 1.6) \cdot (10^{-6} \times 10^{18} \times 10^{-19})$$
$$= 193 \times 10^{(-6+18-19)} = 193 \times 10^{-7} = \boxed{1.93 \times 10^{-5}}$$

2.3.4 Significant Figures

The number of significant figures (or "sig figs") in a number generally correspond to the number of non-zero digits appearing in it, and are related to the precision with which the number is given. Digits of "0" that appear *between* non-zero digits count as significant, as well as 0's that occur before an actual decimal point. For example, the number "12,000" (with 2 sig figs) is less precise than the number "12,040" (with 4 sig figs), and both are less precisely known as the number "12,040." which has 5 sig figs. Zeros that are used to position the decimal point such as in 0.02 or 0.000035 are not considered significant; in this case the first has 1 sig fig and the second has 2. Zero digits that appear after non-zero digits to the right of the decimal point are significant, as in the number 0.06800 which has in total 4 sig figs, which include the 2 non-zero digits plus the 2 zeros to the right of them.

Including an explicit decimal point in a number can render the 0's following non-zero digits significant. For example the number "100" has 1 sig fig, whereas "100." has 3, and "100.000" has 6. Finally, it can help to use scientific notation in cases where ambiguity is possible. For example, if the number 35,000 is actually known to 3 sig figs (which is not implied by the way the number is currently written), then it becomes clear when we express it using scientific notation as 3.50×10^4 which clearly has 3 sig figs.

The number of sig figs in a number generally correspond to how precisely the number is actually known, as in the following example.

Table 2.1: A few numbers and their values expressed in exponential notation

1,600,000	=	$1.6 \times 10 \times 10 \times 10 \times 10 \times 10 \times 10 = 1.6 \times 10^6$
0.000064	=	$6.4 \times \dfrac{1}{10 \times 10 \times 10 \times 10 \times 10} = 6.4 \times 10^{-5}$
0.00314	=	3.14×10^{-3}
51,677,000	=	5.1677×10^7
299,800,000	=	2.998×10^8

Example 2.2

Precision of a Town's Population *Suppose we are given the population of a small town as 18,000 people. To what precision do we know the number of people populating this town? If we are given instead a population of 18,030 people, what is the new precision for the population?*

Solution: Since in the number 18,000 only the first 2 digits are significant, the actual population, which was presumably rounded up or down to give this number, could be as high as 18,499 (in which case it is rounded down to 18,000) or as low as 17,500 (in which case it was rounded up to 18,000). Therefore the number 18,000 has a total uncertainty in its "exact" value of approximately ± 500 so that we could express the town's population as $18,000 \pm 500$ people. In order to calculate the precision of this number, we divide the uncertainty in the number by the number itself to get

$$\frac{500}{18,000} = 0.028$$

which is 2.8%. We therefore know the town's population to a precision of about 2.8%, or about 3 parts in 100.

If we are instead given the population value as 18,030 people, which now has 4 sig figs, the actual number could be as high as 18,034 (to be rounded down) or as low as 18,025 (to be rounded up) so that the number could be expressed as $18,030 \pm 55$ people, and its precision becomes

$$\frac{5}{18,030} = 0.00028.$$

We now know the population of this town to about 0.028% or to about 3 parts in 10,000, a much more precise value.

When working on a problem for which you are given the values of more than one variable, the result of the calculation should generally contain the number of significant figures corresponding to the least precisely known variable. For example, the product $35 \times 4598 = 160,930$ according to my calculator, but should be expressed with only 2 sig figs, so that the final result should appear as 160,000. It doesn't hurt, however to retain one extra digit in cases where judgment would indicate better clarity. In this case, I might choose to represent this number as 161,000 by rounding off the 4th digit. Generally it doesn't usually hurt to retain one more digit than the number of sig figs of the least precise input variable.

2.3.5 Conversion of Units

When we are given a number in non-MKS units and want to convert it to its MKS value, we can use the conversion values given in Appendix A. We want to multiply the number by "1" so that we don't change its value, but in such a way as to change its units. Hence, to change 12.5 feet to meters, we use the conversion factor 1 ft = 0.3048 m to get:

$$\frac{0.3048 \text{ m}}{1 \text{ ft}} = 1 \quad \rightarrow \quad 12.5 \text{ ft} = 12.5 \,\cancel{\text{ft}} \times \left[\frac{0.3048 \text{ m}}{1 \,\cancel{\text{ft}}}\right] = 3.81 \text{ m}.$$

Note that the quantity in the square brackets is numerically equal to 1 so that in multiplying by it we don't change the value of our number. Note also that we need to express the conversion fraction with "m" on the top and "ft" on the bottom so that "ft" cancel, leaving the result in units of "m."

Example 2.3

Speed of Sound *The speed of sound in air is 771 miles per hour (or MPH). In order to produce a sonic boom, an aircraft must exceed this speed. What is this speed, expressed in m/s?*

Solution: We want to convert the units for length from miles to meters, and the units of time time from hours to seconds. The best way to do this is to multiply the initial speed by "1" in such a way as to change the units. Knowing that there are 1610 meters in a mile, we can form two different ratios using these units, both of which are equal to 1:

$$\frac{1 \text{ mi}}{1610 \text{ m}} = 1 \qquad \frac{1610 \text{ m}}{1 \text{ mi}} = 1$$

In order to convert 771 miles to meters, we want to multiply by whichever of the two ratios causes the "mi" unit to cancel and the "m" unit to replace it. The second of the two ratios accomplishes this purpose:

$$771 \text{ mi} = 771 \cancel{\text{mi}} \cdot \left(\frac{1610 \text{ m}}{\cancel{\text{mi}}}\right) = 1,241,310 \text{ m}$$

Similarly, we can convert the time from "h" to "s" using another conveniently chosen factor of 1:

$$1,241,310 \frac{\text{m}}{\cancel{\text{h}}} \cdot \left(\frac{1 \cancel{\text{h}}}{3600 \text{ s}}\right) = 345 \text{m/s}$$

And so the entire conversion can be calculated in one step by applying two multiplicative factors, each of which is equal to 1 and able to change the units the way we need:

$$771 \frac{\cancel{\text{mi}}}{\cancel{\text{h}}} \cdot \left(\frac{1610 \text{ m}}{\cancel{\text{mi}}}\right) \cdot \left(\frac{\cancel{\text{h}}}{3600 \text{ s}}\right) = 345 \text{m/s}$$

2.3.6 Trigonometry

Consider the right triangle depicted in figure 2.1. The length of the hypotenuse is c, and the horizontal and vertical sides have lengths a and b, respectively. The angle the hypotenuse makes with the horizontal x-axis is θ. The basic trigonometric relationships between the three length variables a, b, c and the angle θ are as follows:

$$\boxed{\sin\theta = \frac{\text{side opposite to } \theta}{\text{hypotenuse}} = \frac{b}{c}}$$

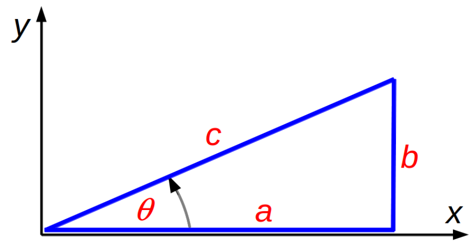


Figure 2.1: A triangle with sides a, b, and c and angle θ.

$$\boxed{\cos\theta = \frac{\text{side adjacent to } \theta}{\text{hypotenuse}} = \frac{a}{c}} \tag{2.1}$$

$$\boxed{\tan\theta = \frac{\text{side opposite to } \theta}{\text{side adjacent to}} = \frac{b}{a}}.$$

Since the triangle in figure 2.1 is a right triangle, the relationship between the three side lengths can be expressed using the Pythagorean Theorem, which states

$$\boxed{c^2 = a^2 + b^2 \quad \rightarrow \quad c = \sqrt{a^2 + b^2}}. \tag{2.2}$$

Likewise, we can express the horizontal and vertical sides of the triangle in terms of θ,

$$a = c\cos\theta$$

$$b = c\sin\theta \tag{2.3}$$

and we can compute the angle θ from the sides a and b,

$$\tan\theta = \frac{b}{a} \quad \rightarrow \quad \boxed{\theta = \tan^{-1}\left(\frac{b}{a}\right)} \tag{2.4}$$

where tan^{-1} is the "arctangent" function.

2.3.7 Logarithms

The use of logarithms is helpful for "squeezing down to size" an extraordinarily large range of numbers.

Definition: **The logarithm of A is the *value of the exponent to which we must raise 10* in order to obtain A**

Read that definition a few more times until you get a good grasp of its meaning. In mathematical form,

$$\boxed{\text{If } \; A = 10^B \quad \rightarrow \quad \log A = B}$$

Logarithms are not difficult to work with once you get a good intuitive sense of what they represent. Here are some examples of some numbers and their logarithms:

$$\begin{aligned}
1000 = 10^3 \quad &\rightarrow \quad \log(1000) = 3 \\
100 = 10^2 \quad &\rightarrow \quad \log(100) = 2 \\
10 = 10^1 \quad &\rightarrow \quad \log(10) = 1 \\
1 = 10^0 \quad &\rightarrow \quad \log(1) = 0 \\
0.1 = 10^{-1} \quad &\rightarrow \quad \log(0.1) = -1 \\
0.01 = 10^{-2} \quad &\rightarrow \quad \log(0.01) = -2 \\
2 = 10^{0.301} \quad &\rightarrow \quad \log(2) = 0.301 \\
5 = 10^{0.699} \quad &\rightarrow \quad \log(5) = 0.699 \\
8 = 10^{0.903} \quad &\rightarrow \quad \log(8) = 0.903
\end{aligned}$$

2.4 Chapter Summary

Key Notes

- The fundamental dimensions with which we are concerned are *length*, *mass*, and *time*, measured in meters (m), kilograms (kg), and seconds (s), respectively.
- Physics is the study of the interaction between *mass* and *energy*.
- Measured quantities have units that must balance on both sides of any equation, and which can be expressed as products and/or quotients of meters, kilograms and seconds.
- Also important is that answers to computations have an appropriate number of significant figures, consistent with the number in the input variables.
- Logarithms are used when we need to make a very large range of numbers fit into a more manageable range (for plotting in a graph, for example). We use logarithms in this course to express the very large range of intensity to which the ear is sensitive (from 10^{-12} to 1 W/m^2, a range extending over a trillion) into a more manageable range from 0 to 120 dB.

♪ ♪ ♪

Exercises

Questions

1) Briefly, why is it important to convert physical variables given in a problem to their MKS unit representation?

2) What are significant figures, and what role do they play in the representation of a numerical value?

Problems

1. A carpet is to be installed in a music studio in order to soften the acoustic character of the room. The room has length 10.55 m and width 6.5 m. Find the area of the room. Make sure your answer retains the correct number of significant figures.
2. How many significant figures are there in (a) 78.9, (b) 3.788, (c) 2.46×10^{-6}, and (d) 0.0032.
3. The speed of light is now defined to be 2.99792458×10^{8} m/s. Express this speed of light to (a) 3 sig figs, (b) 5 sig figs, and (c) 7 sig figs.
4. The edges of a shoebox containing a set of tuning forks are measured to be 11.4 cm, 17.8 cm, and 29 cm. Determine the volume of the box retaining the proper number of significant figures in your answer.
5. Find the height or length of these natural wonders in km, m, and cm: (a) The longest cave system in the world is the Mammoth Cave systems in Central Kentucky, with a mapped length of 348 miles. (b) In the United States, the waterfall with the greatest single drop is Ribbon Falls in California, which drops 1612 ft. (c) At 20,320 ft, Mount McKinley in Alaska is America's highest mountain. (d) The deepest canyon in the United States is King's Canyon in California, with a depth of 8200 ft.
6. Suppose the hair on the first violinist in the Chicago Symphony Orchestra grows at a rate of $1/64^{th}$ inch per day. Find the rate at which her hair grows in nanometers per second (1 nm = 10^{-9} m). Since the distance between atoms in a molecule is on the order of 0.1 nm, how many atoms/sec does this rate of length increase represent?
7. The speed of sound at room temperature is 345 m/s. Express this answer in km/h and mi/h.
8. The cellist of a famous quartet is late for a concert and travels along a highway at a speed of 38.0 m/s where the speed limit is 75.0 mi/h. Is the cellist exceeding the speed limit? Justify your answer.
9. In order for a stage manager to install a spot light, he leans a ladder against the interior wall of the concert hall. The ladder has length 9.00 m and is inclined at an angle of 75.0°to the horizontal. What is the horizontal distance from the bottom of the ladder to the wall?
10. A right triangle has a hypotenuse of length 3.00 m and one of its angles is 30.0°. What are the lengths of (a) the side opposite the 30.0°angle and (b) the side adjacent to the 30.0°angle?
11. Find the following logarithms: a) log(50) b) log(0.5) $\log(2 \times 10^{10})$ d) log(16)
12. Given $\log(x)$, find the number x in each case: a) $\log(x) = 0.3$ b) $\log(x) = 3.0$ c) $\log(x) = 1.3$ d) $\log(x) =$ -0.3

Chapter 3

Sources of Musical Sound

In order for sound to be produced, three elements are required:

1. a ***source*** to generate the sound - ***Production***,
2. a ***medium*** in which the sound can travel - ***Propagation***, and
3. an ***auditory system*** to receive and interpret the sound - ***Perception***

Each of these elements involves important physics principles that we need to consider in order to build a solid understanding of sound and music. In this chapter we will learn several important principles behind how sound is *produced*. In chapter 4 we will consider the second element, important for understanding how sound *propagates* from source to receiver, and in chapters 7 and 8 we will consider the third element, to better understand the *perception* of sound.

3.1 Origin of Sound - Vibrations

Something must *vibrate* in order for sound to occur. When a violinist plays a musical note, what is it that needs to take place in order for sound to be produced? Pulling her bow across a string produces a continual "stick-slip" action on the string, causing it to vibrate. This initial string vibration sets other vibrations into motion (in the body of the violin and the air inside and outside the violin) in a chain of cause and effect until at last the ear perceives the sound by vibration of the eardrum and bones of the middle ear.

A closer look at this sequence of events reveals that the string transmits some of its vibrational energy to the bridge of the violin (see figures 9.2 and 9.3) which causes the body and the air contained within it to vibrate. The air inside the body transmits vibration to the air outside via the f-holes. Vibrating air molecules bump up against their neighbors, which in turn bump up against theirs, and in this way the vibration moves along through the air. Likewise, the vibrating body of the violin bumps up against the adjacent air molecules, which in turn bump up against their neighbors, *etc.*. This vibrational *wave* (called sound), initially set into motion by the bow on the string, **propagates** throughout the room. *Every air molecule in the room* inherits this vibration. In fact, *wherever* the sound of the violin can be heard, whether it be in the next room or outside very far away, air molecules are vibrating as a direct result of the bowing of the string. Some of the vibrational waves enter the auditory canals of a listener, and air molecules adjacent to the eardrum cause it to begin vibration, which sets into motion the three tiny bones in the middle ear which cause vibrational waves to propagate in the fluid of the inner ear. Nerve cells then send electrical signals to the brain via the auditory nerve for processing by the auditory cortex, resulting in ... *voila*! the **perception** of sound.

Notice how central the role of vibrations is to *all three* of the requirements for sound mentioned above, production, propagation, and perception. Since vibrations are so important to the phenomenon of sound, we've devoted this entire chapter to understanding the physics principles behind them.

3.2 Basic Physical Concepts

In order for any system to vibrate and produce sound, it must have *mass*, and additionally there needs to be some *elastic force* that operates on that mass. For sound to propagate from one point to another it moves through a medium that also must have mass and internal elastic forces acting within it. For sound to be perceived, the sympathetic vibrations in the auditory system rely on elements that have mass and internal elastic forces. These two features, mass and elasticity, operate on basic physical principles to which we now turn.

3.2.1 Terms and Their Meanings

In the field of physics, the terms *force*, *pressure*, *work*, *energy*, and *power* have very precise meanings. When used in everyday speech they are commonly applied with a much broader scope than is consistent with their more narrow definitions in physics. For example, we often talk about someone's force of persuasion or the inexorable force of destiny; of being forced into making a decision or of people forcing their way up the corporate ladder; lawyers might apply pressure to witnesses, and we find ourselves under pressure to meet deadlines; we talk about homework, of going to work, or of working out our differences; we consume energy drinks to stay alert (*I* don't, unless you consider tea an energy drink ...); we require energy to finish a project; and we eat power bars and go on power walks. Let's now consider these terms according to their precise definitions in physics.

3.2.2 Mass

On the Earth we find all of the elements of the periodic table, each consisting of atoms that have *mass*. Our solar system, located in the galaxy called the Milky Way, consists of the sun, several planets, comets, asteroids, meteoroids, and much more, all of which have mass. The universe is a very large place, consisting mostly of empty space, dotted with billions of galaxies, each containing tens to hundreds of billions of stars and planets, and each of which has mass.

Physicists divide mass into two categories: *gravitational mass* and *inertial mass*. We will be concerned primarily with inertial mass, and how it affects the production, propagation, and perception of sound. The distinction between the two types of mass is a little subtle but worth exploring briefly.

We perceive the *gravitational* mass of a body at the surface of the earth as *weight*. The earth pulls on a body with a force that is proportional to its gravitational mass. A piano has significant mass; at the surface of the earth it can weigh 1000 lbs or more and you will find it difficult to lift. In contrast to this, we perceive the *inertial* mass of an object by its resistance to being moved. You might say "well of *course* it's difficult to move the piano, because it's so heavy!" and believe that it is the gravitational mass that makes it hard to move. But consider a thought experiment. If a piano were to float out in a remote region of deep space, it would have *no weight* since there would be no planets or stars to pull on it with gravity, and so it simply would drift. If you were to float up to it in your spacesuit and give it a hard shove, you would note that it is quite difficult to get moving. As a matter of fact, your push would set you into much faster motion away from the piano than its motion away from you – a direct result of your having much less inertial mass than the piano. The piano's resistance to being moved derives from its large inertial mass, independent of whether there are any massive bodies nearby to exert gravity on it.

Physicists have determined to a very high degree of precision that the gravitational and inertial mass of an object are directly proportional to one another. That is, if we double the gravitational mass of an object, its inertial mass also doubles. Somehow the two are intimately linked, yet separate properties of the object. A very heavy object (large gravitational mass) is also difficult to get moving (large inertial mass). Comparatively, a flute weighs a lot less than a piano, and therefore also has a lot less inertial mass making it much easier to move.

Galileo Galilei (1564-1642) was the first to articulate a scientific principle of inertia, in which he states that "a body in motion will remain in motion, and a body at rest will remain at rest unless disturbed." Several decades later Sir Isaac Newton (1642-1727) refined this principle into his first law of motion, stating that

> *An object in motion will have a velocity of constant magnitude and direction unless acted upon by a net force.*

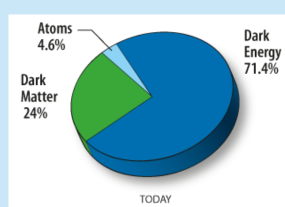
The quantity of inertia to which we refer remains a bit of a mystery; there is as yet no comprehensive understanding of the *origin* of inertia, nor of its substantive difference with gravitational mass. There is no *a priori* reason why inertial and gravitational mass should be the same for any body. Newton developed two very comprehensive theories based on each (kinematics and gravitation), and later Albert Einstein would base his very successful General Theory of Relativity on the assumption that the two are exactly equivalent to one another. Nevertheless, for the remainder of this course, it is solely the *inertial* aspects of any instrument or medium that will concern us for the physics of music.

On a related note . . .

Building Blocks of Matter

Physicists have determined that matter as we know it constitutes a small fraction of the "stuff" in our universe. In addition to regular *atomic matter* (the elements in the periodic table) which constitutes about 4.6% of the known universe, *dark matter* constitutes approximately 24% of the universe, and the remaining 71.4% is taken up by *dark energy.* We do not yet understand the nature of dark matter and dark energy (comparatively recent discoveries), and so they currently remain a mystery. Understanding these two new forms of matter and energy form a very active area of current research. The normal matter that makes up the solar system, including the sun, planets, trees, bees, knees, *etc.* consists of elementary particles called quarks and leptons, whose behaviors are governed by only 4 forces in nature.

Given the incredible richness and variation of form found across the face of the earth and throughout the universe, it is really quite astounding to learn that it all derives from, according to our modern understanding, a total of 12 elementary particles and 4 forces!



The MKS unit for mass is the kilogram (kg). If we are given the mass of an object in any other unit (for example in grams (gm)) then we can easily convert its mass to kg.

3.2.3 Speed

The speed of an object is measured in how much distance it covers in a particular amount of time. The MKS unit for speed is meters/second (m/s). An understanding of speed will be important for our study of the propagation of sound waves in air. The definition of speed v of an object moving in a particular direction is

$$\boxed{v = \frac{\Delta x}{\Delta t}} \tag{3.1}$$

where Δx is the distance through which the object has moved, and Δt is the time it took to move that far.

Example 3.1

Car on the Highway *A car moving along the highway covers a distance of 88 feet in one second. Compute its speed in miles per hour (mph) and in meters per second (m/s).*

Solution: First we need to convert feet to miles and seconds to hours. We are given the speed of 88 ft/s. Using the conversion technique we learned in chapter 2 and the conversion factors given in Appendix A,

$$v = \frac{60.0\ \text{mi}}{\text{hr}} = \frac{60.0\ \cancel{\text{mi}}}{\cancel{\text{hr}}} \times \left[\frac{5280\ \text{ft}}{1\ \cancel{\text{mi}}}\right] \times \left[\frac{1\ \cancel{\text{hr}}}{3600\ \text{s}}\right] = 88.0\ \text{ft/s}.$$

Notice that each of the fractions in the square brackets equals 1 by design so that multiplying the original quantity by them does not change its value but just its unit representation.

Next we convert to MKS units of m/s:

$$\frac{60.0\ \text{mi}}{\text{hr}} = \frac{60.0\ \cancel{\text{mi}}}{\cancel{\text{hr}}} \times \left[\frac{1610\ \text{m}}{1\ \cancel{\text{mi}}}\right] \times \left[\frac{1\ \cancel{\text{hr}}}{3600\ \text{s}}\right] = 26.8\ \text{m/s}$$

3.2.4 Force

In the simplest and most straightforward terms, a force is a *push* or a *pull*, and always involves two bodies interacting with one another. We say that one body *exerts a force* on another, and it doesn't matter whether one (or both) of them actually moves. A force that involves direct contact ("bumping") between two bodies is called a *contact* force. There are other kinds of forces, including *long-range* forces such as gravity or the interaction between two molecules of air "bouncing" off of one another (they don't actually touch one another, but rather interact via their mutual repulsion, which operates in the volume of space between the two molecules). When the force is exerted on or by a stretched string, it is called a *tension* force. The force of *friction* acts between two bodies when they slide against one another. All of these examples (with the exception of gravity) involve the electromagnetic force, and will be of particular interest to us in the study of musical instruments.

The MKS unit of force is the *Newton*. One Newton is defined as the force necessary to cause a body with a mass of 1 kg to accelerate at 1 m/s^2 – that is, to cause the body to have a speed that increases by an additional 1 m/s every second.

Example 3.2

Holding a Guitar Case *Your guitar and case together weigh 13 lbs, which is the force exerted by the earth on their combined gravitational mass. To hold it up by the handle requires that you exert exactly the same amount of force upward, in order to balance the force of gravity downward. What is the weight of the guitar and case in newtons?*

Solution: Using the conversion factors listed in appendix A, we see that

$$13 \text{ lb} = 13 \cancel{\text{lb}} \times \left[\frac{4.448 \text{ N}}{1 \cancel{\text{lb}}}\right] = 57.8 \text{ N}$$

On a related note . . .

Forces in the Universe

There are only four basic forces that physicists have identified in the universe. The four forces are called:

Strong ***Weak*** ***Electromagnetic*** ***Gravity***

Only the latter two of these are ones that we personally experience in everyday life. The strong and weak forces operate at the sub-atomic level and are responsible for holding the particles together that make up the nucleus of the atom. Therefore they are clearly important to our well being, but we do not directly experience their effects as we do the electromagnetic force and gravity. Gravity is by far the weakest of all the forces (being 30 orders of magnitude weaker than the weak force), but we experience it as if it were very strong, owing to the vast size of the earth, and the fact that every single particle of the earth exerts a force on every single particle of our bodies. The electromagnetic force is responsible for all other forces we experience, including that of the ground on our feet that prevents us from falling to the center of the earth, the friction when rubbing a table top, the elastic force of a stretched spring or rubber band, the collision force when we bump into a wall, and many others. It may not sound intuitive that these forces originate from the electromagnetic force, but on the atomic level, it is the microscopic electrostatic force between countless individual atoms that produces these larger macroscopic forces.

In our study of musical instruments, forces producing the initial vibrations which give rise to subsequent waves that propagate as sound are all *elastic* in nature, and derive from the electromagnetic force.

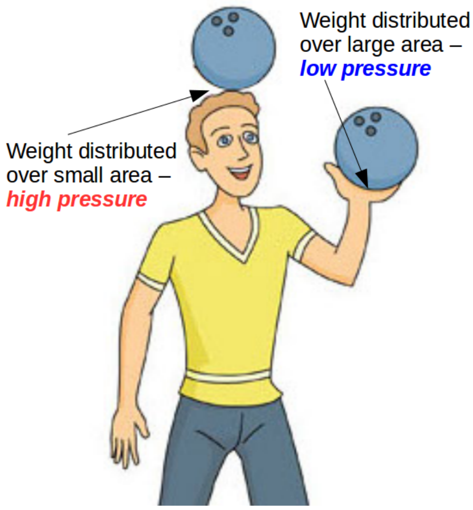


Figure 3.1: A bowling ball being supported in two ways. The weight of the ball distributed over the area of support determines the pressure. The hand supports the ball using much more area than the skull. Since pressure is inversely proportional to area, the pressure on the skull is therefore much higher than on the hand for a ball of the same weight. Adapted from [2].

3.2.5 Pressure

When force is applied to a body, it is necessarily applied over a certain square area. For example, when I push on the backside of my stalled car in order to move it out of the intersection, my two hands cover a certain square area on the body of the car, say about 300 cm^2, or about 0.03 m^2 total. *Pressure* is defined as force per unit area,

$$\boxed{P = \frac{F}{A}}. \tag{3.2}$$

If I push with a total force of 110 lb (= 490 N), the pressure I exert, given the square area of my hands, is 16,300 N/m^2. That sounds like a huge pressure, but keep in mind that I am not applying the pressure over an entire square meter but only a fraction of that area. If I choose to apply the same force, 110 lb, but use my knuckles instead of my palms, the area of contact is significantly reduced, and therefore the pressure significantly increases. This becomes very evident as I experience pain in my knuckles. Even though the force on the car is the same, the pressure is much higher since the area is much smaller (look once more at equation 3.2). This is why, incidentally, it is much more comfortable to sleep with your head on a pillow than on a rock. The weight of your head (*i.e.* the force of gravity on your head) is the same in both cases, but the weight is distributed over a much larger area on the pillow than on the rock. See figure 3.1 for a similar situation where the area over which a force is applied can make a big difference.

Atmospheric pressure is the measure of how much pressure the air exerts on all surfaces with which it comes into contact at the earth's surface. At sea level this pressure is about 14.7 lb/in^2 = 1.013×10^5 N/m^2. The MKS unit of pressure is the Pascal, 1 Pa = 1 N/m^2. The atmospheric pressure at sea-level, when expressed in Pa, is

$$\boxed{\text{Sea level: } P_{atm} = 1.013 \times 10^5 \text{ Pa}} \tag{3.3}$$

A good way to understand where atmospheric pressure comes from is to imagine the following. Draw out a square on the ground outside with area 1 ft^2. Imagine the square column of air resting on that square foot and rising to the top of the atmosphere. The total weight of this column, as it turns out, is about 2117 lb, or a little over a ton! The pressure on the ground applied by the weight of this column divided by the area which comes out to 2117 lb/ft^2 = 14.7 lb/in^2. This means that if instead you drew out a 1 in^2 area on the ground, the weight of the column of air on this smaller footprint extending to the top of the atmosphere would be 14.7 lb. You might wonder why it is that we don't feel this pressure on our skin. A pressure of 14.7 lb/in^2 is approximately what we would feel if we balanced a bowling ball on the top of our head. Multiply this pressure by the number of square inches for a typical human body and you've got a significant force!

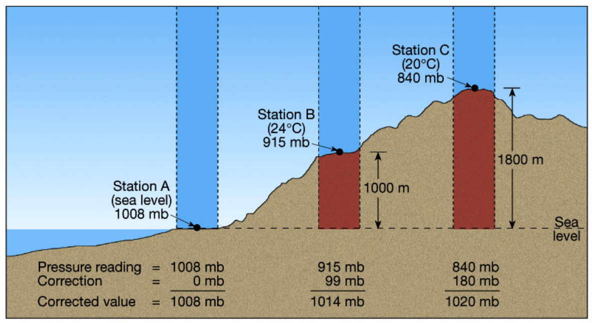


Figure 3.2: An illustration of how air pressure differs at different altitudes. At higher altitudes the column of air supported by the ground has less mass since it is a shorter column [1].

Example 3.3

Atmospheric Pressure on the Skin *The surface area of skin for a typical adult is around 1.75 m^2. What is the approximate total force exerted by the air over the entire body? Start with atmospheric pressure expressed as 14.7 lb/in^2.*

Solution: First we need to compute atmospheric pressure in MKS units:

$$14.7 \text{ lb/in}^2 = 14.7 \cancel{\text{lb}}/\cancel{\text{in}}^2 \times \left[\frac{4.48 \text{ N}}{1 \cancel{\text{lb}}}\right] \times \left[\frac{12 \cancel{\text{in}}}{1 \cancel{\text{ft}}}\right]^2 \times \left[\frac{1 \cancel{\text{ft}}}{0.3048 \text{ m}}\right]^2 = 1.013 \times 10^5 \text{ N/m}^2$$

To calculate the total force on the body we need to multiply this pressure by the body's square area,

$$F_{total} = P \cdot A = 1.013 \times 10^5 \text{ N/m}^2 \cdot 1.75 \text{ m}^2 = 177,300 \text{ N} = 39,860 \text{ lb}.$$

This is an enormous force on a body by any measure, certainly sufficient to crush it. So why is it that we don't experience this force? The answer is that the same pressure exists inside the body, and in the interior of all tissues and bones, so that all pressures completely balance and we don't feel any force. If we were suddenly to step into the vacuum of outer space, the high pressure in our bodies would cause us to explode into the low pressure of vacuum due to the *imbalance* of pressure. Not recommended.

Example 3.4

Standing and Tiptoe *Imagine you weigh 150 lb, and the area of both your feet together when standing barefoot is about 150 cm^2. What is the pressure of your feet on the floor as you stand? How about when you stand on your tip-toes, for which the area is more like 6.0 cm^2? Which is the more painful scenario and why?*

Solution: First we need to convert the force and areas into MKS units

$$150 \text{ lb} = 150 \cancel{\text{lb}} \times \left[\frac{4.448 \text{ N}}{\cancel{\text{lb}}}\right] = 667 \text{ N}$$

$$A_{standing} = 150 \cancel{\text{cm}}^2 \times \left[\frac{\text{m}}{100 \cancel{\text{cm}}}\right]^2 = 0.015 \text{ m}^2$$

$$A_{tiptoe} = 6.0 \text{ cm}^2 \times \left[\frac{\text{m}}{100 \text{ cm}}\right]^2 = 6.0 \times 10^{-4} \text{ m}^2$$

The pressure on the soles of your feet when standing is the force of your weight, divided by the contact area of your feet on the floor:

$$P_{standing} = \frac{F}{A_{standing}} = \frac{667 \text{ N}}{0.015 \text{ m}^2} = 4.47 \times 10^4 \text{Pa}$$

When tiptoeing, the pressure on the tips of your toes is much larger,

$$P_{tiptoe} == \frac{F}{A_{tiptoe}} = \frac{667 \text{ N}}{6 \times 10^{-4} \text{ m}^2} = 1.11 \times 10^6 \text{Pa}.$$

This pressure is about 25 times the pressure of standing, which is why it is so much more painful to stand tiptoe.

A look ahead . . .

Sound as a Pressure Wave

As we will see in chapter 4, sound consists of "pressure waves" that travel over long distances through the air. Small regions of compression (higher than atmospheric pressure) interspersed by regions of rarefaction (lower than atmospheric pressure) move through the air together. The molecules of air don't themselves move very far at all, just back and forth over very short distance, while the high and low pressure zones move long distances through the air together, constituting the sound wave.

3.2.6 Work

Work is a special form of energy. When a force is applied to an object and that object moves along the same direction as the force, a net amount of work is said to be done on the object. The work is done by whatever supplies the force, and the work is positive if the object's motion is in the same direction as the force, and negative if it moves in the opposite direction. Whether or not the force actually *causes* the movement is not important. The amount of work is defined as the product of the force and the distance moved,

$$\boxed{W = F \cdot d} \tag{3.4}$$

where F is the applied force, and d is the displacement *along the direction of the force.* The MKS unit for work is a N·m, referred to as a Joule. If an object does not move, no work is done, regardless of any forces acting on it. I can push against the wall for as long as I like, and get very tired in the process, and yet in the strict physics sense of the term, I do no work on the wall since it does not move. I certainly become tired after a while, owing to the continual exertion of my muscles, but I have done no work *on the wall.*

Note that if the object moves in a direction *perpendicular* to the applied force, no work is done either. The only motion that counts is motion along the same direction as the force. I can do work on my textbook by simply lifting it up or setting it down. When I lift it, I do positive work on the book since the force I apply is upward, as is the movement of the book. If I lower the book, I do negative work on it since my force (up) and the motion (down) are in opposite directions. Work is similarly done by a musician on a musical instrument. When the guitarist pulls on a string in order to pluck it, she applies a force in the same direction as the string moves, resulting in positive work. This positive work (remember it is a form of energy) then becomes manifest in the production of sound, which takes energy to produce.

Example 3.5

Moving a Stalled Car *Your car has stalled ... in the middle of an intersection. No one is available to help you push the car. The ground is level, fortunately, so you decide to push it to the nearest service station, which is 0.8 mi away (!). After aligning the steering wheel in the direction you need to move the car and placing the transmission in neutral, you commence pushing on the back end of the car. On average, you exert a force of 110 lbs, and the total distance you move the car is 0.8 mi. Calculate the total work you've done (in Joules) on the car by the time you get it to the service station. Calculate your answer also in terms of food calories.*

Solution: The first thing we need to do is to convert all quantities to their MKS units. The total distance moved by the car should be in meters:

$$d = 0.8\ \cancel{\text{mi}} \cdot \left[\frac{1610\ \text{m}}{\cancel{\text{mi}}}\right] = 1288\ \text{m}$$

Next we need to calculate the force exerted on the car in Newtons.

$$F_{avg} = 110 \cancel{\text{lb}} \cdot \left[\frac{4.448 \text{ N}}{\cancel{\text{lb}}}\right] = 489 \text{ N}$$

Finally, in order to calculate the work that you've done, we need to use equation 3.4,

$$W = F \cdot d = 489 \text{ N} \cdot 1288 \text{ m} = 630,200 \text{ J}.$$

In terms of food calories, this amount of work comes out to

$$630,000 \cancel{\text{J}} \cdot \left[\frac{\text{Ca}}{4186 \cancel{\text{J}}}\right] = 150 \text{ Ca}.$$

It doesn't sound like much when expressed in terms of food calories, especially for such a tiresome process!

3.2.7 Energy

Energy comes in many forms, and can be transformed from one type to another. Though it is a mystery exactly what energy *is*, the form of energy that is simplest and perhaps easiest to understand is associated with *motion*. In order to get a body moving, one has to exert a force on it to overcome its inertial tendency to remain at rest. In other words, one has to do *work* on the body to get it moving. When in motion, it has a positive amount of energy "stored" in this motion, called *kinetic* energy. The amount of kinetic energy will be exactly equal to the amount of work done on the object to get it moving in the first place. The kinetic energy K for a moving body is defined as

$$\boxed{K = \frac{1}{2}mv^2} \tag{3.5}$$

where m is the body's mass and v is its speed. When m and v are expressed in their MKS units, the kinetic energy is expressed in Joules.

Other common types of energy include various forms of *potential*, or *stored* energy, including chemical energy, nuclear energy, gravitational energy, and spring energy. While we see many forms of energy in nature, we don't understand what it actually is in *essence*. When we burn gasoline in our car, the stored chemical energy is released from the gasoline in the burning process, and transformed into the car's kinetic energy of motion. It also contributes to heating the engine, making noise, and emitting warm exhaust into the environment – all forms of energy. When a book sits on a shelf, it has a particular form of energy called *gravitational* potential energy. When I push a book off of a shelf, this gravitational potential energy is quickly converted into the kinetic energy as the book falls to the ground. The work that gravity does is positive since the force it exerts is in the same direction as the motion. This work done by gravity is directly transformed to kinetic energy.

Another common unit of energy is the *calorie* (note that it is spelled with a small "c"), which is defined as the amount of energy that will raise 1 cm^3 of water 1 degree Celsius. In terms of Joules, 1 cal = 4.186 J. One food Calorie (designated with a capital C) is defined as 1000 calories, or 1 kcal. Thus 1 Cal = 4186 J. The energy stored in food is expended when the body metabolizes, or burns it.

Example 3.6

Kinetic Energy of a Car *A car moves along the road at a speed of 35 mi/hr. If the mass of the car with passengers is 3250 lbs, compute its kinetic energy.*

Solution: Let's first convert the car's mass and speed into their MKS unit representations.

$$3250 \text{ lb} = 3250 \cancel{\text{lb}} \times \left[\frac{4.448 \text{ N}}{1 \cancel{\text{lb}}}\right] = 14,400 \text{ kg}$$

$$35 \text{ mi/hr} = 35 \cancel{\text{mi}}/\cancel{\text{hr}} \times \left[\frac{1610 \text{ m}}{1 \cancel{\text{mi}}}\right] \times \left[\frac{1 \cancel{\text{h}}}{3600 \text{ s}}\right] = 15.65 \text{ m/s}$$

We can then compute the kinetic energy using equation 3.5:

$$K_{car} = \frac{1}{2}mv^2 = \frac{1}{2}(14,400 \text{ kg}) \cdot (15.65 \text{ m/s})^2 = 3.53 \times 10^6 \text{ J}$$

or 3.53 MJ (1 MJ = 1 "mega-Joule" = 1×10^6 J).

3.2.8 Sound Energy

In order to produce sound waves, which require energy, work needs to be done to start the process of vibration. When I speak, the chemical energy stored in my body (deriving from my breakfast, for example) is expended in metabolic processes, allowing me to exercise my diaphragm, which forces the air from my lungs into my trachea and through the vocal folds, which in turn vibrate and do work on the molecules of air in my throat and vocal cavity. These air vibrations propagate out of my mouth in the form of a traveling wave into the lecture hall. Eventually these traveling vibration waves do work on your eardrums, enabling you (hopefully!) to comprehend the words I speak.

3.2.9 Conservation of Energy

One of the most highly cherished of all conservation laws in physics is the conservation of energy. Energy is neither created out of nothing, nor does it disappear into nothing. Rather, it is always transformed from one type into another. It originally derives from some source of stored energy, and then is transformed into other types of energy as processes occur. No process has been seen to violate this conservation principle, which says that before and after a process occurs, the total energy of the system is conserved – *i.e.* the total energy before equals the total energy after even though the energy may change from one type to another in the process.

On a related note . . .

Solar Energy

We talk a lot about the need for developing solar energy as a major source of energy for our world's growing need, and rightly so. It is less appreciated, however, just how much we already rely on solar energy. Besides nuclear and geothermal energy, virtually *all* sources of energy available on earth derive in some way from our sun. When we burn fossil fuels, we are releasing the energy stored there long ago by the sun, from the process of photosynthesis (without which plant life would not grow on earth). When hydro-electric plants produce power, they are simply tapping, storing, and making available energy from water set into motion by the atmospheric water cycle, which is itself driven by the sun's power. When we harvest power from wind, we are tapping energy from air set into motion by atmospheric heating from the sun. When we start each day with breakfast, we derive energy stored in the food, which comes from ... you guessed it ... the sun.

3.2.10 Power

Power is a measure of the *rate* at which energy is expended - that is, energy per unit time,

$$\boxed{P = \frac{\Delta E}{\Delta t}} \tag{3.6}$$

where ΔE is the energy expended over time Δt. When energy and time are expressed in their MKS units of Joules and seconds, the power comes out in the MKS unit of Watts. When a fixed quantity of energy ΔE is expended in a process, the speed at which it is transferred determines the power – the shorter the time, the higher the power output, and the longer the time, the lower the power output.

3.3 Vibrations

Recall the three requirements listed at the beginning of this chapter for musical sound to be possible – *production*, *propagation*, and *perception*. Now that we've considered some of the *physical* principles necessary for understanding sound, let's begin by discussing the first of these three requirements. Sound requires a source, and that source requires motion, in the form of *vibrations*. Sound does not result from stillness. When we speak or sing, air is forced through our vocal folds, which oscillate under this force with high frequency (chapter 13). We hear music from ear buds because miniature diaphragms inside them vibrate in response to electrical signals sent through the wire from the iPod (chapter 16). Air blown over the mouthpiece of a flute sets the air in the cylindrical column into vibration (chapter 10). In short, *all* sound derives from vibrational motion. Sound can arise from a wide variety of sources, such as the following:

Inanimate sources: Thunder, a babbling brook, wind in the trees, rain, ocean waves . . .

Living sources: Birds, howling wolves, whales, human conversation, applause . . .

Manufactured sources: Traffic noises, helicopter blades, jackhammers . . .

And then there's *music* ... a human art form that utilizes sounds from each of these three categories.

3.3.1 Simple Harmonic Motion

Recall from chapter 1 that the foundation of musical sound is *periodic motion*. The simplest example of periodic motion is from the *simple harmonic oscillator*. The signature of a simple harmonic oscillator is *sinusoidal* motion, as depicted in figure 3.3.

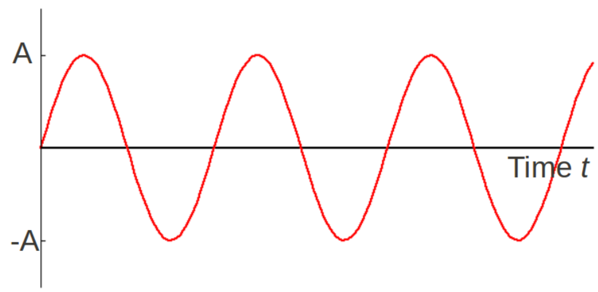


Figure 3.3: Graphical representation of sinusoidal motion from a simple harmonic oscillator.

The vast majority of sounds that we perceive are *not* simple harmonic in nature, but have waveforms that are more complex than sinusoidal. Several examples of non-sinusoidal waveforms are shown in figure 3.4. We will be considering more complex waveforms in much greater detail later.

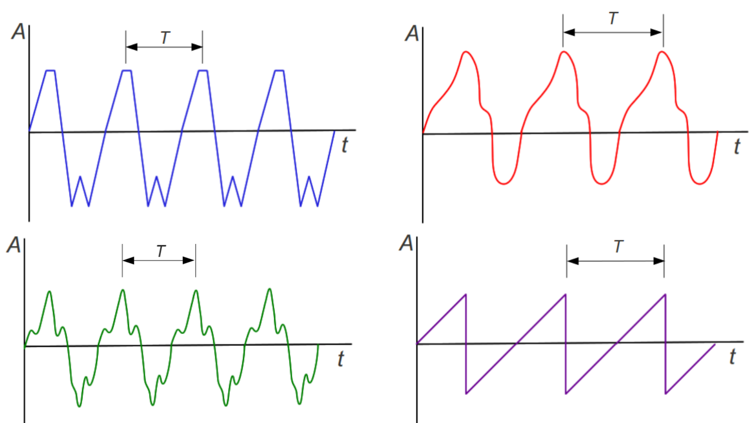


Figure 3.4: Four different periodic complex (i.e. non-sinusoidal) waveforms. Note that while each is clearly periodic in nature, they have complex structures compared with a sinusoidal oscillation with period T.

3.3.2 Mass on a Spring

The physicist's archetypal example of a simple harmonic oscillator is a mass attached to a spring. When the mass is at rest, and therefore still, it resides at its *equilibrium* position. At this location no net force acts on the mass. When we grasp the mass and pull it away from this resting point, the spring stretches in response to our pull and applies an equal and opposite force to the mass. Note that the spring *always directs its force back toward the equilibrium point*, whether we pull or push on the mass (see figure 3.5).

If we denote a displacement of the mass from equilibrium by the variable x, then the force applied by the spring is always directed *opposite* to the direction of this displacement. Thus if the mass is moved in the $+x$ direction (*i.e.* right), the spring force points in the $-x$ direction (left); if the mass is moved in the $-x$ direction (left), the spring force points in the $+x$ direction (right). Since the force applied by the spring is always directed opposite to the displacement, it always has a negative value relative to the position x:

$$\boxed{F_s = -kx} \tag{3.7}$$

where F_s is the force exerted by the spring on the mass, k is the *spring constant* that characterizes the strength of the spring, and x is the distance the mass is displaced from its equilibrium point. When k is a constant, F_s is therefore *directly proportional* to the displacement x. This direct proportionality between the restoring force and the displacement defines a *linear* restoring force. The quality of being linear means that if we double the distance the mass is moved, the restoring force exerted by the spring is doubled; if tripled, then the force is tripled; if halved, the restoring force is halved. This condition is true, by the way, as long as we stay within the *elastic region* of the spring. If we pull the mass *too far* from equilibrium, the spring becomes very stiff and unyielding, and the force is no longer linearly related to position – it rises much more quickly with increased distance. You can demonstrate this effect by pulling on a rubber band so far that you exceed its elastic limit, at which point you will notice that it becomes quite unyielding and eventually breaks under sufficient force.

Other systems that undergo simple harmonic motion include the motion of a marble in a bowl (see figure 3.5) and a child on a swing. In both cases, displacing the marble from the bottom of the bowl or the child from the lowest point on the swing causes gravity to exert a force directed back toward the equilibrium point.

In the case of a child on a swing, which is a type of *pendulum*, the system consists of her weight hanging on the end of a rope. When we pull her to one side and release, the gravitational force on her body accelerates her back toward her equilibrium point at the bottom. When she arrives at that point she is moving at her fastest, and overshoots the equilibrium point, owing to her *inertial mass*. She continues to move past the equilibrium point until she slows to a stop at her maximum displacement on the other side, after which her motion reverses and she returns to the equilibrium point, overshooting it again, and so on. In this way, her periodic motion (*i.e.* her oscillating back and forth about the bottom equilibrium point) is initially produced by first displacing her from equilibrium and then letting her go to oscillate on her own. Note that her actual speed varies throughout her cycle of motion, a characteristic of many oscillators. Also, her kinetic energy of motion and potential energy of displacement in height are in a continual exchange with one another. Her potential energy is at a maximum at the top of her swing (where kinetic energy is zero), and her kinetic energy is at a maximum as she passes through the bottom point (where potential energy is zero).

A look ahead . . .

Elasticity of the Ear Drum

We will see in chapter 8 that the ear drum, an elastic system, can be driven beyond its elastic limit by very loud sounds, in which case distortion of the perceived sound results.

Example 3.7

Mass on a Spring *A 0.5 kg mass is connected to a spring with spring constant k = 150 N/m. The spring is stretched so that the mass is 3.0 cm from equilibrium. a) Calculate the force the spring exerts on the mass. b) Why does the result not depend on the value of the mass?*

Solution: To start with, what do we know? We know the value of the spring constant k, the mass m, and the distance x the mass is pulled from equilibrium. What we don't know is the force F_s.

a) This is a "one-step" problem. The equation we need to solve for F_s (equation 3.7) has three variables, two which we know and one that is unknown. But before we calculate the force, we need to make sure that all our quantities are expressed in their MKS units so that the force will come out in Newtons. First we convert the distance x to meters:

$$x = 3.0\ \cancel{\text{cm}} \cdot \left[\frac{\text{m}}{100\ \cancel{\text{cm}}}\right] = 0.03\ \text{m}$$

Since k is already expressed in newtons/meter, we are ready to calculate the force:

$$F = -kx = -\left(15\ \text{N/m}\right) \cdot \left(0.03\ \text{m}\right) = -0.45\ \text{N}.$$

The *value* of the force is 0.45 N, and the *direction* is opposite to the displacement x as denoted by the negative sign.

b) The value of the force is independent of the mass attached to the spring (note that the variable m does not appear in the equation), since the spring force based only on how *far it is stretched*. The mass attached to its end does not affect the force exerted by the spring, but if the mass is released, the subsequent motion of the mass *will* depend on its value, as we will see in equation 3.11.

3.3.3 Elasticity and Inertia

For a simple harmonic oscillator to work, it must have two features in order that it can exhibit periodic behavior. It must have a quantity of "springiness" or *elasticity*, as well as a quantity of "massiness" or *inertia*. In the system consisting of a mass on a spring, the element of elasticity comes from the spring and the element of inertia comes from the attached mass. The function of these two quantities and how they work together can help us to understand how the system oscillates. When the mass is pulled away from its equilibrium point and held, the spring pulls back on the mass with an equal and opposite force (owing to its *elasticity* - see for example figure 3.5). When the mass is released, the spring force causes it to accelerate back toward its resting point until it reaches it with considerable speed. Since it is moving, it overshoots that point owing to its *inertia*. It then slows down to a rest opposite its starting point by the spring force, only to turn around and head back toward equilibrium. Again it overshoots this point, and the cycle continues as oscillation about the equilibrium point. This back and forth oscillation continues because of the interplay between the elastic and inertial components of the system. The oscillation will continue indefinitely until someone or something stops it.

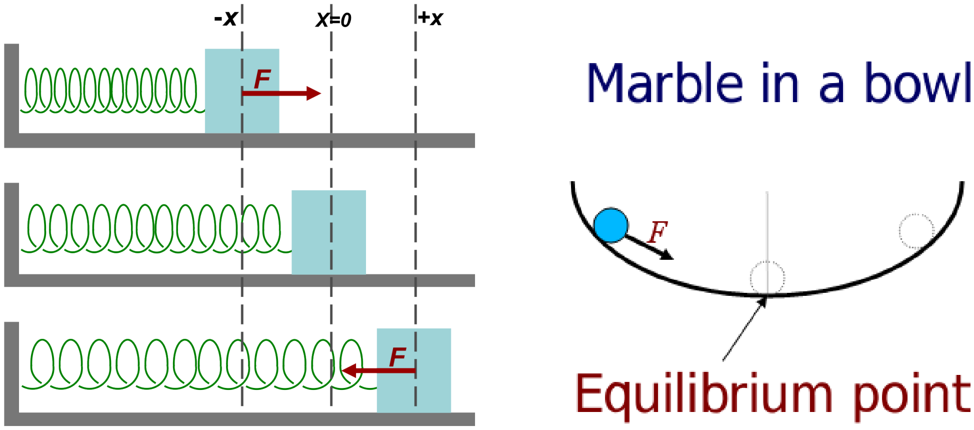


Figure 3.5: (***For color version see Appendix E***) Left: A mass attached to a spring. In the middle figure the mass is at rest at the equilibrium position. Note that the force F exerted by the spring is always directed toward the equilibrium point. At the top the mass is pushed to the left (compressing the spring), to which the spring responds with a force to the right (toward the equilibrium point). At the bottom, the mass is pulled to the right (stretching the spring), to which the spring responds with a force to the left (again toward the equilibrium point). Right: A marble in a bowl. At the center of the bowl the marble is in equilibrium - no net force acts on it. Here the restoring force is provided by gravity. As soon as the marble is displaced from the center, gravity applies a force back toward the center. When let go, the resulting motion of the marble is simple harmonic motion.

3.3.4 Energy Considerations

If we pull on the mass and stretch the spring, we perform *work* on the system, since the force we apply moves the mass over a specific distance (remind yourself of the definition of work outlined in section 3.2.6). When we stop pulling, the mass is then displaced by a fixed distance from center. At that point the amount of work we have done has become *stored in the spring* as potential energy, U_s, which is given by

$$\boxed{U_s = \frac{1}{2}kx^2} \tag{3.8}$$

When we then release the mass, the stored spring energy is transferred to kinetic energy, K, of the moving mass. After the mass overshoots the equilibrium position and becomes displaced by the same distance on the opposite side of equilibrium, it stops briefly and the energy is once again stored as potential energy in the spring. The original work energy we did by pulling the mass away from its equilibrium point and then letting go is has become *stored in the oscillating system*, and the energy oscillates between potential energy and kinetic energy.

The sum of the potential and kinetic energy *always equals a constant value*, owing to the conservation of energy:

$$\boxed{E_{total} = U_s + K}. \tag{3.9}$$

This sum corresponds precisely to the amount of work we put into the system by stretching the mass to its starting point. After release, the mass oscillates, and the energy of the system is transformed back and forth between stored spring energy and energy of the mass's motion. At the precise moments when the mass passes through its equilibrium point, all of the system's energy is stored in its motion (kinetic energy K), and at the precise moments when the mass comes briefly to rest at its farthest positions from equilibrium, all of the system's energy is stored in the stretched or compressed spring (potential energy U_s). At all other times, the system's energy is shared between the potential and kinetic energy.

There are two different oscillations going on in this system, one easy to visualize and the other more abstract. While the mass physically oscillates in *space*, the *energy* of the system also oscillates potential energy stored in the spring and kinetic energy of the mass's motion. Each of these oscillations has the same periodicity.

3.3.5 Period and Frequency

In order to properly characterize the periodicity of oscillating systems, we need to understand the concepts of *period* and *frequency*. Vibrations that are important for music are *periodic* in nature. A periodic vibration has a well-identified *cycle* that repeats itself with regularity. Many features in nature exhibit *periodicity*: the earth's rotation has a regular periodicity that brings day and night; a child on a swing moves from one extreme to another with regular periodicity; the high and low tides of the ocean each appear twice daily, owing to the gravitational effects of the moon and the sun.

The *period* of the vibration T is defined as the time corresponding to the execution of one full cycle. The *frequency* f of the vibration is defined as the number of cycles executed per second, which is the inverse of the period. The period T has units of *seconds per cycle* and frequency f has units of *cycles per second*. The two are the inverse of one another:

$$\boxed{T = \frac{1}{f} \qquad\qquad f = \frac{1}{T}}. \tag{3.10}$$

A unit we will be using for frequency is the Hertz, named after the 19^{th} century physicist Heinrich Hertz, and abbreviated "Hz". The Hz is defined as 1 *cycle/second*.

Example 3.8

Calculating Frequency *A typical child on a swing has a period of around 2 seconds, the period of a person clapping has a period of 0.25 seconds, and the period of a vibrating string producing a "concert-A" pitch is 0.002273 seconds. Calculate the frequencies for each of these examples.*

Solution: All of these examples can be calculated using equation 3.10. The frequency of the child on the swing is

$$f = \frac{1}{T} = \frac{1}{2\ \mathrm{s}} = 0.5\ \mathrm{Hz}.$$

Clapping with a period of 0.25 s results in a frequency of

$$f = \frac{1}{0.25 \text{ s}} = 4.0 \text{ Hz}.$$

The frequency of the concert-A pitch is

$$f = \frac{1}{0.002273 \text{ s}} = 440. \text{ Hz}.$$

Example 3.9

Headache ... *A woodpecker taps at the trunk of a tree, making 8 pecking sounds every second. What is the time corresponding to one full cycle of her peck?*

Solution: One full cycle of the woodpecker's head motion corresponds to one audible tap on the tree trunk. She makes 8 taps per second, and therefore the frequency of her head motion is 8 taps/second, or 8 Hz. Since the period T is inversely proportional to the frequency f, then

$$T = \frac{1}{f} = \frac{1}{8 \text{ Hz}} = 0.125 \text{ sec}$$

Thus the time it takes her head to move through one cycle is just over 12 hundredths of a second. Just try tapping your head against the wall that fast – I dare you!

3.4 Resonance

Consider the mass hanging on a spring as depicted in figure 3.6. When the mass is lifted up above its equilibrium point and released, it moves down and then up and then down again, continuing in sinusoidal motion. The system oscillates at a frequency that is "natural" to it, the value of which is determined by the physical properties of the system. This brings us to the topic of *resonance*, which will become very important in our study of musical instruments. The specific nature of sound emitted by real, physical instruments is dependent on their resonant structure, which is typically quite complicated. So before we study musical instruments in detail let's consider the topic of resonance by looking at much simpler systems.

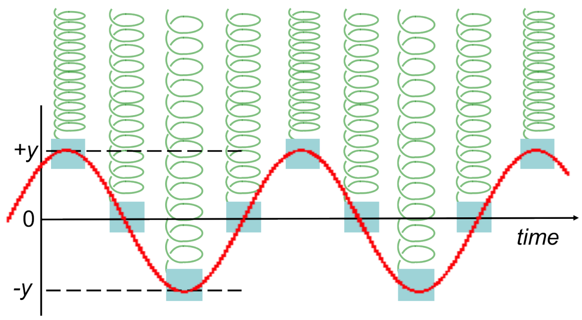


Figure 3.6: Demonstration of the periodic nature of simple harmonic motion. The oscillating mass on the spring executes sinusoidal motion as it oscillates back and forth about its equilibrium point. The frequency at which it executes this motion is natural to the system, dependent only on its physical characteristics.

3.4.1 Natural Frequency of Oscillation

Oscillating systems typically have a frequency or a collection of frequencies at which they naturally vibrate. A system at rest and left to itself will not oscillate - it needs an external source of energy to get it going. But pluck a guitar string and it quickly adopts the frequency of oscillation that is natural to it, giving off a recognizable pitch. Pull a mass attached to a spring away from its equilibrium point and let go, and the system will vibrate at its natural frequency of oscillation. Note that when we let go of the mass, we are not *causing* it to oscillate at this particular frequency, nor when we pluck the guitar string are we *forcing* it to adopt any particular frequency. But both of these systems automatically and quickly migrate toward oscillation at their natural frequencies. When I pull the child back on the swing and let go, the system quickly adopts the frequency that is natural to it. I can choose to force the swing to oscillate at a frequency that is *not* natural to it (say, by pushing on the swing once every half second), but you will quickly find that it won't build up much oscillation amplitude in response (nor will the child be very happy about it!). This is because I am forcing the system to oscillate at a frequency far away from its natural frequency, and its response is accordingly weak. On the other hand, when I drive the system at its natural, or *resonant* frequency, large oscillation amplitudes occur.

The resonant frequencies of a system depend on two important factors discussed in section 3.3.3, namely the stiffness (elasticity) and the mass (inertia) of the system. For the case of a mass on a spring system, increasing the spring constant (i.e. making it stiffer) *increases* the natural frequency of oscillation, and attaching a larger mass to the spring *decreases* the natural frequency. Different strings on a guitar have different natural frequencies and therefore different pitches when plucked, because the strings have different thicknesses (inertia) and are tied at different tensions (elasticity). The process of tuning a guitar involves adjusting the string tension, which changes the string's elasticity, which therefore changes its natural frequency of oscillation. Understanding the nature of resonance will help us build a better understanding of how musical instruments operate and produce the specific sounds they do.

3.4.2 Frequency of Mass on a Spring

So far we've seen that when a mass connected to a spring is pulled away from equilibrium, and released, it oscillates at what is called its natural frequency. This frequency can be calculated from the spring constant and the mass using the following equation:

$$\boxed{f_s = \frac{1}{2\pi}\sqrt{\frac{k}{m}}} \tag{3.11}$$

where f_s is the oscillation frequency of the system, k is the spring constant, and m is the mass. From this equation we can now see why increasing the spring constant (thus making the system more stiff) *increases* the natural frequency, while increasing the mass (giving it greater inertia and therefore more resistance to movement) *decreases* the natural frequency.

Example 3.10

Natural Frequency of a Mass-Spring System *By what factor does the natural frequency for a mass-spring system change when the spring constant is tripled in value? Reduced by a factor of 4? By what factor does it change when the mass is replaced with one 5 times the mass?*

Solution: The natural frequency of the mass-spring system is $f_s = 1/2\pi\sqrt{k/m}$. If we replace k with $3k$, the new frequency is

$$f_s' = \frac{1}{2\pi}\sqrt{\frac{3k}{m}} = \sqrt{3}\left(\frac{1}{2\pi}\sqrt{\frac{k}{m}}\right) = \sqrt{3}f_s.$$

As a result of tripling the spring constant, the natural frequency of the system increases by $\sqrt{3}$. When we reduce the spring constant by a factor of 4, the natural frequency changes to

$$f_s' = \frac{1}{2\pi}\sqrt{\frac{k}{4m}} = \frac{1}{\sqrt{4}}\left(\frac{1}{2\pi}\sqrt{\frac{k}{m}}\right) = \frac{1}{2}f_s.$$

The frequency becomes half the original. Finally, when the mass is replaced with one 5 times the mass, the new frequency becomes

$$f_s' = \frac{1}{2\pi}\sqrt{\frac{k}{5m}} = \frac{1}{\sqrt{5}}\left(\frac{1}{2\pi}\sqrt{\frac{k}{m}}\right) = \frac{1}{\sqrt{5}}f_s.$$

Note that the relationship between the natural frequency and the spring constant and mass is *not linear*. The frequency is proportional not to the spring constant but to its square root. It is also proportional to the *inverse* square root of the mass, meaning that the frequency increases with an increasing spring constant and decreases with an increasing mass.

The natural frequencies of real musical instruments are more difficult to capture in such simple mathematical form. But a basic system, such as the mass on a spring, can serve as a helpful model for understanding more complicated systems such as musical instruments.

3.4.3 Driving a Resonant System

So far we've discussed systems that can be made to resonate at their natural frequencies simply by pulling or pushing on them and simply "letting go." Pull the mass on the spring back and release it, and it quickly adopts its natural frequency; give the child on the swing a brief push and the swing naturally oscillates at its natural frequency, *etc.*

Many of us have experienced situations where sound coming from some local source (it might be a loud music sound system) can cause certain objects to buzz in sympathetic vibration – a portion of the wall, a hanging picture, a window, a plate on the table, a metal cabinet, *etc.* You've probably noticed that these items don't buzz continually, but only from time to time, when certain notes in the music (usually in the bass region) occur. These are examples of *driven resonance*, where one vibrating system can generate vibrations in another. The sound waves filling the air cause *everything* in the room to vibrate with very small amplitude, but we don't notice these vibrations until one particular object is driven into much larger vibration and as a result begins to noticeably buzz. The buzzing occurs when the driving frequency (the note in the music) has frequency at or near the object's natural resonant frequency. The very fact that we can hear music coming from an adjacent room is clear evidence that sound sets the walls of both rooms into vibration, since it is the walls that separate the rooms and transmit sound from one room to the next.

We can drive a system to vibrate at any frequency we choose, provided we have a mechanical means to do so. This involves applying a *periodic* force to an object, that is, a force that varies sinusoidally in time. Depicted in figure 3.7 is a mass on a spring, whose natural frequency is f_0, driven by a periodic force F of frequency f_d. When the frequency of the driving force f_d is well below the system's natural frequency (*i.e.* $f_d \ll f_0$), the resulting amplitude of oscillation is small. Figure 3.8 depicts the typical response for a system like this to the value of the driving frequency. The x-axis corresponds to the frequency at which we drive the system, and the y-axis corresponds to the magnitude of its oscillation amplitude. As the driving force frequency is increased to a value closer to the system's natural frequency ($f_d < f_0$), the response becomes larger, until the driving frequency is very near or at the natural frequency ($f_d \approx f_0$), at which point large amplitude oscillations occur. When the driving frequency increases above the object's natural frequency ($f_d > f_0$), the resulting oscillation amplitude once again decreases until it becomes very small and the driving frequency continues to move well beyond the natural frequency ($f_d \gg f_0$).

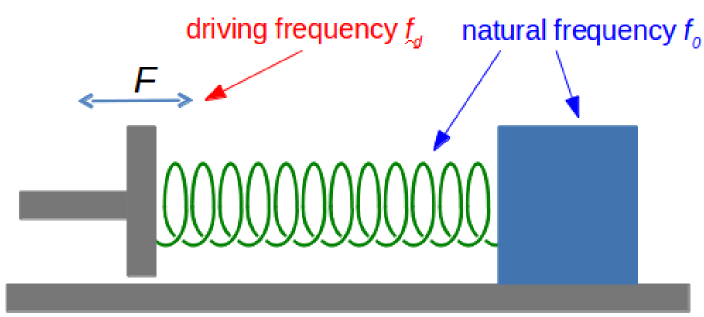


Figure 3.7: (***For color version see Appendix E***) Mass on a spring driven by a periodic force F. When the piston moves back and forth, the connecting spring causes the mass to oscillate as well. The amplitude of the mass will depend on how the driving force frequency f_d compares with the natural oscillation frequency of the mass-spring system f_0.

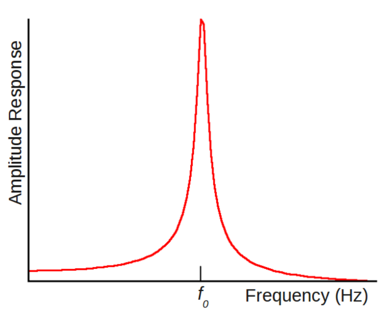


Figure 3.8: The typical amplitude response of a resonant system being driven at different frequencies. When the driving frequency is far from the system's natural frequency (far to the left or right of f_0), the resulting amplitude of oscillation is weak. When the driving frequency is at or near the resonant frequency f_0 the response of the system is to oscillate with large amplitude.

As another example of driven resonance, consider a large metal plate suspended in front of a loudspeaker cone, as depicted in the left half of figure 3.9. Let's say that the metal plate has resonant frequency f_0. (We could measure this frequency by hanging the plate by a string and tapping lightly on the plate with a rubber mallet, measuring the pitch of its response tone.) By adjusting the input signal we can vary the frequency of the speaker output to any value we choose. For most frequencies the vibrational response of the plate will be minimal, and we might not notice it vibrating in response to the speaker signal unless we have a very sensitive pickup attached to it. But it does vibrate with low amplitude at the same frequency as the speaker cone with low amplitude. When the speaker frequency f_d comes closer in value to f_0, we notice that the plate begins to respond with larger amplitude, and in a very narrow range surrounding f_0 the plate will vibrate considerably. When the speaker is driven at exactly f_0, the plate will have its largest amplitude response.

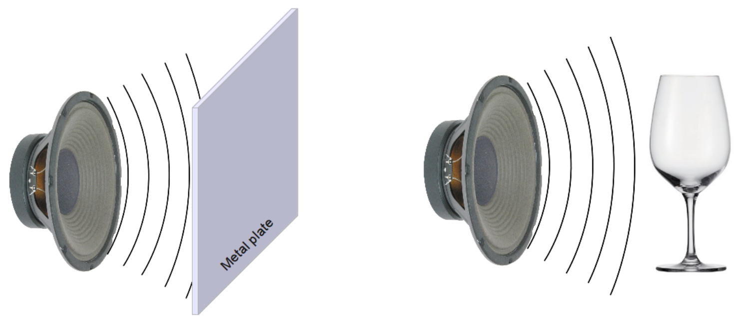


Figure 3.9: Driving a metal plate (left) and a wine goblet (right) into vibration using sound waves. When the frequency of sound is equal to the resonant frequency of the plate, it vibrates with large amplitude and buzzes. When the goblet is set into resonant vibration, the amplitude of oscillation may become large enough to cause instability in its mechanical structure, and if there is not sufficient damping present in the goblet's structure, it may shatter.

3.4.4 Damping Mechanisms

When a system is set into resonant vibration and then allowed to oscillate freely, its oscillations will eventually die out because of internal *friction forces* that serve to damp the resonance. Damping forces take energy away from the oscillations and typically turn it into heat. For example, whenever a spring is stretched or compressed, it produces a little bit of heat. The moving mass pushes the air out of its way as it moves, creating drag. If the mass is sliding on a tabletop, surface friction takes energy away from the moving mass. All of these frictional forces eventually cause the mass to run out of energy, and it slows to a stop. Systems with low damping tend to sustain their oscillations for a long time, whereas systems with high damping die out quickly.

Another damping mechanism by which the energy of oscillation is diminished is the emission of sound. A struck tuning fork rings loud at first, but dies away in volume as the energy of oscillation is converted to sound emitted into the surrounding air. Examples of damping in musical instruments include a piano whose key is held down after striking, so that the tone sustains for a long time (low damping), and a flute for which the tone disappears as soon as the blowing stops (high damping). Figure 3.10 depicts the diminishing amplitude of oscillation as a function of time for three systems, one with low, one with medium, and one with high damping.

Most vibrating resonant systems have sufficient internal damping forces so they can oscillate at resonances with fairly large amplitude without experiencing structural damage. If an oscillating system has insufficient internal damping to keep the oscillation amplitude at sustainable values, destruction of the system can occur. One example of such a system is a wine goblet, driven into oscillation by rubbing a moistened finger around its rim, causing a clear pitch whose frequency corresponds to its resonant frequency. This method of exciting the goblet into oscillation will not likely cause it to break.

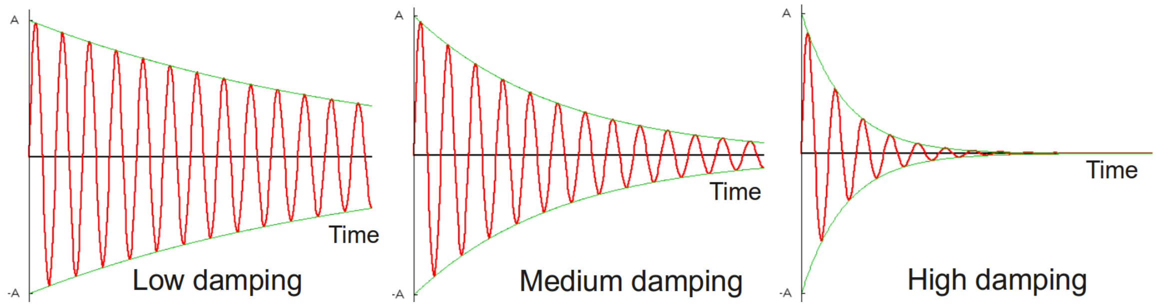


Figure 3.10: Three examples of damping in an oscillator: low, medium, and high damping. The lower the damping, the longer the oscillations persist before dying out.

Another way to drive its oscillation is by exposing it to sound waves of the proper frequency. In the right half of figure 3.9 we have replaced the metal disk with a wine goblet. We could find the goblet's resonant frequency with a slight tap and listen to the pitch it emits. We could then scan the speaker's output frequency for a range of frequencies which include this resonant frequency, and map out a resonance curve similar to that depicted in figure 3.8. The wine goblet has relatively weak internal damping forces, as evidenced by the long duration of sound it emits when tapped (certainly longer than that of the metal plate!). If the sound at resonance is sufficiently loud, sound waves impinging on the goblet can produce oscillations that are prohibitively large, exceeding the internal atomic forces that hold the goblet together, and the goblet can shatter. Some singers have been known to break goblets using just their well-trained voice at the right pitch.

When musical instruments are driven into oscillation, they produce sound and heat, both of which serve to remove energy away from the oscillations. There are a number of ways in which musical instruments are driven into resonant vibration, including drawing a bow across strings, plucking strings, hammering strings, blowing into a mouthpiece, striking with a mallet, and others. It's helpful to envision the dynamic energy balance that occurs when a musical instrument produces a long, sustained note at constant loudness. The conservation of energy tells us that all of the energy put into the instrument by the performer must be accounted for through energy being expended by the instrument. As energy is pumped into the instrument by whatever the external driving mechanism is (for example by drawing the bow across the violin strings), it is continually transformed into the output of sound and heat. The sound output for some instruments accounts for the majority of their energy output (in which case the heating represents a relatively small portion of the spent energy), while for other instruments the production of sound accounts for very little of the energy output (in which case the heating represents a much larger fraction of the spent energy). Remember that power is defined as energy expended per unit time. Therefore, a more helpful and proper way to describe this dynamic balance is to say that power being pumped into the instrument is balanced by the power being emitted and expended by the instrument.

Some instruments are constructed with a body designed to support several natural or resonant frequencies. In this case the body serves as a *resonator* to amplify the sound output of the instrument. Examples of this are the violin and the guitar, in which rather weak sounds produced by vibrating strings set the wooden body into resonant vibration, thus serving to amplify the sound delivered to the surrounding air. We will take up the topic of musical instruments in chapter 5.

3.5 Chapter Summary

Key Notes

- Three important ingredients in the production of sound: Production, Propagation, Perception. Sound is initiated by the creation of vibrations; these vibrations propagate through the air; they eventually activate vibrations in the auditory system of a listener.
- **Mass**: the quality of a body that resists being moved, measured in kilograms, (kg).
- **Speed**: the rate of change of a body's position in time, $v = \dfrac{\Delta x}{\Delta t}$, measured in meters/sec (m/s).
- **Force**: a "push" or "pull" by one body on another, measured in Newtons (N).
- **Pressure**: the force applied over an area, $P = \dfrac{F}{A}$, measured in Pascals (Pa).
- **Work**: the product of force applied to a body and the distance moved, $W = F \cdot d$, measured in Joules (J).
- **Energy**: a quantity belonging to a system that represents either the potential do work, or inherent in the motion of a body.
- **The Conservation of Energy**: states that energy can neither be created nor destroyed in any process. Although energy may change from one form to another, the total amount of energy in a closed system remains constant.
- **All sound**: begins with vibrations which propagate through a medium, requiring energy to get set into motion.
- **Periodic Motion**: the type of motion that produces sound is *periodic*. The **period** T is the time it takes an oscillator to complete one full cycle of motion. Examples of periodic motion include the earth's rotation around the sun (period = 1 year), the pendulum motion in a grandfather clock (about 1 second), the rise and fall of the tides (about 12 hours), and a child on a swing (about 3 seconds).
- The simplest type of periodic motion is called simple harmonic motion, and motion of the oscillator in time is sinusoidal in shape. Periodic motion can be much more complicated in shape.
- **Frequency**: the number of complete cycles (for a periodic system) that occur per second, typically measured in units of "Hertz" (1 Hz = cycle/second). Related to period by $T = 1/f$, $f = 1/T$.
- **Amplitude**: the "extent" of motion as measured by the medium's displacement from equilibrium. As an example, the weight on the end of a grandfather clock pendulum may swing a total distance of 10 cm from its equilibrium position, which corresponds to its amplitude.
- For some periodic motions, the larger the amplitude, the larger the period (as is the case for the planets going around the sun), while for other systems (the kind we are in fact interested in for musical purposes) the period is a constant, independent of the amplitude (as in the case of a pendulum or vibrating string). This latter case (*i.e.* where period is independent of amplitude) is true of simple harmonic motion (SHM).
- Periodic motion takes place about an **equilibrium position**. The force acting on a body that causes it to execute periodic motion is always *directed toward the equilibrium point*, no matter which way we displace the object from equilibrium. Hence, whenever we disturb a system from its equilibrium position, the restoring force attempts to return it to that point. The force F (**spring force**) increases with increasing distance of displacement from equilibrium, and is represented by the algebraic relation $F_s = -kx$, where k is a constant characterizing how difficult it is to displace the oscillator from equilibrium (the "springiness" of the system) and x is the measure of distance of the object from its equilibrium point.
- Note that the spring force is always *negative*, since it is always directed oppositely from the displacement (*i.e.* back toward the equilibrium position of the body).

- The maximum distance that an object moves away from its equilibrium point during oscillation is its amplitude.
- All physical systems that can undergo vibration must have the two qualities of *elasticity* and *inertia*. The medium's elasticity causes it, once stretched from equilibrium, to return to this state, and the medium's inertia corresponds to its particle's resistance to being moved.
- The period and frequency of a periodic system are the inverse of one another. The period is the time through which an oscillating system moves through one complete cycle, and the frequency is the number of cycles the system executes per second.
- When a mass on a spring is set into oscillation, energy is transferred from potential energy U_s to kinetic energy K and back again, with a period equal to the mass's oscillation period. The potential energy for a stretched spring is $U_s = \frac{1}{2}kx^2$ and energy conservation states that the system's energy is at all times equal to a constant, $E_{total} = U_s + k$.
- While we can "drive" a system at any frequency we choose (i.e. you can push a child on a swing with whatever frequency you desire), a system will oscillate at its *natural frequency* if first displaced from equilibrium and then allowed to oscillate on its own.
- The natural oscillation frequency for a mass on a string is given by $f_s = \frac{1}{2\pi}\sqrt{\frac{k}{m}}$.
- *Resonance* occurs when a system is driven at a frequency near or at its natural frequency. A system might have several natural frequencies.
- When a system is driven at its natural frequency, the amplitude of oscillation becomes large. If internal damping mechanisms are not sufficient to keep the amplitude small, the system may become unstable and break if the driving force drives the amplitude too large.
- All oscillations are damped, meaning that they eventually die away, owing to the presence of dissipative forces such as friction. The struck tuning fork rings loudly at first, and dies away in volume as the energy of oscillation is emitted into the surrounding air in the form of sound. Examples of damping in musical instruments might include a piano with its key held down after striking a note (relatively low damping, tone sustains for a long time), and a flute (high damping, the tone disappears as soon as the air stops producing it).
- Damping is very important in vibrating systems, taking energy away from the oscillation so that the amplitude does not become prohibitively large and cause physical damage. When a musical instrument resonates, the emission of sound is one form of damping (internal friction is another).

Exercises

Questions

1) What characterizes the inertial mass of a body?

2) In words, how is pressure related to force?

3) When I push on a wall with a force of 100 N for 5 minutes, am I doing work? Why or why not?

4) What does the principle of the conservation of energy say? How does it apply to musical instruments?

5) Can a waveform be periodic and yet not simple harmonic? Explain briefly
6) What characterizes a simple harmonic oscillator? Give an example of one.
7) What is the character of a restoring force? How is it useful in the function of musical instruments?
8) What defines an equilibrium point for a simple harmonic oscillator? What is its importance?
9) What two physical features of a system will enable it to undergo simple harmonic motion? Explain briefly the function of each.
10) What is the role of damping in oscillators?
11) What is a natural frequency of oscillation? Can a system have more than one? Explain briefly.
12) What might happen to a resonant system if it has insufficient damping? Explain briefly.

Problems

1. Draw a 100 Hz sine wave with amplitude A as a function of time. Be sure to mark the x-axis with tick marks and values appropriate to the problem. For each of the following parts, draw a wave that differ from this one in the characteristic specified: a) higher frequency b) lower amplitude, c) different wave shape (non-sinusoidal), d) larger period.

   Solution:

2. A serious question can be raised about one graph in figure 3.4 as to whether it can really represent motion. Which graph is it, and why is it that it cannot represent a true motion?

3. Calculate the average speed in each of the following cases:

   a) A flute on the back seat of a moving car travels 25 m in 3 s.
   b) A train carrying a the Juilliard Quartet travels 2 km, the first km at an average speed of 50 km/h and the second km at an average speed of 100 km/h. (*Note:* the average speed is not 75 km/h)
   c) A saxophonist on a run who covers a distance of 1 km in 3 min and a second kilometer in 4 min.
   d) A note dropped from a height of 75 m which strikes the ground in 4 s.

4. Suppose a xylophone mallet with a hard-rubber head remains in contact with the wooden bar for only a few thousandths of a second, but during that short time it exerts a force of 500 N. If that force is concentrated in a contact area of only 5 square millimeters (*i.e.* 5×10^{-6} m^2), how much pressure is being exerted on that part of the bar? Express your answer in Pa (N/m^2).

5. Calculate the kinetic energy of a 1500 kg automobile with a speed of 30 m/s If it accelerates to this speed in 20 s, what average power has been developed?

6. If a steady force of 2 N applied to a certain spring stretches it 0.2 m, what is its spring coefficient k?

7. The force law $F = -kx$ means that the larger the displacement from equilibrium (x), the larger the restoring force (F). If a spring has a spring constant of 71 N/m, what is the force required to hold the spring in a stretched position 0.075 m beyond its equilibrium position?

8. If you push against a car with a force of 400 N but the brakes are set and the car goes nowhere, how much work are you doing? If the brakes are released and you push this hard while the car moves a distance of 30 m, how much work have you done?

9. A block of mass $m = 2.0$ kg is connected to the end of a horizontal spring whose other end is mounted to a wall. The mass sits on a frictionless table and is free to oscillate about the spring's equilibrium position with no friction. The spring constant for the spring is 16 N/m. If the mass is pulled from equilibrium by 6.0 cm and then released,

   a) What is the force required to hold the mass in place before letting go?
   b) In what direction is the force directed?

10. a) What is a linear restoring force? b) How is SHM related to linear restoring forces?

11. You are at a concert, and before the performance begins, the musicians on stage tune their instruments. In order to get it all started, the principle oboist plays a long "Concert *A*", at a frequency 440 Hz. Calculate the period of time between successive vibrations of the oboe air column - that is, calculate the time between successive vibrations for this sound.

12. A hummingbird can flap its wings at around 80 Hz. What is the time interval between flaps?

13. When a bassoon of unknown mass is attached to a spring with spring constant 1,200 N/m, it is found to oscillate with a frequency of 6.00 Hz. Find a) the period of the motion, and b) the mass of the body.

14. A harmonic oscillator consists of a 0.500 kg mass attached to a spring with spring constant 140 N/m. Find a) the period of oscillation, and b) the frequency of oscillation.

15. Hanging a mass of 1 kg on a certain spring causes its length to increase by 0.2 m.

    a) What is the spring constant k of that spring?

    b) At what frequency will this mass-spring system oscillate on its own when started?

16. Given that the lower limit of human hearing consists of tones with a frequency approximately 20 Hz, what is the period of such a tone? The upper limit is around 20,000 Hz. What is the period of this tone?

17. a) Draw a curve of damped harmonic motion as a function of time. b) Draw a graph of motion that decreases in *both* period and amplitude as time progresses. Is this damped motion harmonic? Why or why not?

18. a) Draw a graph of two sinusoidal waves of the same frequency and amplitude that differ in phase by 180°. Draw a graph of two sinusoidal waves of the same frequency but b) with different amplitude, c) with a phase difference of 90°, d) with a phase difference of 45°.

19. Describe physically the relationship between the motions of two pendulums whose oscillations a) are in phase, b) are out of phase, and c) differ in phase by 90°.

20. Define *resonance*. Give two examples of resonance from music, and two from other fields.

21. Will the amplitude of a struck tuning fork decay more quickly when the stem of the fork is *on* or *off* a resonator box that vibrates at the same frequency? Why?

22. A moderate earthquake hits downtown Los Angeles, and as a result of the quick movement of the ground underneath one of the tall buildings, its top sways back and forth with a period of 8 seconds. The building is equipped with natural damping such that its oscillation dies out in approximately 30 seconds. Draw a plot representing the motion of the building top as a function of time, if the building top's initial displacement from equilibrium is 2.5 m.

23. The shape of the resonance response curve is dependent upon the amount of damping present in the oscillator. The smaller the damping, the taller and more sharply peaked the response curve. In other words, the smaller the damping, the larger the response of the oscillator when it is driven at the resonant frequency.

    a) Which would be more responsive at frequencies off resonance, an oscillator with large damping or one with small damping?
    b) Can you anticipate the acoustical significance of the answer you gave in a)? (We shall understand the significance of this later.)

24. Imagine a simple harmonic oscillator with a natural frequency of 200 Hz. A periodic driving force is applied to the oscillator to

set it into vibration. For each of the following periods of the driving force, calculate the frequency, and describe the nature of the amplitude of the oscillator:

a) 0.00556 s
b) 0.00455 s
c) 0.00500 s
d) 0.00375 s
e) 0.00625 s

Make a plot of the approximate amplitude of the oscillator as a function of frequency.

References

[1] The Atmosphere, Lutgens and Tarbuck, 8th edition (2001)

[2] http://classroomclipart.com/clipart-view/Clipart/Sports/Bowling_Clipart/holding_a_bowling_ball_in_hand_jpg.htm

Chapter 4

Propagation of Sound

4.1 Waves

If a mechanical vibration is to generate sound that can be perceived, the energy associated with the vibrations must propagate in the form of a *wave* through a *medium*. Several types of media, including solids, liquids, and gases are capable of supporting the propagation of waves. For musical sound that medium is *air*.

Definition: A wave is a *disturbance* that moves through the medium.

Imagine, for example, that you are holding the end of a long piece of rope that is tied to a wall. If you were to quickly snap the rope up and back down again, the hump you form in the rope would move along the rope's length, reach the wall, and bounce back toward you (see figure 4.1).

What is it that actually travels along the length of the rope? It's clearly not the rope itself, but the hump that you've created. More precisely, the force of your snap adds energy to the rope, disturbing it from its equilibrium position. This disturbance moves along the rope at a specific speed as a form of moving energy. Notice that the rope itself does not go anywhere (except briefly to move up and back down to equilibrium), but a *wave* propagates the entire length of the rope, consisting of the original disturbance you caused. We can therefore see that the wave consists of traveling energy in the form of a disturbance that moves through the medium.

The most important wave we will consider in this course is the *sound wave*. Remember back at the opening of chapter 3 we considered briefly the sound produced by a violin. Vibrations in the string and body of the instrument, including the air within its volume, cause disturbances in the air molecules immediately adjacent to the violin, which bump up against their neighbors, which bump against theirs, and so on, and in this fashion the initial disturbance induced by the vibrating violin propagates through the air as a wave with a specific speed. The air molecules themselves do not move much, just slightly back and forth, but the wave moves out a great distance from the violin. In this way the *source* of sound (the vibrating violin string, body, and enclosed air) produces a wave that *propagates* outward in the form of traveling energy, and a sound wave is born.

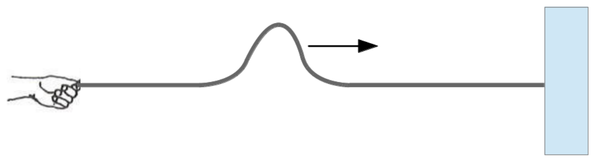


Figure 4.1: A pulse moving along a rope. The pulse represents a "disturbance" in the medium (which in this case is the rope) which travels along the medium.

Nature is filled with waves of many different kinds. Think of waves moving along the surface of the ocean. While the actual water does not move, except briefly to rise and fall as the wave passes, the *disturbance* in the water's surface propagates along the surface for a long distance at a specific speed.

In summary, then we see that an important characteristic of waves through a medium is that the particles of the medium do not move much (except "locally"). The energy of the deformation travels through the medium with a characteristic speed and direction.

4.1.1 The Medium

A *medium*, for our purposes, is a large continuous body of matter, be it solid, liquid, or gas. The particles that make it up (i.e. atoms or molecules) are in relatively close proximity to each other, and they interact with one another, either by mutual attraction (as in solids or liquids) or repulsion (as in gases). The common feature these three types of media share is that they can each support the propagation of a disturbance in the medium. The disturbance can move through the medium by virtue of the interaction between neighboring atoms or molecules. For example, imagine the surface of a quiet lake. When I cast my fishing lure out onto the surface, the impact of the small lure causes the initial disturbance (or deformation) in the glassy surface. This disturbance spreads outward from its original point. The very reason it can move is that molecules are connected to their neighbors, and when one of them is displaced from its quiet point of rest, its neighbors get pulled into the action as well. As one molecule moves up, it lifts its neighbor, which in turn lifts its neighbor, and so on. As it reaches the top of its motion and begins to descend, it pulls its neighbor back down, which in turn pulls its neighbor, and so on. The original disturbance propagates outward in this fashion. Figure 4.2 depicts how the medium movement is passed along. The medium is depicted as a series of masses connected by springs. The masses represent the inertial component of the medium, and the springs represent the elastic component of the medium.

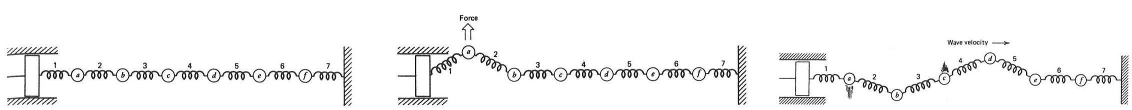


Figure 4.2: An array of masses connected by springs, as a model of a medium carrying a wave. The initial disturbance propagates through the medium by virtue of the connection between neighboring masses.[1]

The speed with which the disturbance moves outward on the water, or in any medium for that matter, is related to two important features of the medium: the "stiffness" (or elasticity) of the forces between neighboring particles (attractive or repulsive), and the "massiness" (or inertia) of the constituent particles. The stiffer the medium, the faster a wave moves through it. This is because the elasticity determines how *quickly* the medium "snaps back" to equilibrium after being disturbed. Likewise the more massive the constituent particles, the slower the wave moves through it. This is because the more massive an object, the more sluggish or *slow* it is to respond to the restoring force attempting to move it back to equilibrium. Water, when disturbed from equilibrium on the surface of a still lake, is fairly slow to move, compared with the stiff rope. Waves move very quickly through a crystalline material, since the atoms are bound with strong forces and therefore snap back to equilibrium very quickly.

4.1.2 Medium for Sound - Air

The medium in which sound usually travels is air. In order to appreciate how fine the medium of the air is (i.e. how small the molecules that make it up are, and how many of them there are), consider the following. The question is, how big is a molecule? An alternative way of looking at this is asking how many of them would fit in a particular volume. It turns out that the number of molecules in the volume of 22.4 liters, about the same size of a medium UPS package (28 cm on each edge), is around 6×10^{23}. Put in terms of dollar bills, we could imagine this number as follows. Imagine stacking clean, new dollar bills. How tall would the stack be in order to contain 6×10^{23} dollar bills? Answer: A stack of 1 dollar bills large enough to circle the earth *3.5 trillion times*. Quite difficult for us to imagine such a large number.

4.1.3 Longitudinal and Transverse Waves

We've established that a wave is a disturbance that travels through a medium, but note that there are really only two directions that the medium can move, relative to the direction of the wave motion. The particles in the medium can either move along the same direction as the wave (a *longitudinal wave*), or perpendicular to it (a *transverse* wave).

As an example of a longitudinal wave, consider a stretched slinky arranged horizontally on a tabletop. Regions of compression and stretching between the rings move along a slinky when its end is periodically pushed and pulled. As the resulting pulse, or wave, moves along the slinky's length, its rings move briefly from their equilibrium positions along the length of the slinky, and back again. Figure 4.3 depicts such a longitudinal wave.

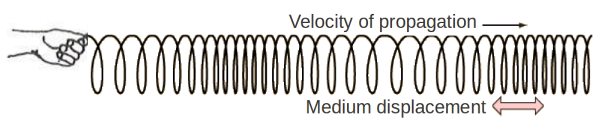


Figure 4.3: Compression wave moving along a slinky. The slinky loops move back and forth in the same direction as the wave propagation

An example of a transverse wave is the one discussed earlier concerning the section of rope. As the pulse moves along the rope's length, each piece of the medium (*i.e.* the rope) moves briefly up and back to its equilibrium position, *perpendicular* to the direction of wave propagation. Figure 4.4 depicts a transverse wave.

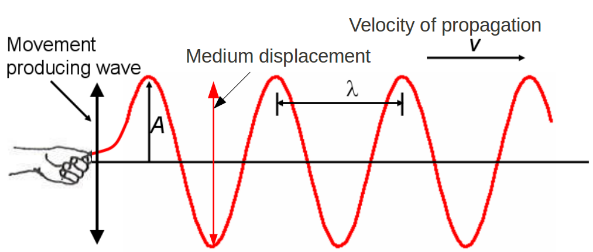


Figure 4.4: Transverse wave – the portions of the rope move in a vertical direction, perpendicular to the direction of wave propagation, which is horizontal.

Sound waves are initiated by vibrations, as we have learned already. These vibrations get communicated to the air by contact between a vibrating surface and the air. As the surface vibrates, it bumps up against molecules immediately adjacent to it, and these molecules then communicate this bump to their neighbors. Figure 4.5 depicts the effects of a moving piston on a column of gas. The initial piston movement bumps up against the first molecules on the left, which then turns into a pulse that propagates down the tube by the intermolecular forces between the molecules. If we drive the piston using a tuning fork, as in figure 4.6, the principle is the same, except that now the pulses communicated to the gas are *periodic*, resulting in waves with particular frequency and wavelength.

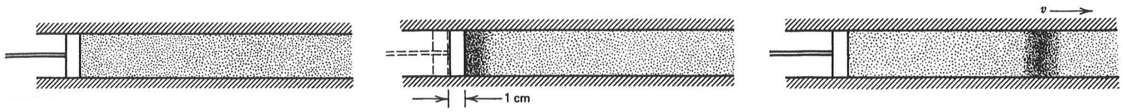


Figure 4.5: The pulse initiated by the moving piston gets communicated down the column by the forces between molecules. Before the piston moves (left) the air pressure is uniform. When the piston moves, a compression zone appears adjacent to it (middle), which then propagates down the column by the intermolecular forces (right)[2].

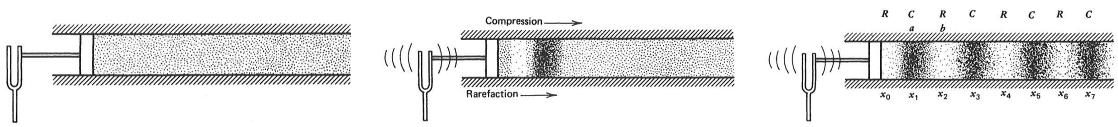


Figure 4.6: The figure on the left depicts the uniform air pressure in the column before the tuning fork is struck. The middle figure shows the first compression waves moving down the column shortly after it is struck, and the right figure shows the same waves a little later with several successive compression zones moving down the tube with a specific wavelength[3].

The pressure variations in the gas inside the cylinder consist of regions of compression and rarefaction, compared with atmospheric pressure. Standard air pressure at sea-level is about 1.013×10^5 Pa. The variations in pressure between the compression zones and normal zones, for a sound of moderate intensity (normal conversation, for example), is only about 2×10^{-2} to 2×10^{-3} Pa, or a few thousandths of a Pa! For sounds at the very bottom of our ability to hear, the pressure variations are on the order of 2×10^{-5} Pa, and for sound loud enough to cause hearing damage, the variations are around 20 Pa. For this latter case, the compression zones depart from atmospheric pressure by around two hundredths of 1%. This means that the longitudinal wave in figure 4.7 shows compression zones that are not to scale. The variations shown in the figure are far larger than those typical of sound in air.

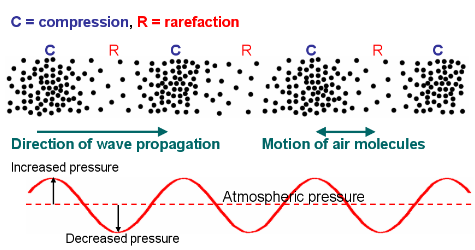
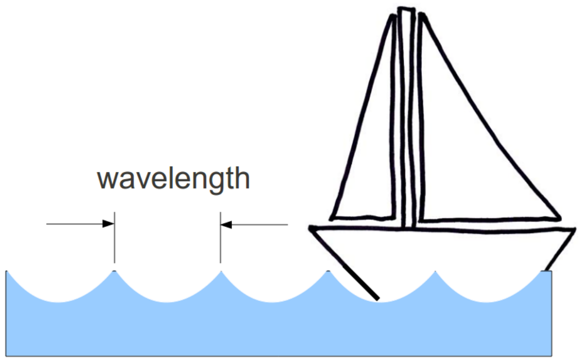


Figure 4.7: (***For color version see Appendix E***) 'The medium for sound is air. Waves are longitudinal, meaning that the molecules of air move back and forth in the direction of the wave propagation. The physical distance between neighboring compression zones is called the wavelength (see section 4.1.4). The above is a graphical representation of the pressure variations in the medium corresponding to the moment the wave is depicted. Compression regions are *above* atmospheric pressure, and rarefaction regions are *below* atmospheric pressure.

4.1.4 Wavelength

When a source of vibrations produces a wave that travels through a medium, the wave appears "stretched out" in space as it moves. Consider, for example, a wave moving through water, consisting of a large number of equally spaced water crests marching along, perhaps produced in the wake of a passing ocean liner. We can separately analyze the *space domain* or the *time domain* characteristics of the wave (fancy ways of saying that we can choose to look at periodic motion as it spreads out in space, or as it evolves in time). When we sit in a boat and let the wave pass underneath us, we can time the period of the rising and falling of the boat, giving us the period of the moving wave. Alternatively, we can take a "picture" of the wave on the water and use it to analyze the distance between neighboring crests. This spatial information gives us the *wavelength* of the wave (see figure 4.8).

The wavelength is defined as the distance between successive peaks *corresponding to one full cycle* of periodic motion. In the case of the water waves in the figure, the wavelength is the physical distance between neighboring crests on the water, and is measured in units of distance, *i.e.* meters. Note that the wavelength can be determined using *any* two points on the curve provided they are separated by one full period, whether they be at the peaks, the troughs, the zero crossings, or some other points (see the right half of figure 4.8).



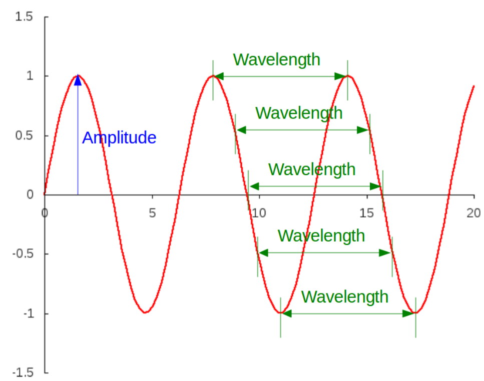


Figure 4.8: Water wave with periodic shape. The wavelength is defined as the distance between successive crests, or the distance between any two similar points on the curve that are separated by one cycle.

4.2 Wave Speed

The speed of an object is defined as the distance it moves in a specific time. A wave is not an object, of course, and therefore it's reasonable to ask how to understand the speed of a wave. The speed of a water wave is perhaps easiest to measure since we can fix our attention on a wave crest, and measure how far it advances in a fixed amount of time. Dividing the distance by the time results in the value of the wave speed. On the other hand, we can't "see" the crest of a sound wave. However, the concept is exactly the same. If we were able to fix our attention on one of the crests in the pressure wave moving through the air, we could simply measure the distance it moves over a fixed time to calculate its velocity. There is a particularly easy way to calculate the speed of a wave if we know both the wave's frequency and wavelength. Imagine sitting on the sidelines with a stop-watch as a wave moves by. We could start the watch when one of the crests is in front of us, and then wait until the next one is in front of us, at which point we would stop the watch - see figure 4.9.

The distance the wave has moved is, by definition, one wavelength, which corresponds to the length between neighboring crests. The time it takes to move that distance is, by definition, the period T, which corresponds to the time it takes for a wave to advance through one cycle. The velocity is obtained by dividing the distance, λ, by the period T. Since the period is the inverse of the frequency, we can see that the speed is calculated using

$$v = \frac{d}{t} = \frac{\lambda}{T} \qquad \rightarrow \qquad \boxed{v = f\lambda}. \tag{4.1}$$

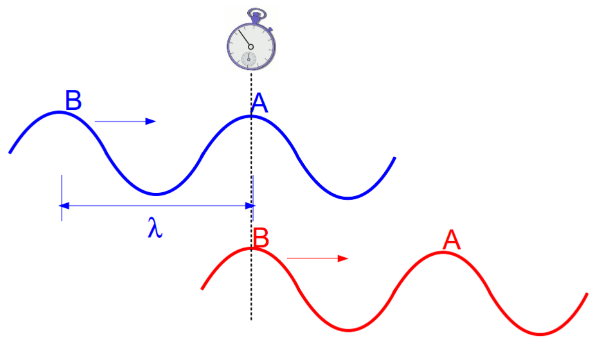


Figure 4.9: Measuring the speed of a wave by watching it advance one complete cycle. The watch is started when the wave crest "A" is at the watch's location (upper wave). The wave advances forward one full cycle, at which point wave crest "B" is at the watch (lower wave). The time it takes to advance one full cycle is the period T, and the distance it has moved is the wavelength λ.

4.2.1 Speed of Sound

The temperature of a body is a macroscopic measure of how vigorously its individual molecules are moving – its value is a measure of the average kinetic energy of the molecules. For a solid material, this kinetic energy is that of the molecular vibration about their equilibrium points. The two most common temperature scales used in daily life are Fahrenheit (in the US) and Celsius (pretty much everywhere else!). Both of these scales have positive and negative values. The origin (or zero point) of the Celsius scale corresponds to the freezing point of water, for which the molecules still have non-zero vibrational kinetic energy. Negative temperature values simply correspond to situations where the average molecular vibrational kinetic energy is less than that corresponding to zero degrees. The *absolute temperature scale,* on the other hand, measured in "Kelvin" has no negative values, since 0 Kelvin corresponds to a *complete cessation* of all molecular motion. In this sense it really is "absolute." Room temperature is generally accepted to be about 70°F which is 21°C, or 295 K.

The following formulas are helpful for converting from one scale to another. Before calculating any wave speeds in air, we need to express the temperature in the absolute scale.

$$\boxed{T_F = \frac{9}{5}T_C + 32 \qquad T_C = \frac{5}{9}(T_F - 32) \qquad T_K = T_C + 273.15}. \tag{4.2}$$

So how cold is absolute zero? In Fahrenheit, absolute zero corresponds to -460°F! As a comparison, air begins to liquefy at 77 K, or -321°F,a and the coldest temperature recorded on the earth was -129°F, in Vostok, Antarctica [4].

For sound waves in air the speed is a function of temperature:

$$\boxed{v = 20.1\sqrt{T_A}} \tag{4.3}$$

where T_A is the *absolute temperature*, measured in units of *Kelvin*, and the number 20.1 results from the natural properties of air, and gives the speed in MKS units as long as the temperature is in Kelvin. The speed of sound at room temperature is

$$\text{Room Temperature, 21°C: } v_s = 20.1\sqrt{T_A} = 20.1\sqrt{294 \text{ K}} = 345 \text{ m/s}. \tag{4.4}$$

When temperatures are expressed in Kelvin, the *variations* we typically experience on the surface of the earth are relatively small. For example, between a cold and a hot day in Santa Barbara (say from 40°F to 95°F), the difference in these temperatures expressed in Kelvin amounts to about 10%. Since the speed of sound depends on the square root of the temperature, this means that sound will vary in speed by about 5% on a very cold day compared with a very hot day.

Example 4.11

Speed of Sound at an Outdoor Concert *It is a very hot day, 100°F, and we are sitting at an outdoor concert, 45 m from the stage. How long does it take for the sound of the tympani to reach us from that distance? Calculate the difference in travel time for a very cold day, 32°F. Which time is shorter, and why?*

Solution: First we need to calculate the absolute temperature for the hot day in K,

$$T_C = 100F \cdot \frac{5}{9}(100\text{ F} - 32) = 37.8\text{ C}$$

$$T_K = T_C + 273.15 = 37.8 + 273.15 = 311\text{ K}.$$

The speed of sound is then

$$v_s = 20.1\sqrt{311} = 354.5\text{ m/s}$$

The travel time for the sound of the tympani to travel 45 m is then

$$v_s = \frac{d}{t} \quad \rightarrow \quad t = \frac{d}{v_s} = \frac{45\text{ m}}{354.5\text{ m/s}} = 0.127\text{ s}$$

Let's now compute the same number for the cold day. The absolute temperature corresponding to 32°F is

$$T_K = T_C + 273.15 = \frac{5}{9}(32\text{ F} - 32) = 273.15\text{ K}$$

so that the speed of sound for this cold day comes out to

$$v_s = 20.1\sqrt{273.15} = 332.2\text{ m/s}.$$

The travel time for the sound on this cold day is then

$$t = \frac{45\text{ m}}{332.2\text{ m/s}} = 0.135\text{ s}$$

so that the difference in travel time between the hot and the cold day is

$$\Delta t = 0.135\text{ s} - 0.127\text{ s} = 0.008\text{ s}$$

or only 8 thousandths of a second. The speed is greater, and therefore the travel time shorter, on the hot day, since the speed of sound depends on the interactions of molecules in the air. The higher temperature means that the molecules are moving faster and therefore translate their collisions to one another more quickly.

4.3 Wave Characteristics

4.3.1 Reflection

When a traveling sound wave encounters a surface, part of it is absorbed by the surface, and part of it is reflected. If the size of the surface is small compared with the wavelength of sound, then instead of reflecting the sound, it will cause it to be dispersed throughout the room. Sound waves can also bend around corners. We will treat these wave behaviors in much greater detail when we consider the topic of Acoustics later in the course.

4.3.2 Intensity

When a musical instrument vibrates, it causes disturbance in the surrounding air, which moves out in all directions from the instrument in the form of *acoustical energy.* This energy spreads out into space in the form of an expanding sound wave. The rate at which the energy "flows" from the instrument, *i.e* the amount of energy emitted per unit time, is the power output of the instrument. This power moves outward from the instrument in all directions in the shape of an expanding sphere. Figure 4.10 depicts the surface over which the energy is spread at three different times (corresponding to the three surfaces) as the wave moves outward.

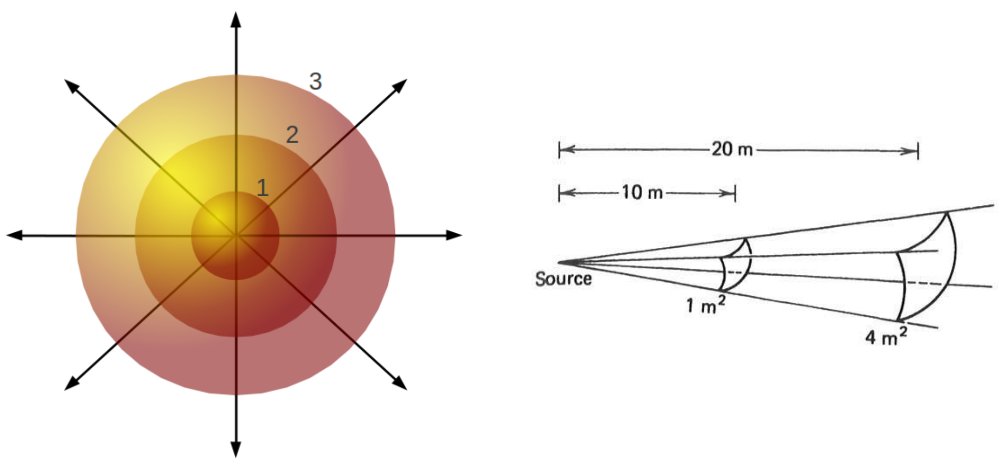


Figure 4.10: The figure depicts the area of a sphere as it expands outward from the instrument, at three successive times. As acoustical power spreads out over that sphere, whose area increases as the square of the distance from the instrument, the power *density* correspondingly *decreases* with the square of the distance. The figure on the right shows how a 1 m^2 portion of area quadruples to 4 m^2 when the distance from the origin doubles from 10 to 20 m, meaning that the power density drops by a factor of 4 over the same distance.

Since the surface area of the sphere ($4\pi r^2$) over which the energy is spread increases as the square of the distance from the origin, the power density must therefore *decrease* in strength by the same factor. The intensity I of a sound wave is defined as its power density, and is measured in units of Watts/m^2.

$$\boxed{I = \frac{\text{Power}}{\text{Area}} = \frac{P}{4\pi r^2} \qquad \rightarrow \qquad I \propto \frac{1}{r^2}}. \tag{4.5}$$

Considerations of power are very important in the production and perception of music. In order for sound to be perceived by our ears, it must rise above the threshold, or lower limit, of the ear's sensitivity. If the power exceeds the so-called "pain threshold" of the ear, it can produce significant damage, including a bursting of the eardrum.

Example 4.12

Intensity of Sound at Different Distances *A loud siren far removed from any reflecting surfaces has an intensity measured to be 0.25 W/m² at a distance of 20 m from the source. What would be the intensity of this siren if it were measured at a distance of a) 10 m, and b) 30 m from the siren? c) What is the total power output of the siren?*

Solution: a) Let's call the total power output of the siren P_{siren}. In terms of this power output, the intensity at 10 m would be

$$I_{10\text{ m}} = \frac{P_{siren}}{4\pi(10\text{ m})^2}$$

The ratio of intensities at 10 and 20 m is then

$$\frac{I_{10\text{ m}}}{I_{20\text{ m}}} = \frac{P_{siren}/4\pi(10\text{ m})^2}{P_{siren}/4\pi(20\text{ m})^2} = \frac{(20\text{ m})^2}{(10\text{ m})^2}$$

$$\to I_{10\text{ m}} = I_{20\text{ m}} \cdot \frac{(20\text{ m})^2}{(10\text{ m})^2} = 0.25 \cdot 4 = 1.0\text{ W/m}^2.$$

b) In moving to 30 m from the source, the intensity becomes

$$I_{30\text{ m}} = I_{20\text{ m}} \cdot \frac{(20\text{ m})^2}{(30\text{ m})^2} = 0.11\text{ W/m}$$

c) We can compute the power output of the siren knowing what its intensity is at any of these distances. Let's use the number given in the problem statement. At a distance of 20 m, the intensity of the siren is given as 0.5 W/m², so that

$$I_{20\text{ m}} = \frac{P_{siren}}{4\pi(20\text{ m})^2} = 0.25\text{ W/m}^2$$

$$\to P_{siren} = 0.25\text{ W/m}^2 \cdot 4\pi(20\text{ m})^2 = 1260\text{ W/m}^2$$

From this example we can conclude a useful relationship between intensities at different distances from the source. The change in intensity of a sound moving from distance d_1 to d_2 from the source is

$$\boxed{\frac{I_{d_2}}{I_{d_1}} = \left(\frac{d_1}{d_2}\right)^2 \qquad \text{or} \qquad I_{d_2} = I_{d_1} \cdot \left(\frac{d_1}{d_2}\right)^2}. \tag{4.6}$$

The enormous range of intensities perceivable by the human ear is astounding. The minimum intensity detectable is around 10^{-12} W/m^2, and the maximum intensity that can be heard before damage sets in is around 1 W/m^2. This corresponds to a "dynamic range" of twelve orders of magnitude, or 10^{12} – a factor of a thousand billion in difference. If the minimum intensity threshold were 1 inch, then the maximum hearing intensity level would be 15.8 million miles, or 17% of the earth-sun distance! It is very difficult to handle such a large range of numbers. As an interesting exercise, try and plot on one linear axis the numbers 1, 150, 1200, 10^4, 10^8 and 10^{12}. Once you've drawn your y-axis, you will quickly find that the lowest numbers are hard to locate separately on that axis – they are far too close in value to one another to be separately identified on the y-axis (see figure 4.11). For this reason, it will be helpful to develop a new measure of sound based on the *logarithm* of the intensity. This new measure of sound strength is called the *sound intensity level*, and it is expressed in units of decibels (dB). First a word about logarithms.

4.3.3 Sound Intensity Level

Table 4.1: 12 levels of sound over the range of hearing

Intensity (W/m^2)	Sound Level (dB)	Source
1	120	Vuvuzela horn at 1 m
10^{-1}	110	Jet engine at 100 m
10^{-2}	100	Jackhammer at 1 m
10^{-3}	90	Subway station
10^{-4}	80	Hair dryer
10^{-5}	70	Traffic, city street
10^{-6}	60	Conversation, TV set
10^{-7}	50	Classroom, dishwasher
10^{-8}	40	Quiet library
10^{-9}	30	Very quite room
10^{-11}	10	Leaves rustling, quiet breathing
10^{-12}	0	I_0, Threshold of hearing

The sound intensity level (SL) is defined as

$$\boxed{SL = 10 \log \frac{I}{I_0}} \tag{4.7}$$

where I is the intensity we are interested in characterizing, and I_0 is the threshold intensity of human hearing. We will often refer to it simply as the "sound level" for short. This is sometimes referred to as the *absolute* sound intensity level since it is in reference to the hearing threshold. Notice that instead of dealing with a range of intensity values that differ over 10^{12} between the minimum and maximum, the sound intensity level gives us a more convenient means of characterizing the same range, using a scale from 0 to 120 dB – see figure 4.11

If we wish to determine the sound intensity level *difference* between two sound levels I_1 and I_2, we can use

$$\Delta SL = SL(I_2) - SL(I_1) = 10 \log I_2/I_0 - 10 \log(I_1/I_0) = 10 \log \frac{I_2}{I_1}. \tag{4.8}$$

The sound intensity level is a *comparative* measure, *i.e.* the sound intensity level of one source relative to another. The "absolute" SL level is actually a comparative measure of a sound relative to the threshold of hearing.

Example 4.13

Loud Clarinet *What is the SL of a sound wave from a clarinet whose intensity is* $I = 10^{-3}$ W/m^2*?*

Solution:

$$\begin{aligned} SL = 10\log\frac{I}{I_0} =& 10\log\frac{10^{-3}\ \mathrm{W/m^2}}{10^{-12}\ \mathrm{W/m^2}} \\ =& 10\log(10^9) \\ =& 10(9) \\ =& 90\mathrm{db} \end{aligned}$$

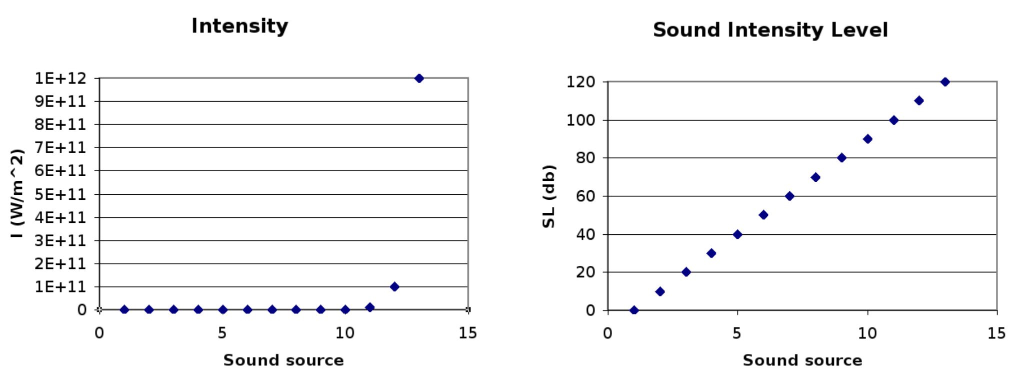


Figure 4.11: Comparison of a plot of intensity values over the entire range of human hearing, and a plot of sound level over the same range. Twelve points are plotted in each, ranging from the low to the high end of hearing range. The majority of points (whose values are of interest for music) are so small compared with the two highest values that they are squashed against the axis, and we can't distinguish their different values on the y-axis. The sound intensity level provides us a very useful means of depicting these same 12 levels since they all fit nicely on the plot when expressed in decibels.

Example 4.14

SL **Comparison** *What is the SL difference between two tones,* I_2 *and* I_1*, which differ in intensity by a factor of 2? That is,* $I_2 = 2I_1$.

Solution:

$$
\begin{aligned}
SL = 10\log\frac{I_2}{I_1} =& 10\log\frac{2I_1}{I_1} \\
=& 10(0.301) = 3.0 \text{ db}
\end{aligned}
$$

We can find a useful relationship for the change in sound level ΔSL in moving from distance d_1 to d_2 by using the intensity relationship in equation 4.6:

$$\Delta SL = 10\log\frac{I_2}{I_1} = 10\log\left(\frac{I_1\left(d_1^2/d_2^2\right)}{I_1}\right) = 10\log\left(\frac{d_1}{d_2}\right)^2$$

$$\boxed{\Delta SL = 20\log\frac{d_1}{d_2}} \tag{4.9}$$

where we have used the principle that $\log\left(a^b\right) = b\log a$.

4.3.4 Relation Between Pressure and Intensity

A sound wave is a longitudinal *pressure* wave with periodicity. The pulses of pressure and rarefaction move through the air at the speed of sound, and the separation between pressure peaks corresponds to the wavelength. Since this pressure wave carries energy, and the intensity of the wave depends on this energy per unit time per unit area, the question naturally arises concerning the relationship between pressure and intensity. When no sound wave is present, all of the air is at room pressure, which is $p_{atm} = 1.013 \times 10^5$ N/m^2. When sound moves through the air, the peaks and valleys of pressure (see figure 4.7 once again) are measured relative to the equilibrium room pressure. The peaks represent pressure *above* room pressure, and the valleys represent pressure *below* room pressure. For a typical sound, how much different is the pressure in the peaks and valleys relative to room pressure?

The minimum pressure difference to which the ear is sensitive is less than one billionth (10^{-9}) of one atmosphere. This means that the ear can (barely) sense a wave whose pressure peaks and valleys differ from atmospheric pressure by less than 1 part in a billion. This *threshold of hearing* is approximate, and differs from one person to the next. The ear is most sensitive at a frequency of around 1000 Hz, and at this frequency the threshold of hearing corresponds to deviations Δp from p_{atm} of 2×10^{-5} N/m^2. At the loud end of the scale, the ear can hear pressure differences of about one million (10^6) times this pressure difference, or 20 N/m^2, before pain and hearing damage are possible. Interestingly, this pressure difference is still only about 0.02% of atmospheric pressure.

The relation between pressure and intensity is simple:

$$\boxed{I = \frac{p^2}{\rho v_s}} \tag{4.10}$$

where ρ is the density of air and v_s is the speed of sound. If we call the smallest pressure difference to which the ear is sensitive $p_0 = 2 \times 10^{-5}$ N/m^2, then the *sound pressure level* is defined as

$$\boxed{L_p = 20 \log \frac{p}{p_0}} \tag{4.11}$$

where it is important to remember that both p and p_0 are pressures values *relative to standard atmospheric pressure*. The variable p is the pressure amplitude for the sound wave relative to standard atmospheric pressure, and the variable p_0 is the pressure amplitude for smallest wave to which the ear is sensitive, relative to atmospheric pressure.

The intensity and pressure vary with temperature. At most normal temperatures, the sound pressure level and sound intensity level have very similar values. The sound pressure level can be measured with a sound-level meter.

Example 4.15

Sound Pressure Level *A sound wave moves through the room whose pressure peaks and valleys differ from atmospheric pressure by 0.25 N/m^2. What sound intensity level does this correspond to?*

Solution: Since sound intensity level and sound pressure level are essentially equal to one another, we will calculate the sound pressure level.

$$L_p = 20 \log \frac{0.25 \text{ N/m}^2}{2 \times 10^{-5} \text{ N/m}^2} = 82 \text{ dB}$$

Hence the sound intensity level for this sound wave is about 82 dB.

4.4 Chapter Summary

Key Notes

- **Wave Propagation** – Sound requires a medium – air – for transport of waves. A wave is a disturbance that travels through a medium. Note that the motion of the actual medium under the influence of the passing wave is very limited, whereas the wave itself can move great distances.
- Many other types of media can carry sound waves, such as water, metal bars, string, etc. They all have a certain springiness and density that characterizes the speed with which the wave moves.
- **Longitudinal and Transverse Waves** - Longitudinal waves occur when the actual medium moves along the same direction as the wave propagation. Examples include sound and compression waves in a metal bar. Transverse waves occur when the actual medium moves perpendicular to the direction of the wave propagation. An example is the wave that travels along a piece of rope, or on a guitar string. Water waves turn out to be a combination of both longitudinal and transverse waves.

- Waves have *amplitude*, *wavelength*, *period*, and *frequency*. The wavelength λ is the distance between successive corresponding points, for example, between successive maxima.
- **Speed of a wave** – The speed v of a wave is defined as the distance it advances, divided by the time it takes to advance. For example, in order for a wave to move the distance corresponding to its wavelength λ, it takes an amount of time corresponding to its period T. Hence, $v = f\lambda$.
- For sound waves in air, in units of m/s, the speed is a slight function of temperature, $v = 20.1\sqrt{T_A}$, where T_A is the absolute temperature in Kelvin and the number 20.1 results from the natural properties of air.
- **Wave Reflection** – Waves can reflect from surfaces. The law of reflection states that the angle of reflection equals the angle of incidence. When a sound wave encounters a boundary such as a wall, it can be partially reflected and partially absorbed or transmitted. Waves can also refract around corners.
- **Intensity of Sound Waves** – Musical instruments radiate acoustical energy. The rate at which energy is radiated is the power output, $P = \Delta E/\Delta t$. Power is measured in Watts. The sound intensity is defined as the power transmitted through an area of 1m^2 oriented perpendicular to the propagation direction. See table 4.1 for the range of hearing for the human ear along with the intensity values for several familiar sounds.
- **Sound Intensity Level (SL)** – A convenient scale on which to register sound intensity is that of the sound level (SL), defined as $SL = 10 \log I/I_0$, where I_0 is the threshold of human hearing, around 10^{12} W/m^2.We can also characterize the relative sound level, say between sounds with intensity I_2 and I_1, $SL\text{relative} = 10 \log I_2/I_1$.

Exercises

Questions

1) What is a wave? Explain briefly.
2) Explain briefly the role of the particles in the *medium* in the propagation of a wave through it.
3) Define a longitudinal wave.
4) Define a transverse wave.
5) Can a wave be both longitudinal and transverse? If so, give an example.
6) What two physical properties of a medium are essential if it is to support the propagation of a wave? What are their roles?
7) Why is it that, all other things being equal, waves propagate through "stiffer" media at higher speeds than less stiff?
8) Briefly, why is it, from a physical standpoint, we would expect the speed of sound to increase with increasing temperature?
9) Briefly, why is it that intensity of a sound wave changes as the inverse distance squared from the source? Does this mean it increases or decreases with distance?
10) Why is the Sound Level (SL) a more convenient measure for characterizing the range of human hearing?
11) Doubling the frequency of a sound wave multiplies the wavelength by what factor? Is this always the case? Explain briefly.
12) Does a loud sound travel faster than a softer sound? Explain briefly.

Problems

1. The maximum audible range is about 15 to 20,000 Hz. Determine the range of wavelengths for sound over this range of frequency. Assume room temperature.

2. We measure the intensity of a sound source in open air to be 0.3 W/m^2 when we are located a distance of 15 m away from the source. If we were to move to a distance of 25 m, what would be the intensity of the sound? How about at a distance of 45 m?

3. The intensity of sound in a typical classroom is approximately 10^{-7} W/m^2. What is the sound level (SL) for this noise?

4. You are standing 1.5 m from a source of sound, and far from reflecting surfaces. The sound has, according to your measurement, a sound level of 70 db at your position. If you move further back to a distance of 4.5 m, what is the new sound level you would measure?

5. Sound coming from a siren has a frequency of 550 Hz. If the temperature of the surrounding air is 25°C, what is the wavelength of this sound?

6. What is the speed of sound on a very cold day whose temperature is 0°C? This is called Mach 1. If a supersonic airplane is traveling at Mach 2 (*i.e.* at two times the speed of sound at 0°C), find its speed in miles per hour.

7. How much will the speed of sound in a trumpet change as it warms up? Assume a change from room temperature (21°C) to body temperature (37°C). If the wavelength remains essentially the same (the expansion in length will be very small), by what percentage will the frequency change?

8. Suppose a sound wave has a pressure amplitude 0.3 N/m^2 above normal atmospheric pressure. What is the amplitude expressed as a number of atmospheres? What are the minimum and maximum pressures (in atm) in the presence of this wave? How about in the MKS unit of Pa?

9. A sound has an intensity of 10^{-5} W/m^2. What is the SL of this sound?

10. A trombone plays a note with a sound level of 85 db. If three more trombones join the first instrument and each plays with a SL of 85 db, what is the sound level of all four instruments combined?

11. If two sounds differ in level by 46 dB, what is the ratio of their sound pressures? Intensities?

12. A loudspeaker is supplied with 5 W of electrical power, and it has an efficiency of 10% in converting this power into sound power. This means that $1/10^{th}$ of the electrical power ends up as sound power. a) What is the amount of sound power emitted by the loudspeaker? b) If we assume that the sound radiates equally in all directions, what is the sound pressure level at a distance of 1 m? 4 m?

13. Find the sound pressure and the intensity of a sound with $L_p = 50$ dB.

References

[1] Rigden, Physics and the Sound of Music, 2nd Edition, pg. 27-8 (1985)

[2] Rigden, Physics and the Sound of Music, 2nd Edition, pg. 29 (1985)

[3] Rigden, Physics and the Sound of Music, 2nd Edition, pg. 30 (1985)

[4] Earth Science Revised Third Edition. New York. Prentice Hall, pg. 217 (1987)

Chapter 5

Musical Instrument Groups

To summarize what we've learned so far, in order for sound to be created there must be

- *motion*, in the form of vibrations, and
- *waves*, initiated by these vibrations that propagate outward as sound waves.

For musical sound, in particular, these vibrations needs to be *periodic*. The simplest form of periodic vibration in nature is that of a simple harmonic oscillator, which we studied in chapter 3, and whose vibrational pattern consists of sinusoidal variations. But in order for musical sound to be interesting and engaging, it will be necessary to move beyond simple sinusoidal motion to produce a wider, more complicated variety of vibrational patterns, because in *complex, non-sinusoidal* waveforms resides the richness of musical sound, the separate and unique personalities of the numerous musical instruments.

Musical instruments naturally produce complex waveforms. In this chapter we shall see that when an instrument plays a *single musical note*, it vibrates at not one but several frequencies simultaneously, frequencies that are related to one another mathematically. Remarkably, the human auditory system does not hear this collection of simultaneous vibrations as a multitude of sounds, but rather as a single note with a distinct pitch corresponding to the fundamental. The collection of the higher frequencies add to the sound of the fundamental in a way that gives the instrument its distinct "personality," which allows us to distinguish it from other types of instruments.

5.1 Instrument Attributes

The specific ***geometry*** and ***material constitution*** of an instrument determine its particular resonant qualities, which in turn determine the nature of sound it produces. Instruments range from very simple to very complex, and the nature and complexity of sound they produce varies widely.

Consider a very simple instrument – the triangle. It consists of a small hanging triangular-shaped metal bar and a small striker. The "body" of the instrument consists of this one bent bar, and it produces a family of resonant frequencies that result from its (rather simple) geometry and material construction. It doesn't play many notes – just one, in fact! And it doesn't require extensive training and practice (as far as I know!) for one to become a triangle aficionado. Compare this with a much more complicated instrument, such as a bassoon. The body consists of several materials, including wood, felt, and metal, and the shape of the body and positions and sizes of the keys give rise to a much more complicated resonant structure, one which allows for *several* natural frequencies and pitches, a feature that turns out to be very important in the operation of musical instruments.

In this chapter we'll study the way in which two important features of instruments determine their resonant character and therefore their sound output, namely the **geometry** (which determines the *boundary conditions* of the system) and the **material construction** of the instrument (which determines *speed and amplitude* of the propagating waves in the instrument). Simply put, the geometry of an instrument determines the family of vibrational wavelengths that are supported, and the material construction determines the speed with which these vibrational waves move throughout the body, which then determines the family of frequencies that the instrument produces.

5.1.1 Instrument Geometry – Boundary Conditions

The specific frequency or frequencies at which any object naturally vibrates, whether it be a taut string, a crystal goblet, an air column, *etc.*, are in part dependent on the object's specific geometry. In order to understand how musical instruments function and how they produce the variety of pitches they do, it's important for us to understand how their geometry affects their unique function – that is, we need to understand the "boundary conditions" that govern the function of the instrument. In the particularly simple case of a taut string, the two ends that are tied down form boundaries of the system. A slightly more complex case involves the flute, for which the boundaries consist primarily of the cylindrical tube, the mouthpiece, and the opening at the end. By operating the keys on the flute the flautist can alter the effective length of the instrument by adjusting its *boundary points*, which in turn determines the resonant frequencies at which the enclosed air vibrates and therefore the notes that it produces. The trombone has a continuously changeable length by operation of the slide, which allows for a continuous adjustment of its dimensions and therefore a continuous variation in its resonant frequencies. The dimensions of a drumhead's circular perimeter forms the boundary that contributes to the vibrational frequencies it produces.

5.1.2 Physical Constitution – Wave Speed

The second important feature that determines the function of musical instruments is that of their physical makeup, or material construction. When vibrations are produced in the body of the instrument (by plucking a guitar string, drawing a violin bow across the strings, blowing into the mouthpiece of a clarinet, *etc.*), the vibrations produced in the body and/or in the enclosed air propagate throughout the instrument, moving back and forth as they reflect from the boundaries. The speed at which the vibrational waves move is determined by the *elastic* and *inertial* characteristics of the medium in which the wave travels. Both the body of the instrument (made of some solid material) and any volume of enclosed air constitute the media in which the vibrational waves move.

5.2 Traveling vs. Standing Waves

In chapter 4 we considered *traveling* waves, the type that consists of a disturbance that moves from one place to another inside a medium. The most important of these in the physics of music is the *sound wave*, which propagates through the air. We learned that sound waves have their origin in a vibrational source that bumps up against nearby air molecules during each cycle of its vibration, which in turn bump up against their neighboring molecules, *etc.*, giving rise to a sequence of compression and rarefaction zones that move away from the source and constitute a traveling sound wave.

Another type of wave important to music is the *standing wave*, which has the appearance of oscillating "in place" as if it were are standing still (hence the name). Standing waves are also set into motion by some vibrational source, and the reason they appear to stand still is that they are confined to reflect back and forth between very well defined boundaries within a medium. Standing waves can exist in the air enclosed within an instrument, along the physical body of the instrument, or both. One of the most common examples of a standing wave is that which is supported by a vibrating string. On a guitar, the ends of the vibrational portion of the strings occur at the nut (located at the top of the fingerboard) and the bridge (below where the strings are strummed). The string can vibrate with positive amplitude anywhere on the string except at the ends where it is tied down.

Imagine a guitarist who plucks one of the horizontal guitar strings by pulling it up and away from equilibrium and then letting go. This initial disturbance in the string then propagates in *both directions* along its length with a speed dependent on the string's linear mass density and tension. When each wave arrives at an end (one at the nut and the other at the bridge), it reflects from the boundary point and reverses direction. The initial pluck from the guitarist's finger produces *two* waves that move in opposite directions and reflect back and forth between the two ends. The complex interplay between the two waves results in what *appears to be a single wave* standing in place and oscillating vertically. We don't see the individual left and right moving waves, but only the standing wave, which is the summation, or *superposition*, of the two waves. We will take up the topic of wave superposition in chapter 6.

We can summarize the role of standing and traveling waves in the production of musical sound by the following two observations:

- Vibrations produced by musical instruments result in a family of **standing waves** in the body and/or in the air enclosed by the instrument, and
- These standing wave vibrations generate **traveling waves** that propagate through the air to fill the surrounding environment.

Our task now is to better understand how instruments with well-defined boundaries produce a family of standing wave frequencies. These frequencies will be called either *harmonics* or *overtones*, depending on the mathematical relationship between members of the family. We will consider a few basic physical systems upon which many musical instruments are based. We will start with the vibration of *taut strings*, which will lay the foundation for understanding the stringed family of instruments. We will then consider vibrating *air columns*, the basis of the woodwind and brass family instruments, and conclude with an analysis of *stretched membranes*, upon which the percussion instruments are based.

5.3 Vibrating String

Any point in an instrument which is constrained to behave in a certain "fixed" way corresponds to a *boundary point*. In the case of a taut string, the two ends are boundary points, and the specific *boundary condition* for those points is that the string is constrained *not to vibrate* at these ends. Any other portion of the string is free to vibrate, and as a result several possible vibration patterns are possible.

Consider the arrangement in figure 5.1. The string is under some tension T due to the hanging weight. The two ends of the string (the left end tied to the wall and the right end in contact with the pulley) are constrained not to vibrate, and each therefore corresponds to a displacement *node* of vibration – a point on the string where the displacement amplitude is always zero, at rest. The dotted lines depict the shape of the string at the points of maximum amplitude in its up-down motion. This simplest mode of vibration on the string is called the *first harmonic* of vibration, and has the longest wavelength of all the possible vibration modes that satisfy the boundary conditions. For this reason it is also referred to as the *fundamental* mode of vibration. The string shape corresponds to 1/2 of a full sine wave, meaning that the wavelength for this mode of vibration (i.e. the length of one full wave cycle) is equal to twice the string's length (see figure 5.2):

$$\lambda_1 = 2L. \tag{5.1}$$

Note that it is impossible for any vibration with a *longer* wavelength to satisfy the boundary conditions. This *fundamental* mode of vibration is identified by the subscript "1." The only nodes for this wave exist at the ends – all other portions of the string vibrate with non-zero amplitude. The point at the exact middle vibrates with maximum amplitude, and therefore corresponds to a displacement *antinode* – a point where the string moves with maximum amplitude.

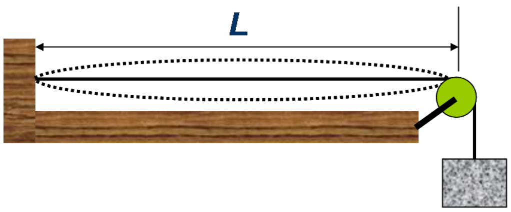


Figure 5.1: Fundamental vibrational mode for a string under tension. The dotted lines represent the vibrating string, each of which is 1/2 of a full sine-wave cycle (see figure 5.2). Therefore the wavelength of the standing wave, i.e. the length of one full cycle, is twice the length L between the ends.

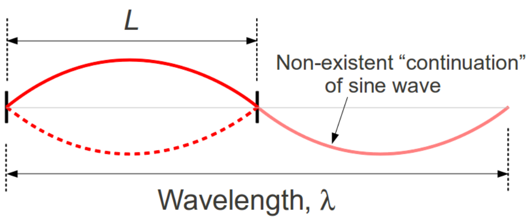


Figure 5.2: Demonstration of how the fundamental mode of vibration on a string corresponds to 1/2 of a full sine wave. Therefore the wavelength of the fundamental is twice the length of the string, $\lambda_1 = L/2$

Now that we've been able to determine the longest wavelength that can exist on this string, how do we determine its frequency? We've learned that the frequency and wavelength of a wave are related to one another by the speed with which it moves along the string, $v = f\lambda$.

In order to calculate the frequencies produced by the string, we first need to calculate the wave speed. The physics involved in the derivation of the speed of a transverse wave along a string is not discussed here, but it turns out to be simply related to the *tension* T in the string and its *mass density* μ:

$$\boxed{v = \sqrt{\frac{T}{\mu}}}. \tag{5.2}$$

The force of tension T is measured is N, and the mass density μ in kg/m. The mass density is defined as the mass contained in one unit of length. In this case, it corresponds to the mass of 1 meter of string length. The concept can be related to an example from everyday life, that of the price per unit length. If you were considering buying some ribbon at the craft store, and the sign said "$0.18/ft," you would know that each foot of material you wished to purchase would cost 18 cents. So if you wanted to purchase 6 feet of ribbon, the total cost would be 6 × $0.18 = $1.08. Likewise, the mass per unit length for a string allows us to calculate the total mass of a piece of string, given its length. Note that in determining the speed of a wave on this string, it's *not the total mass* of the string that matters, but the mass *density*.

Example 5.1

Linear Mass Density of a String *We are given a string of total length 350. cm and a total mass of 30.0 g. What is the linear mass density of this string in MKS units?*

Solution: First we wish to convert all quantities to MKS.

$$350.\,\cancel{\text{cm}}\cdot\left(\frac{1\text{ m}}{100\,\cancel{\text{cm}}}\right)=3.50\text{ m}\qquad 30.0\cancel{\text{g}}\cdot\left(\frac{\cancel{\text{g}}}{1000\cancel{\text{mg}}}\right)\cdot\left(\frac{1\text{ kg}}{1000\,\cancel{\text{g}}}\right)=3.00\times10^{-2}\text{ kg}$$

We can then compute the linear mass density of the string.

$$\mu=\frac{m}{L}=\frac{3.00\times10^{-2}\text{ kg}}{3.50\text{ m}}=8.57\times10^{-3}\text{ kg/m}$$

The fundamental frequency at which this string will vibrate (denoted by f_1 since it is the fundamental, and remembering that $v=f\lambda$), is given by

$$f_1\lambda_1=v\quad\rightarrow\quad f_1=\frac{v}{\lambda_1}=\frac{1}{\lambda_1}\sqrt{\frac{T}{\mu}}=\frac{1}{2L}\sqrt{\frac{T}{\mu}}.\tag{5.3}$$

5.3.1 Quantization of Wavelengths and Frequencies

In addition to this fundamental mode of vibration, there are many more vibrational patterns that this string supports. The common requirement for all of them is that they abide by boundary condition requirements at each end. The first three modes of vibration are depicted in figure 5.3. Note that each has nodes at the two ends (as required), but that they have a different number of nodes along the string length. The first mode (fundamental) has no nodes along its length between the two ends, and its wavelength is twice the string length, $\lambda_1=2L$. The second mode has one additional node at the center, and its wavelength is equal to that of the string, $\lambda_2=L$ (since one complete wave cycle fits between the two ends). Note that there are *no* vibrational pattern solutions with wavelengths *in between* those of the first and second harmonics, *i.e.* with values anywhere between L and $2L$. The third mode has two nodes, in addition to the two at the ends, and its wavelength is $\lambda_3=\frac{3}{2}L$.

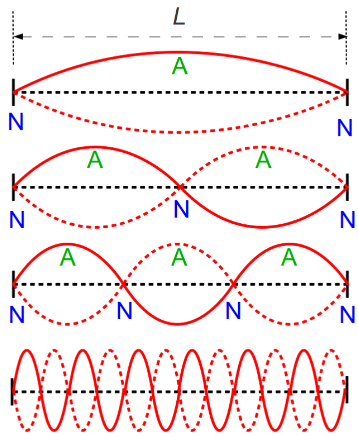
The set of possible vibrational patterns supported by the string is *quantized* in nature, only able to adopt a discrete set of possible wavelengths. This quantization property comes from the boundary condition that requires that the string have zero vibrational amplitude at each end. As we move from the fundamental mode to the higher modes of vibration, each additional increase in the integer n involves an increase of one additional node and one additional antinode, compared with the previous mode. Notice the pattern: for the n^{th} mode of vibration (depicted at the bottom of figure 5.3), the total number of antinodes is n, and the total number of nodes is $n+1$.

Given that the wavelength of the second mode of vibration is $\lambda_2=L$, frequency f_2 for this mode of vibration is then

$$f_2\lambda_2=v=\sqrt{\frac{T}{\mu}}\quad\rightarrow\quad f_2=\frac{1}{\lambda_2}\sqrt{\frac{T}{\mu}}=\frac{1}{L}\sqrt{\frac{T}{\mu}}.\tag{5.4}$$

For the third mode of vibration the wavelength is $\lambda=2/3L$ and the frequency is

$$f_3=\frac{3}{2L}\sqrt{\frac{T}{\mu}}.\tag{5.5}$$



$$n = 1: \quad \lambda_1 = 2L/1 \quad f_1 = \frac{1}{2L}\sqrt{\frac{T}{\mu}}$$

$$n = 2: \quad \lambda_2 = 2L/2 \quad f_2 = \frac{2}{2L}\sqrt{\frac{T}{\mu}}$$

$$n = 3: \quad \lambda_3 = 2L/3 \quad f_3 = \frac{3}{2L}\sqrt{\frac{T}{\mu}}$$

$$n: \quad \lambda_n = 2L/n \quad f_n = \frac{n}{2L}\sqrt{\frac{T}{\mu}}$$

Figure 5.3: Several vibrational modes for a stretched string, indexed by the integer $n = 1, 2, 3 \ldots$; $n = 1$ corresponds to the first harmonic, or fundamental, with two displacement nodes and one antinode. Each successive wave in the sequence ($n = 2$, second harmonic, $n = 3$, third harmonic, *etc.*) adds one additional node and antinode. Thus for the n^{th} vibrational mode the number of antinodes is n and the number of nodes is $n + 1$.

Looking carefully at this trend, you should be able to identify a general formula for the quantized values for the wavelength and frequency of the n^{th} mode of vibration on the string:

$$\boxed{\lambda_n = \frac{2L}{n} \qquad\qquad f_n = \frac{n}{2L}\sqrt{\frac{T}{\mu}}} \tag{5.6}$$

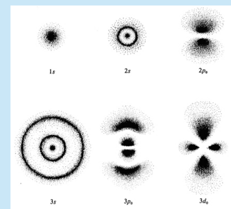
where L is the length of the string and n is any positive integer, $n = 1, 2, 3 \ldots$.

When the string is plucked, several modes of vibration are simultaneously set into motion. Most often the first, or fundamental mode has the highest amplitude. But it is important to note that the higher harmonics are present as well even though the ear does not hear them as separate tones (we'll discuss *why* in a future chapter). According to equation 5.6, all of the modes of vibration have wavelengths and frequencies that are mathematically related to one another. All higher mode frequencies are integer, or whole number, multiples of the fundamental frequency, since n only takes on integer values. This relationship between the higher modes and the fundamental defines the entire family of frequencies as *harmonics*.

On a related note . . .

Quantization in Atoms

The concept of the resonant modes of vibration for strings is helpful for understanding the quantum mechanical description of matter on the atomic scale. Electrons in the atom are constrained by boundary conditions (a little more complicated than those of the one-dimensional taut string!) since they are bound to a point nucleus by electrostatic attraction in three dimensional space. The shapes of the orbitals for the various states in the atom are the direct results of these boundary conditions.



5.3.2 Harmonics and Overtones

Some oscillating systems have very few natural frequencies, such as a pendulum (it has only one). The frequency at which it oscillates depends on its geometry (length of the string) and material makeup (mass of the bob influenced by the force of gravity). More complex systems, including most musical instruments, are able to vibrate at several natural frequencies simultaneously, even when they are playing a single note. For instruments we identify as *tonal*, that is, ones which produce identifiable pitches, the relationship between the members of the family of frequencies can be expressed in a particularly simple mathematical form, and they are referred to as *harmonics*. Other instruments with less recognizable or no tonal output (drums, for example), typically produce families of frequencies with much more complex relationships, and are referred to as *overtones*.

The sound that musical instruments make falls generally into two categories. There is, however, some overlap between the two. Tonal instruments (*i.e.* ones for which a clear pitch is identifiable) are *harmonic* instruments. The harmonics (as we will see in a later chapter) serve to support and emphasize the vibration of the fundamental tone, and as a result we hear a clear tonal center for the sound. Non-tonal instruments, ones for which a pitch is not so simple to identify, or may not exist at all (*e.g.* most drums) create sound consisting of a family of *overtones*, where there is little or no support for the fundamental tone, the result of which is an absent or ambiguous tonal center for the sound.

> **Harmonics**: simultaneous resonance vibrations whose frequencies are *integral multiples* of the fundamental, or lowest, frequency. Instruments that produce harmonics in their output we will refer to as *tonal* instruments.

> **Overtones**: simultaneous resonance vibrations whose frequencies are *non-integral multiples* of the fundamental. Instruments that produce predominantly overtones in their output we will often refer to as *non-tonal* or *atonal* instruments.

Even though we draw these clear distinctions between different types of instruments, we will soon see that some instruments which produce very clear, unambiguous pitches (the piano is a good example) nonetheless technically produce overtones that are close to being harmonics, but which are not exactly integral multiples of the fundamental. Tubular bells and tympani drums are examples of instruments for which a pitch is also pretty clearly identifiable, but which produce strong overtone series in their sound output.

A look ahead . . .

Timbre

The geometry and material construction of an instrument not only determine the various fundamental frequencies at which it operates (which we associate with the pitches it produces), but they also determine the "personality" of the instrument, *i.e.* that which gives the instrument its unique and recognizable "voice," or *timbre.* The latter arises from the particular family of frequencies that the instrument produces in addition to the fundamental. When a tonal instrument plays a single note, it fills the surrounding air with a family of frequencies corresponding to that one note, consisting of a fundamental (the lowest frequency) and a set of harmonics (higher frequencies mathematically related to the fundamental). Remarkably, as we shall later learn, our ear/brain system does not perceive these extra frequencies as additional tones, but rather perceives the entire family of frequencies as a single pitch (with a value corresponding to the fundamental), and with timbre (the personality produced by accompanying harmonics). We'll study this feature of perception in detail in later chapters.

Example 5.2

Waves on a String *We are given a string with a mass density of 0.00065 kg/m. It is stretched between two points a distance 0.85 m apart, with a tension of 90 N (about 20 lbs). What are the wavelengths of the first three harmonics for this string? What are their corresponding frequencies?*

Solution: We first identify the basic equation for the harmonics on a string:

$$f_n = \frac{n}{2L}\sqrt{\frac{T}{\mu}} = nf_1.$$

For the first harmonic, the wavelength is twice the length of the string, since 1/2 of a full sine wave exists between the two ends. The second and third harmonics have wavelengths of L and $2/3L$, respectively (see figure 5.3). Thus

$$\lambda_1 = 2\cdot(0.85\text{ m}) = 1.70\text{ m} \qquad \lambda_2 = 0.85\text{ m} \qquad \lambda_3 = \frac{2}{3}(0.85\text{ m}) = 0.57\text{ m}$$

Using the values we are given, the frequency of the fundamental is

$$f_1 = \frac{1}{2\,(0.85\text{ m})}\sqrt{\frac{90\text{ N}}{0.00065\text{ kg/m}}} = 219\text{ Hz}$$

and then the higher harmonics are related to the fundamental by

$$f_n = n\cdot f_1$$

so that the first three harmonics for this string have frequencies 219 Hz, 438 Hz, and 657 Hz, respectively.

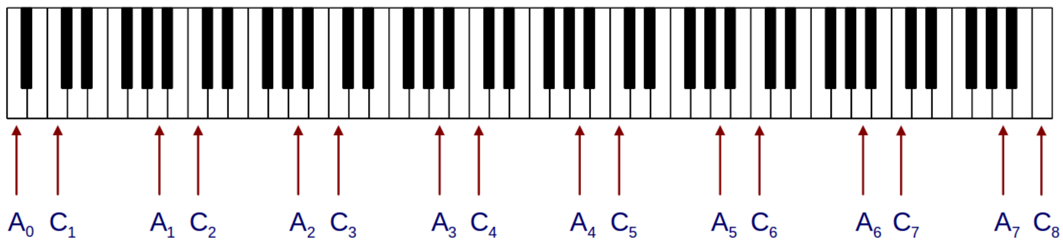


Figure 5.4: The full piano keyboard. The lowest note is denoted "A_0"
.

At this point it is worth pausing and thinking about our results thus far. Our equation for the frequencies of various harmonics depends on both tension and mass density of the string. It suggests that the frequency should be proportional to the square root of the tension, and inversely proportional to the square root of the mass density. Strings on the guitar, as well as for the violin and other stringed instruments, are thicker for the lower frequency range strings. Is this consistent with our predictions from the equation? Likewise, if we tighten the string, adjusting it to higher tension, the frequency should go up. Is this in accord with your experience? You should try this yourself ... find a guitar somewhere and experiment with different string thicknesses and with adjusting them to different tensions to see if these predictions are correct.

On a related note . . .

Harmonics Experiment:

When a string is plucked, several of its harmonics are set into motion. Even though our ear hears one pitch, associated with the frequency of the fundamental, several other harmonics are present in vibration as well, and our ear hears these extra harmonics as contributing to the timbre of the instrument. In order to find out whether the higher harmonics are in fact present in the string vibration, we can conduct a little experiment. Find a piano somewhere and do the following. Find the C_2 key (see figure C.2) and depress it slowly, so that it doesn't make any sound. This will lift the damper from the string leaving it free to vibrate. While holding this key down, strike the C_3 key sharply and then let go. After the sound from C_3 dies away, you should be able to hear its pitch continue, but now it will be coming from the second harmonic of the C_2 string. The harmonic has been set into vibration by sympathetic resonance, *i.e.* the vibrations in the air have set it into motion.

Other harmonics on this string can also be set into motion. While holding the C_2 key down, briefly strike the G_3, C_4, E_4 and G_4 keys, each of which corresponds to the next higher harmonics. Keys which do not correspond to harmonics of the C_2 string will not produce the same effect. Try, for example, doing the same thing with the D_3 or F_3 keys and you should hear nothing. One final test you can conduct is to set several of its harmonics into motion all at once. While holding the C_2 key down, depress briefly as many keys as you can with your forearm loudly and quickly, perhaps using two or three arm motions. What lingers afterwards should be a full dominant seventh chord, consisting of C_3, G_3, C_4, E_4, G_4, $B_4^\flat$, and C_5. Brief note: this will NOT work with a digital piano. Why?

5.4 Vibrating Air Column

A column of air also vibrates with a family of frequencies that includes a fundamental and higher harmonics (or overtones, depending on the geometry of the column). In order to understand the resonances that can exist in tubes, we will first, as we did with strings, determine what *allowed wavelengths* can exist, constrained by boundary conditions. We will then calculate the allowed frequencies that the tube can support.

5.4.1 Standing Sound Waves

In order to understand standing sound waves in tubes, it helps to compare and contrast them with waves on a string. In the case of a string, the material medium, i.e. the string, moves in a direction perpendicular to the direction of wave motion (*i.e.* a *transverse* wave). The string moves a small distance in the vertical direction, while the wave moves a great distance horizontally along the length of the string and bounces back and forth between the tied-down ends.

For sound waves, the medium is air, and the waves are *longitudinal*. This means that the molecules of air making up the medium in which the wave propagates move back and forth along the same direction as the wave motion. As a wave passes through, each air molecule moves a very small distance in the forward-backward direction in periodic motion. When we drew the sinusoidal waveforms for string vibration resonances, we naturally associated the amplitude of the waveform with the vertical position of the string, *i.e.* transverse to the wave motion. Displacement nodes for the string vibration are identified at the regions where the string stays still, and antinodes at regions where it moves with maximum displacement. We would also like to represent standing sound waves graphically, identifying nodes and antinodes, and developing a mathematical nature for the resonant frequencies supported by the geometry of the system and its boundary conditions.

We have basically two choices for how to characterize a standing sound wave in order to identify the location of nodes and antinodes. We can either choose to focus on the *motion*, or displacement of the air molecules from their equilibrium positions (analogously to the way we treated the string), or we can focus on the *pressure* of the air at various regions along the standing wave. This second option involves a little more abstraction in the way we picture the waves. But it's simpler than it may sound. For any standing sound wave, there are always certain fixed locations where the air pressure remains at a constant value (its equilibrium value of atmospheric pressure), which we will refer to as pressure nodes. As well, there are also fixed locations where the air pressure varies regularly from a maximum to a minimum pressure, which we will refer to as pressure anti-nodes.

Take a look at figure 5.5 to help visualize a pressure node and an antinode in a standing sound wave. Points where pressure nodes occur correspond in space to the places of motion *antinodes*, and points where pressure antinodes occur correspond to motion *nodes*. If, for example, you are a molecule sitting at the location of a pressure *node* (left side of figure), the pressure at your location remains the same (meaning that the average spacing between you and your neighbor molecules remains constant). You and your neighbors move as a group to the left and to the right in periodic motion, while the pressure remains at atmospheric pressure. If, on the other hand, you are a sitting at the exact center of a pressure *antinode* (right side of the figure), the pressure at that location varies from the maximum to the minimum value, as a result of your neighbors to the left and right moving first away from you (creating low pressure), and then back toward you (creating high pressure at your location) in periodic fashion – while you stay in place during the entire cycle. You are therefore situated at a pressure antinode and a position node. (But maybe you already node that.)

One of the benefits of viewing standing waves from the pressure point of view is that the pressure

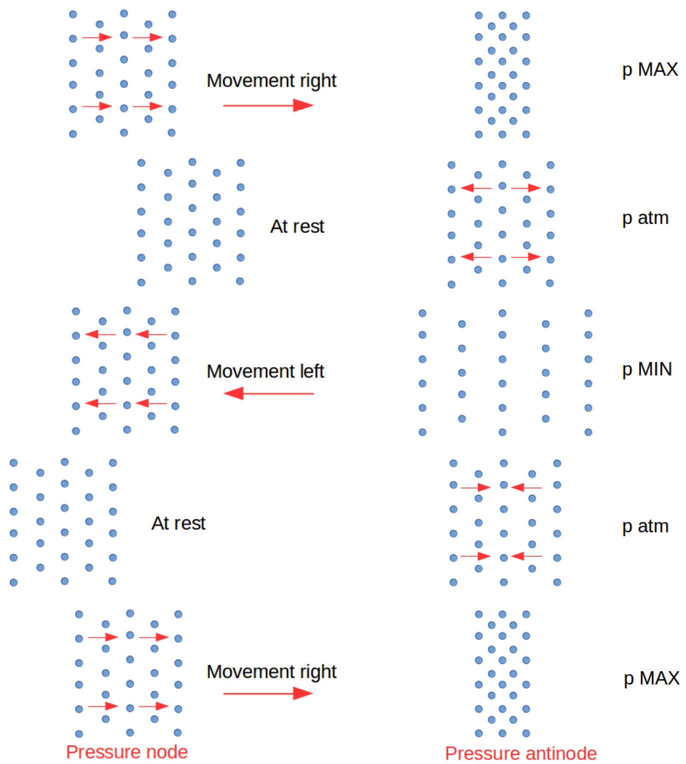


Figure 5.5: Depiction of a pressure node (left) and a pressure antinode (right) in a standing sound wave. The dots are air molecules, and the 5 vertical frames represent one complete cycle of the standing wave evolution. Note in the pressure node how the spacing between molecules (corresponding to a pressure of 1 atmosphere) remains constant through the complete cycle, but the motion of the molecules left and right is maximal (motion antinode). In the pressure antinode the pressure varies from maximum to minimum and back to maximum through the complete cycle, and the molecules at the very center of the zone remain at rest (motion node).

standing wave patterns for vibrating air in an "open-open" cylindrical tube (one that is open at both ends) look similar to those for the vibrating string. Let's now turn to vibrational resonances in a cylindrical tube.

5.4.2 Open-open tube

Consider first the case for an open-open cylindrical tube (we will often refer to this type of tube as simply "open"). This tube supports standing sound waves that include a pressure node at each end (meaning that the air pressure at these ends always remains at 1 atm). The fundamental mode of vibration for this tube has a single pressure antinode at its center and a pressure node at each end. .

Figure 5.6 shows a few standing wave patterns that are supported by an open tube. Note that the wavelength of the fundamental (upper left in the figure) is related to the length of the tube by $\lambda_1 = 2L$, as was the case for the fundamental of the stretched string. Note also that the sequence of wavelengths relative to the tube length match those of the string, and therefore the harmonic structure of this tube should be similar to that of the string.

5.4.3 End Correction

There is one important difference between the open cylindrical tube and the taut string. The pressure node at the open end of a tube is not, in fact, located at the exact end of the tube, but extends slightly outside this point by a distance $0.613R$, where R is the tube's radius (see figure 5.7). The open end radius affects the exact location of the node: the larger the radius, the further out the node is located. The physics behind the placement of this node is rather complicated, but the formula used to determine its location is very straightforward. It is good to keep in mind that the distance

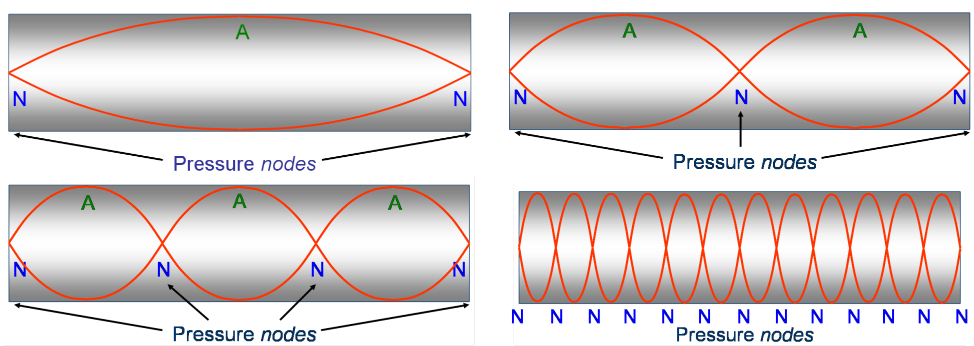


Figure 5.6: Harmonic series for sound standing waves in an open-open tube. Their geometric shape, including location of the nodes and antinodes are the same as for the standing waves on a stretched string. Here, the nodes and antinodes correspond to *pressure* variations of the enclosed air - nodes for where the pressure remains constant (that is, at atmospheric pressure), and antinodes where the pressure changes from the highest to the lowest values relative to atmospheric pressure.

between neighboring nodes determines the wavelength of vibration for a given resonance. For properly calculating the wavelengths of the resonances, then, we need to express the length of the tube not by its physical length, but by its *effective* length:

$$\boxed{\textbf{Open tube:} \qquad L_e = L + 1.226R} \tag{5.7}$$

where $1.226 = 2(0.613)$, since the correction applies at each end. This will be be the appropriate length for us to use in calculating the harmonic wavelengths and frequencies of the tube. For long, narrow tubes (i.e. small radius) the physical and effective lengths are very similar, but for short wide tubes the difference is more pronounced.

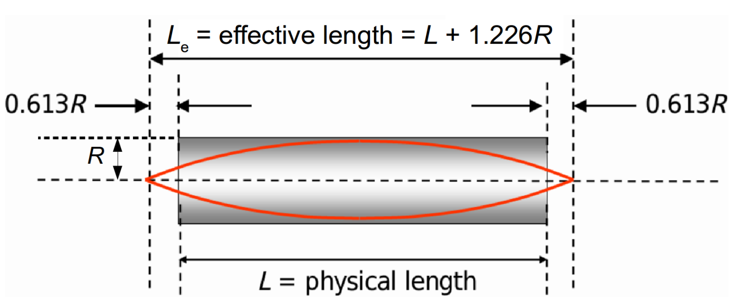


Figure 5.7: The nodes for an open tube are positioned slightly outside the physical ends of the tube, and are a function of the tube's radius. The *effective*, or corrected length is the proper length to use in order to calculate the resonant structure of the tube accurately.

The series of resonances that are supported by this open tube can be expressed in terms of the speed of sound and its effective length:

$$
\begin{aligned}
&n=1 \qquad \lambda_1 = 2L_e && f_1 = \left(\frac{1}{\lambda_1}\right) v_s = \left(\frac{1}{2L_e}\right) 20.1\sqrt{T_A} \\
&n=2 \qquad \lambda_2 = L_e && f_2 = \left(\frac{2}{2L_e}\right) 20.1\sqrt{T_A} \\
&n=3 \qquad \lambda_3 = 2L_e/3 && f_3 = \left(\frac{3}{2L_e}\right) 20.1\sqrt{T_A}
\end{aligned} \tag{5.8}
$$

where $n = 1, 2, 3\ldots$ Thus we can identify the trend that allows for calculating the n^{th} vibrational harmonic,

$$
\boxed{\textbf{Open tube:} \quad \lambda_n = 2(L + 1.226R)/n \qquad f_n = \left(\frac{n}{2(L+1.226R)}\right) 20.1\sqrt{T_A} \qquad n = 1, 2, 3, \ldots}. \tag{5.9}
$$

5.4.4 Semiclosed tube

A semiclosed tube is open only on one end and closed off on the other. The open end of the semiclosed tube corresponds to a pressure *node*, but the closed end corresponds to a pressure *antinode*. Figure 5.8 (upper left) shows that, with a node at the open end and an antinode at the closed end, the longest wavelength that satisfies these boundary condition has exactly $1/4^{th}$ of a complete wave inside the tube. This means that $\lambda_1 = 4L$, which leads us to conclude that the wavelength for the fundamental mode of vibration in the semiclosed tube is twice as long as that for an open tube (of equal length).

When we add the end correction for the location of the node outside the open end (only one!) of the semiclosed tube, we see that the effective length of the semiclosed tube is

$$
\boxed{\textbf{Semiclosed tube:} \qquad L_e = L + 0.613R} \tag{5.10}
$$

The series of resonances supported by this tube can be expressed in terms of its effective length and the speed of sound in air. Review figure 5.8 in order to verify that only an odd number of 1/4 wavelengths are supported as resonances in the tube:

$$
\begin{aligned}
&n=1 \qquad \lambda_1 = 4L_e && f_1 = \frac{1}{\lambda_1} v_s = \frac{1}{4L_e} 20.1\sqrt{T_A} \\
&n=3 \qquad \lambda_3 = 4L_e/3 && f_3 = \frac{3}{4L_e} 20.1\sqrt{T_A} \\
&n=5 \qquad \lambda_5 = 4L_e/5 && f_5 = \frac{5}{4L_e} 20.1\sqrt{T_A}
\end{aligned} \tag{5.11}
$$

where $n = 1, 3, 5, \ldots$, *i.e.* n *only takes on odd-integer values* for the semiclosed tube. Since only an odd number of 1/4 wavelengths can be supported in the tube, then only *odd values* of n are present in the harmonic spectrum. Thus we see the trend for calculating n^{th} harmonic:

$$\boxed{\textbf{Semiclosed tube:}\quad \lambda_n = 4(L+0.613R)/n \quad f_n = \left(\frac{n}{4(L+0.613R)}\right)20.1\sqrt{T_A} \quad n = 1,3,5,\ldots}. \tag{5.12}$$

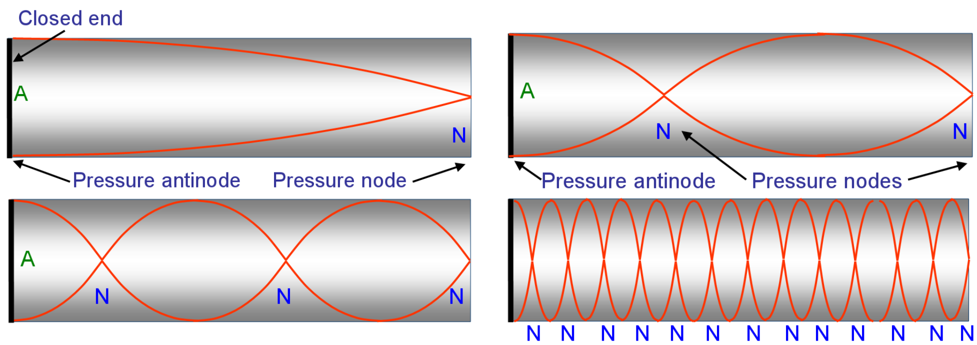


Figure 5.8: Harmonic series for sound standing waves in an semiclosed tube. The open end on the right always corresponds to a pressure node, and the closed end on the left always corresponds to a pressure antinode. The wavelengths for the various supported modes of vibration follow the pattern $\lambda_n = 4L/n$ where n = 1, 3, 5, 7...

5.4.5 Other Tube Shapes

So far, then, we see that open tubes support all even and odd integral harmonics, and that semiclosed tubes support only odd integral harmonics. The next question we can ask is, do any other fundamental tube shapes have musical value? Figure 5.9 shows three more possibilities in addition to the cylindrical. Of these, only the conical and the exponential tube types occur in musical instruments (for the most part). Of the woodwind instruments, the clarinet and bassoon are cylindrical in shape, whereas the oboe and the saxophone are conical bore instruments. The bells on the ends of several brass instruments are exponential-like in shape. We will consider the exponential bells on brass instruments in a later chapter, but let's consider briefly the conical shape and the resonances it supports.

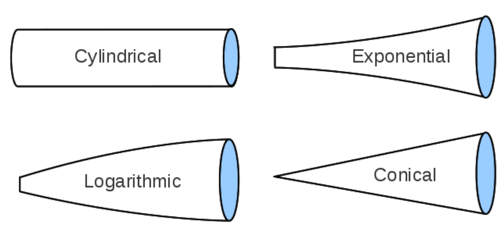


Figure 5.9: Four basic tube types. The cylindrical and conical varieties are used in woodwind instruments, and the exponential shape appears in the bells on the ends of brass instruments

5.4.6 Conical Tube

Since the pointed end of the conical tube is closed, we might assume that it acts like the closed end of a semiclosed tube, *i.e.* that its position would correspond to a pressure *antinode*. Interestingly, the opposite is true! Again, for physics reasons that are more complicated than we care to get into, the closed pointed end of a conical tube acts as if it is an open end, and therefore corresponds to a pressure *node*, just as if it were an open tube of the same length! The harmonic structure of a conical tube is the same as it is for a cylindrical open tube, and the harmonic series for this tube is therefore described by equation 5.9. Some wind instruments have internal tubes with conical shape (the oboe and the saxophone, for example), but none have a complete cone all the way to the point. They are *truncated*, meaning that the small end of the cone has been removed and closed off (see figure 5.10).

Interestingly, the same harmonic structure holds true for the conical tube even if we truncate the pointed end of the cone (see figure 5.10), as long as we don't remove more than about 25% of the original cone's length. The harmonic structure of this truncated cone is identical to the untruncated cone. If the length we remove exceeds 25% of the original cone's length, then the mathematical relationship between the higher frequencies is not so mathematically simple. In this case, the higher natural frequencies, or resonances, are *not* integer multiples of the fundamental frequency, and are therefore not harmonics, but instead form *overtones*, meaning that the higher resonances are non-integral multiples of the fundamental.

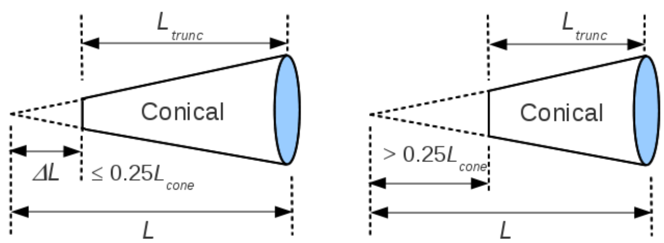


Figure 5.10: Truncated conical tubes. If less than 25% of the full cone length L_{cone} is cut off at the pointed end and closed off, the remaining truncated tube has a similar harmonic structure to an open cylindrical tube, as described by equation 5.9. If more than 25% of the full cone length is truncated, the relation of the upper resonances to the fundamental is no longer integral in relationship, and they are no longer harmonics but become *overtones*.

5.5 Vibrations in Membranes

A number of percussion instruments use a membrane (usually circular) stretched over a frame. The membrane is two-dimensional, meaning that the initial disturbance can propagate along the membrane in a variety of directions. Thus, unlike both the vibrating string and the vibrating air column which are one-dimensional, it is a much more complicated structure. Instead of producing harmonics, drumheads produce a family of overtones.

For a membrane of uniform thickness stretched over a circular ring, the values for the vibrational resonance frequencies are much more complicated to derive than for strings or tubes. Waves formed on the membrane surface can propagate in two dimensions (in any direction along the plane of the membrane). In this two-dimensional case, the boundary of the system is the continuous circumference where the membrane is tied down. All the way around this perimeter the membrane cannot vibrate – its amplitude remains zero.

Unlike the string and tube, because this system is two-dimensional, the nodes occur along *lines* (see figure 5.11). The nodal lines divide the membrane into symmetric parts that move in opposite directions. The entire circular rim must necessarily be a vibrational node, since that is where the membrane is physically tied down. The frequency for the fundamental mode of vibration is

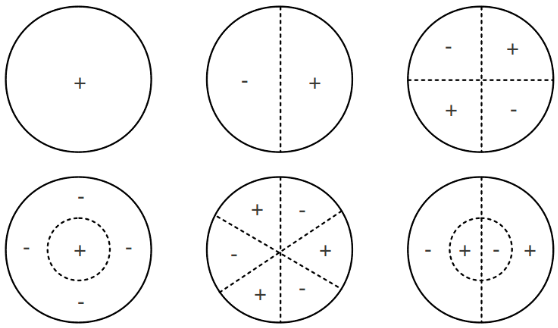


Figure 5.11: Schematic of several vibrational modes for a stretched membrane. Since the membrane is two-dimensional, the nodes occur along the dotted lines, and not at points. The + and - symbols indicate portions of the membrane that are moving up or down at the moment of the depiction of the vibrational mode.

$$\boxed{f_1 = \frac{0.383}{R}\sqrt{\frac{T}{\mu}}} \tag{5.13}$$

where T is the membrane tension (expressed in N), R is the membrane radius, and μ is its areal mass density (expressed in kg/m^2). Note that the vibrational frequencies increase with increasing tension and decrease with increasing membrane areal mass density. This formula for the fundamental frequency of the circular drumhead is fairly simple. Calculation of the frequencies of the higher modes of vibration becomes much more complicated, and so rather than show the formulas for those modes, we'll just list the frequencies of the vibrational modes shown in figure 5.11 for comparison. Each is characterized in relation to the fundamental mode, and they are arranged in two rows and three columns to correspond to the images in the figure.

$$\begin{array}{ccc} f_1 & f_2 = 1.59 f_1 & f_3 = 2.13 f_1 \\ \\ f_4 = 2.29 f_1 & f_5 = 2.65 f_1 & f_6 = 2.91 f_1 \end{array}$$

Whereas the string and air column instruments produce families of frequencies harmonically related to one another, the drumhead produces an overtone series. This gives drums a non-tonal character, meaning that the drum produces a sound with a very weak or non-existent sense of pitch. One notable exception to this is the tympani drum, which has a specially shaped cavity below the membrane that allows for resonances that partly support the fundamental and produce a recognizable pitch.

5.6 Chapter Summary

Key Notes

- The two features of musical instruments that determine the nature of sound they emit are the *geometry* and *material makeup*.
- Physical geometry determines the *boundary conditions* by which the allowed vibrational modes are dictated.

- The material (*i.e.* physical) makeup of the instrument determines the speed with which vibrational waves move throughout the instrument, determining which frequencies will emanate from the instrument given its physical geometry.

- Standing waves are produced in the instrument, which propagate out into the surrounding environment as sound waves.

- The sounds produced by string instruments originate in the vibration of a string. Since both ends are tied down, the string is forced to adopt vibration patterns with nodes at each of the ends. The wavelengths and frequencies able to be supported on this string are a function of the string length, the tension, and mass density.

- A column of air can also vibrate with frequencies related to the fundamental.

- A cylindrical tube with two open ends (open tube) supports standing waves that have a pressure node at each end (corresponding to points where the air pressure remains at a constant value). Note that the wavelength of the fundamental is related to the length of the tube by $\lambda_1 = 2L$. Note also that the sequence of possible wavelengths match those of the string, and therefore that the harmonic structure of this tube is similar to that of the string.

- The pressure node is not located at the tube's exact end, but rather slightly outside it, giving the instrument an *effective* length which we use in determining its resonant frequencies. The effective length of the open tube can be expressed as $L_e = L + 1.226R$ where R is the radius of the tube and L is the physical length of the tube. For a semiclosed tube $L_e = L + 0.613R$.

- A cylindrical tube with one end closed (semiclosed tube) has a different harmonic structure than the open tube. The fundamental mode of vibration for this tube has the relationship $\lambda_1 = 4L$. Only odd numbered harmonics, or odd integral multiples of the fundamental, exist for this tube geometry.

- The air contained within a perfect cone of length L vibrates with exactly the same frequencies as the open cylindrical tube of the same length. We can even cut the pointed end and truncate the cone by length l_0, making it blunt, without changing the harmonic structure, so long as l_0 is not more than about $1/4^{th}$ of the total cone length.

- If we cut more than $1/4L$ from the cone, the harmonic structure changes to one of overtones, *i.e.* non-integral multiples of the fundamental.

- For a membrane of uniform thickness stretched over a circular ring, the vibrational frequencies are much more complicated than for strings or tubes, since the waves formed on the membrane surface can propagate in two directions (i.e. along a plane). Since these vibrations occur in a plane, the corresponding "nodal points" for the strings and tubes become "nodal lines" on the drum surface. The nodal lines divide the membrane into symmetric parts that move in opposite directions to each other. The rim must necessarily be a vibrational node, since the membrane is tied down physically.

- The vibrational frequencies for the membrane constitute overtones since they do not have a simple mathematical relationship with one another, and increase with increasing tension and decrease with increasing membrane density . For this reason, most drums have much less of a distinctive "pitch" to their sound.

Exercises

Questions

1) Briefly describe the role of *geometry* and *material construction* in helping determine the resonant frequencies of an instrument.
2) What is a boundary condition for an instrument? What is its purpose?
3) What qualities of an instrument determine the speed with which waves propagate through it?
4) Explain briefly why vibrating strings naturally support certain frequencies and not others - in other words, explain why the vibrational solutions are *quantized*.
5) Why is a standing wave called "standing"? What is actually going on in a standing wave as far as wave propagation?
6) What two specific (*i.e.* not general) features of a string system determine the speed with which a wave moves along it?
7) Can a string support the vibration of several natural frequencies at once, or only one at a time? Explain briefly.
8) Describe the difference in the family of frequencies that an open tube and a semiclosed tube can support. Why does this difference occur?
9) Why is it important to take into account the end correction for a tube when calculating the resonance frequencies it supports?
10) Why is it that the semiclosed tube has a smaller end correction than an open tube?
11) What makes vibrating membranes more complicated than vibrating strings or air columns? Is there a difference in the character of vibrational modes it supports? Explain briefly.

Problems

1. One of the strings of a piano is 96 cm in length. If the string is vibrating in its sixth vibrational mode,
   a) Draw the shape of the standing wave. Indicate the nodes and antinodes for this mode of vibration.
   b) How many antinodes exist along the string?
   c) What is the wavelength of the standing wave on the string?
2. Consider a violin string of length 0.316 m (31.6 cm). Waves travel on this string with velocity of 277 m/s.
   a) What is the largest period a wave can have if it is to be accommodated by the string as a standing wave?
   b) Waves of other periods can also exist as standing waves on the string. What are the next two largest periods?
   c) If the string has a linear density of 0.00065 kg/m, with what tension is it stretched?
3. A string has a linear density of 0.035 kg/m. If it is stretched to a tension of 667 N, what is the velocity of the wave on the string?
4. Consider the string in the previous question. If the string is 1.0 m long, what is the frequency of the fourth resonant mode?
5. A harp string is 52.0 cm long and tuned to a frequency of 660 Hz. Find the wavelength
   a) of the fifth harmonic on the string
   b) of the sound waves in the air (22°C) that result from the fifth harmonic.
6. A stretched string vibrates with a frequency of 440 Hz (A_4 on the piano). If the tension in the string is doubled, what is the new pitch?

7. A nylon guitar string has a mass per unit length of 8.3×10^{-4} kg/m and the tension is 56 N. Find the speed of tranvserse waves on the string.

8. The distance from the bridge to the nut on the nylon string guitar in the previous problem is 63 cm. If the string is plucked at the center,

   a) how long will it take the pulse to travel to either end and return to the center? You can use the speed you calculated in the previous problem.

   b) Calculate the frequencies of the first four harmonics (including the fundamental) for this string.

9. An open cylindrical tube is 0.05 m long. If you ignore the end correction factor, what is the frequency of the third harmonic? Assume a temperature of 10°C.

10. Two open organ pipes have lengths of 0.60 and 0.64 m. These two pipes are sounded together. If we assume the speed of sound is 345 m/s, what is perceived?

11. An open cylindrical tube has a radius of 3 cm. What is the end correction?

12. Find the difference in the fundamental frequency, calculated with and without the end correction, of an open organ pipe 2.0 m long and 10 cm in diameter.

13. A semiclosed cylindrical tube has a length of 0.5 m and a radius of 6 cm. What is the frequency of this tube's first harmonic?

14. You are given an instrument that has strings of length 76 cm. Part of the string is suspended over a fingerboard. When the string is pressed to the fingerboard:

    a) What happens? Explain briefly.

    b) When the string is fingered a distance of 8cm from the end, the string vibrates with a frequency of 300 Hz. Where would the finger have to be placed to obtain a vibration frequency of 311 Hz?

Chapter 6

Superposition

6.1 Complex Waveforms in Music

As we learned in chapter 5, when a musical instrument produces a single note, it actually produces a family of frequencies consisting of a fundamental frequency plus a series of harmonics whose frequencies are integral multiples of the fundamental. Our ears hear the tone as a single pitch with personality, or *timbre*. The unique timbre of the instrument derives from the specific collection of harmonics with differing amplitudes, relative to the fundamental. When two different instruments play the same musical note, their fundamental frequencies are identical, but the specific amplitudes of each of their higher harmonics, relative to the fundamental, differ between the two, giving rise to different timbres. In this chapter we will learn that the individual harmonics of an instrument add together to produce a complex, non-sinusoidal waveform that moves out from the instrument as sound.

It is important to emphasize that each of the harmonics that an instrument produces is individually simple harmonic in character. It is the *combination* of them all playing simultaneously, when added together, that gives rise to the complex periodic waveform that is emitted from the instrument. Our goal in this chapter is to understand these complex periodic waveforms, how they are produced and perceived, and how to represent them graphically. In order to do this we'll need to understand principle of *superposition*.

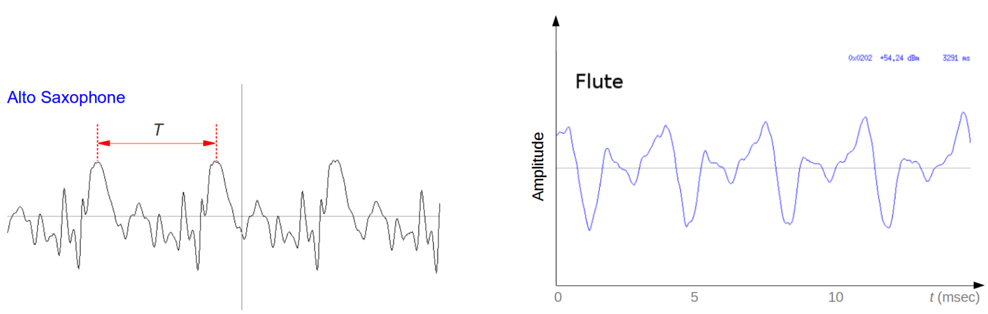


Figure 6.1: Waveform of an alto-saxophone (left) and a flute (right). Note that both of these instruments have a clearly defined periodicity, which gives rise to the strong sense of pitch associated with these instruments. The waveforms are complex, each consisting of a fundamental frequency accompanied by additional harmonics. The superposition of all the harmonics for a particular instrument, each of which is sinusoidal, results in a complex wave.

Superposition is a process of "adding" sinusoidal waves or vibrations together. For our purposes, this process of adding will be primarily *graphical* (though there are mathematical methods to do so that are beyond the scope of this course). Since real physical instruments put out sound that is much more complex than sinusoidal (see figure 6.1), it is therefore of interest to us to know how these complex waveforms can be "broken down," or "deconstructed" into the individual sinusoidal harmonic components that add together to give rise to them. The specific collection of frequencies and amplitudes that the instrument produces forms a unique sort of "fingerprint" for that instrument, making it different from all other instruments. When we hear the same note played by a saxophone and by a flute, the personality or tonal quality of each is quite different from the other, enabling us to tell them apart with our eyes closed. Therefore it will be of great interest to us to better understand how these complex waveforms can be represented by their individual sinusoidal component frequencies and amplitudes to better understand the origin of *timbre*.

There is a powerful mathematical theorem that says that *any* complicated periodic waveform can be constructed from a set of simple sinusoidal waves of differing frequencies and amplitudes. The exact combination of frequencies and amplitudes one needs to add together to produce a particular complex waveform is calculable using a technique called *Fourier analysis*, but this method is beyond the scope of our discussion. We do want to understand, however, that no matter how complex, any waveform can in principle be represented by a sum of individual sinusoidal waves of differing frequencies and amplitudes.

6.2 Simultaneous Vibrations

Superposition consists of adding the amplitudes of two simultaneously played waveforms at each moment in time. In order to see how superposition works, let's first consider the situation in figure 6.2, where two pulse waves approach each other from opposite directions. When they overlap with each other at the center of the string, their amplitudes add together to yield a result that is different than each of them. In the left figure (where both amplitudes are positive), the two pulses add *constructively* at the exact center, so that the resulting pulse is instantaneously twice the amplitude of each. In the right figure (where the amplitudes have opposite sign) they add *destructively*, instantaneously canceling each other out at the exact center. Note that after passing through one another they continue on their way unaffected by the brief encounter.

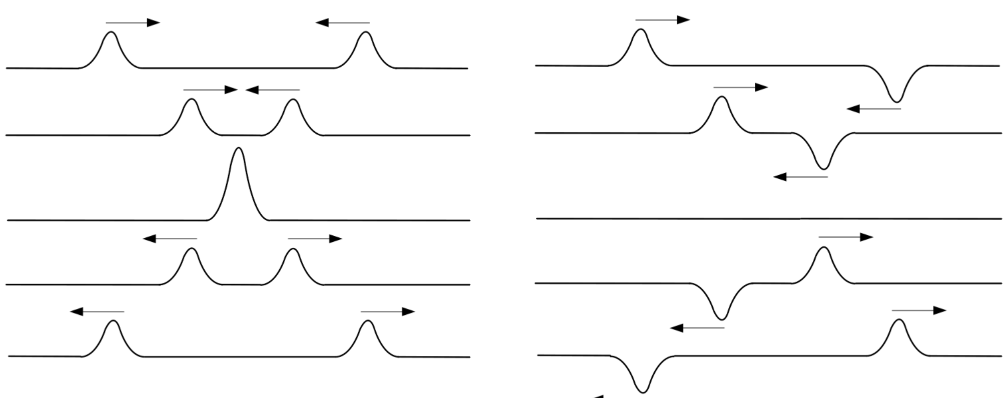


Figure 6.2: Two pulses on a string approaching each other. When they overlap, their amplitudes add. On the left the pulses both have positive amplitudes and therefore add to give a resulting waveform with twice the amplitude of the individual pulses. On the right, the pulses have opposite sign amplitude and therefore temporarily cancel one another when they overlap (yielding zero amplitude). After they pass through one another, they continue on just as they did before the encounter.

6.2.1 Physical Motion of the String

Note in particular that when the pulses are at the exact center, the string itself can only move in one direction at a time even though two waves are passing through one another. This is a very important point, since *all* vibrating systems that oscillate in a complex way share this feature. Consider, for example, the string of a guitar as a single note is played. The string vibrates with a fundamental frequency (associated with its pitch) accompanied by a family of harmonics (associated with the timbre of the instrument). Hence all of the harmonics are "present" in the sound output. And yet at any particular instant in time, a small portion of the string can only move in one direction at a time, either up or down, so that it is physically unable to vibrate with each individual harmonic separately. The string moves as the *superposition* of all the harmonics, *i.e.* in a complex fashion that represents the sum of all the harmonics at once. If we were to watch the string in slow motion, it would move up and down in a complex, but decidedly *non*-sinusoidal fashion (see figure 6.3). The harmonics are really "present" in the tone, in the sense that they all add to give the result, but can't be individually seen in the motion of the string.

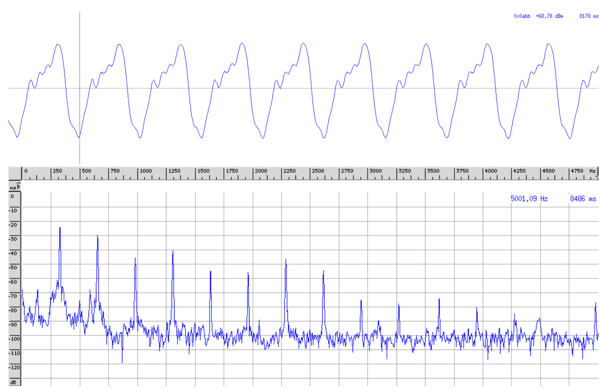


Figure 6.3: Waveform for a nylon guitar playing a single note. While the waveform has strong periodicity, it certainly is not sinusoidal. The complex waveform shape is the superposition of all of the harmonics that are produced by the string and guitar body as they vibrate together. Below is its frequency spectrum.

Figure 6.3 shows the complex tone of a guitar playing a single note. The complex waveform (in the upper portion of the figure) corresponds to the specific way in which the string *actually moves* in the up and down direction. This motion is clearly complex, and has a well-defined periodicity (corresponding to the pitch). The vibrations from this string are communicated to the body of the instrument, which subsequently takes up this complex vibrational motion and communicates it to the surrounding air. The air molecules then also move in this complex fashion as they communicate the vibrations to their neighbors as the waveform propagates through the air. When the sound enters the ear, the eardrum also vibrates with this same complex motion. The particles of any medium, be it the string, the instrument body, the air molecules, or the eardrum, move in the same complex fashion as the original sound that is produced, and therefore carry the "personality" of the guitar from the source to the receiver.

6.2.2 Phase

The concept of *phase* enables us to characterize the location of an oscillating body within its periodic cycle. Consider the mass hanging on a spring in figure 3.6. We can represent its oscillation graphically using a sine wave. When the mass m moves in the upward direction and passes through its equilibrium point (the fourth mass-spring image in that figure), we characterize its phase as 0°, where the sine wave crosses the axis in an upward moving direction. A phase of 90°corresponds to 1/4 the way into its complete cycle, where the mass is at the maximum height of its oscillation (the fifth mass-spring image). A phase of 270°corresponds to the spring being 3/4 the way through its complete cycle, where the mass is at the bottom of its travel, *etc.* For a correspondence between angles on the unit circle and a sinusoidal wave see figure 6.4.

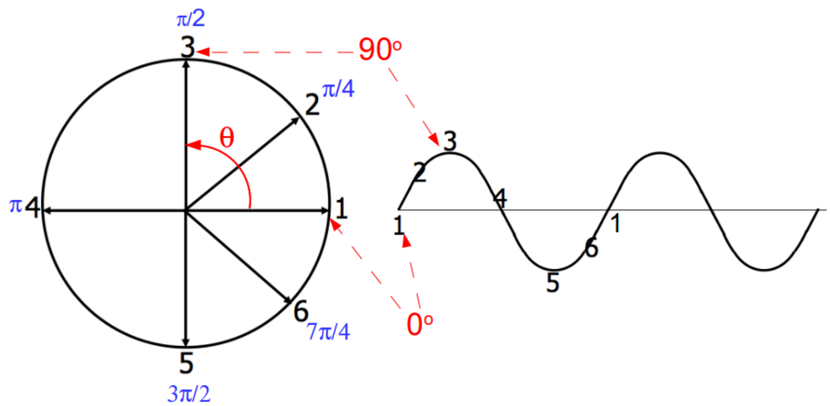


Figure 6.4: (***For color version see Appendix E***) The unit circle with points identified on its perimeter corresponding to certain phases. The corresponding points are shown on the sine curve to indicate phase and its relationship to points along the cycle of an oscillatory system. The circumference of the circle is $2\pi r$, and once around corresponds to a total angular journey of 2π radians = 360°.

When two identical systems are oscillating simultaneously with the same period (for example, two identical masses hanging on two identical springs), they can be set into motion such that they have a *phase difference* between them. Figure 6.5 shows three situations: the first depicts two oscillators differing in phase by 45°, the second by 90°, and the third by 180°.

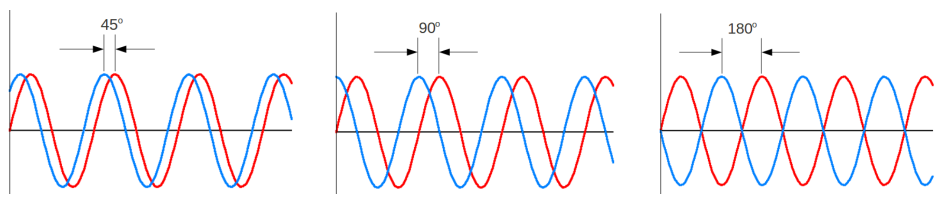


Figure 6.5: (***For color version see Appendix E***) Three situations where two identical oscillating systems differ from one another by a net phase.

The concept of phase will be very important in the coming chapters, especially as we talk about the addition of sound waves by a process called *superposition.*

6.2.3 Adding Vibrations Together

Let's now see how the process of superposition works graphically. Consider figure 6.6, in which the two (solid) curves are being added graphically. They might represent, for example, two simultaneous vibrations from two pure tones impinging on the eardrum. The eardrum can't move in two directions at once from each of two separate tones, so it moves according to their *sum.* In the figure, the curves have equal amplitude but one has twice the frequency of the other (resulting in a musical octave). The auditory system can discern *two* separate, distinct tones, even though the eardrum can only move according to their single superposition curve. (We will take up the mystery behind this ability of the ear later.)

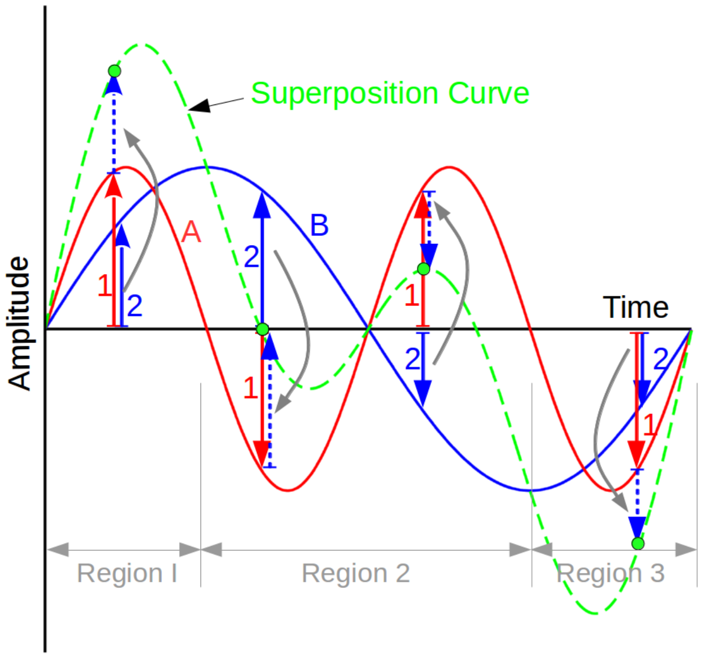


Figure 6.6: (***For color version see Appendix E***) Example of adding two curves by superposition. Arrows show the graphical sum of the two waves in four sample time locations. Arrows for the amplitudes of curves A and B are added head-to-tail to give the result for the dotted superposition curve amplitude. In Region 1 both curves A and B have positive amplitude so that the superposition is above both. In Region 2 the curves A and B have opposite sign amplitude so that the superposition curve is between the two. In Region 3 both A and B have negative amplitude meaning that the superposition curve is below both.

The process by which we find the resulting superposition curve is by "graphically" adding the amplitudes of the two curves together.

> In **Region 1** of figure 6.6, find the point along the horizontal axis where the two solid curve amplitudes are denoted by arrows "1" and "2." The curved arrow shows how arrow 2 can be lifted placed so that its tail connects to the head of arrow 1. Their combined length is then the amplitude of the superposition curve. This is what we mean by the graphical sum – simply adding the length of the two arrows together to arrive at their sum. Since in Region 1 both curves have *positive* amplitude (arrows pointing up), the superposition curve, which results from their sum, will always have *larger* positive amplitude than either curve, at every point along the horizontal axis in that region. Confirm this for yourself by verifying that the dotted curve amplitude is positive and larger than each of curves 1 and 2 throughout all of Region 1.
>
> In **Region 2** the two curves have amplitudes of opposite sign to one another (arrows always pointing in opposite directions) at all points along the horizontal axis . The resulting superposition curve therefore has an amplitude *in between* the two curves. The first graphical summation point in Region 2 (find the first pair of arrows in Region 2) results in total amplitude zero, since the two curves have *equal and opposite* magnitude and therefore the arrows sum to zero. At this point the superposition curve crosses the axis, going from positive to negative amplitude. The second graphical summation point in Region 2 results in a total amplitude which is small and positive and in between the amplitudes of the two curves.
>
> In **Region 3** both curves have negative amplitude (arrows pointing down), so that their superposition sum has greater negative amplitude than each.

The 4 graphical illustration points used in the figure are insufficient to determine the shape of the entire superposition curve with much precision. It is therefore necessary to perform this summing operation at several additional points along the horizontal axis in order to more fully determine the shape of the dotted curve in the figure.

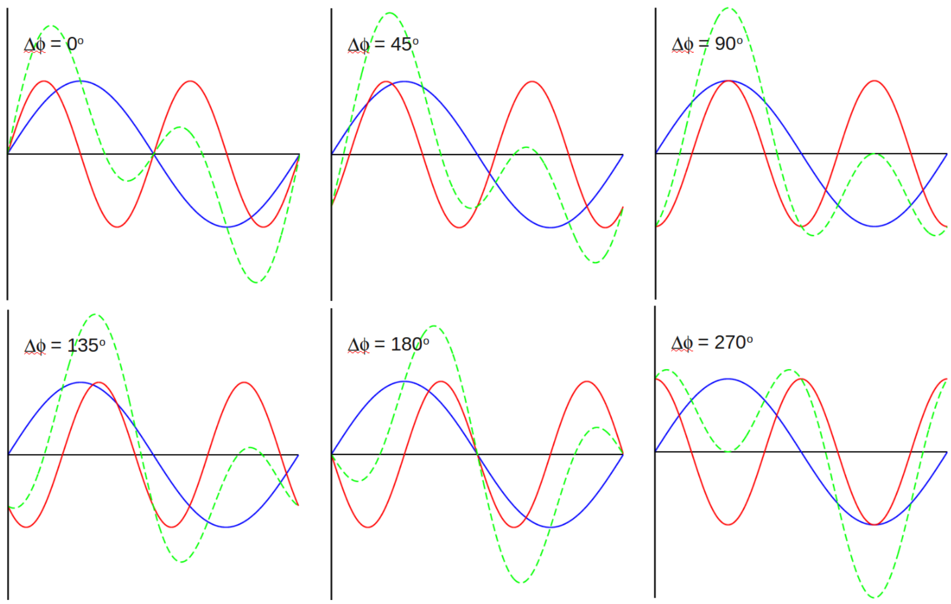


Figure 6.7: (***For color version see Appendix E***) Superposition of two pure tones at various relative phases between the two solid curves. The phases $\Delta\phi$ marked on each pane correspond to the phase difference between the two curves at the origin. Note how the superposition curve changes shape as the relative phase between the two changes.

6.2.4 Superposition and Phase

Note that in figure 6.6 both solid sine curves cross the horizontal axis with positive slope at the origin, meaning that they have zero relative phase at that point and are therefore in phase with one another. If we were to shift one of the two curves to the right or to the left, there would be a non-zero phase difference between them at the origin. Figure 6.7 shows the resulting superposition curves for the octave combination with relative phases of $\Delta\phi = 0^\circ$, 45°, and 90°(upper row), and 135°, 180°, and 270°(lower row).

Note that the shape of the superposition curve changes as a function of the initial relative phase between the two. The remarkable thing is that the ear still hears the two tones as a musical octave for each case, even though the superposition curve has a different shape. Even though the precise motion of the eardrum differs for these various initial phase differences (governed by the specific shape of the superposition curve that drives it), the perception of the tones is the same. We will find out later that the auditory system can pick up on subtle relative shifts in phase between sounds entering each ear and use that information to determine the location of a sound source.

Figure 6.8 depicts the superposition curves for two tones that differ in frequency by a factor of $\frac{3}{2}$ (musical fifth), $\frac{4}{3}$ (musical fourth), and $\frac{5}{4}$ (musical third).

6.3 Sound Spectrum

A powerful graphical tool used to characterize complex waveforms is called a *sound spectrum*. Sound spectra are typically measured using a "spectrum analyzer," a device that accepts as input a complex waveform and "decodes" it into its individual harmonic components. Simply put, a sound spectrum is a plot of all of the individual harmonics that add by superposition to yield that complex waveform.

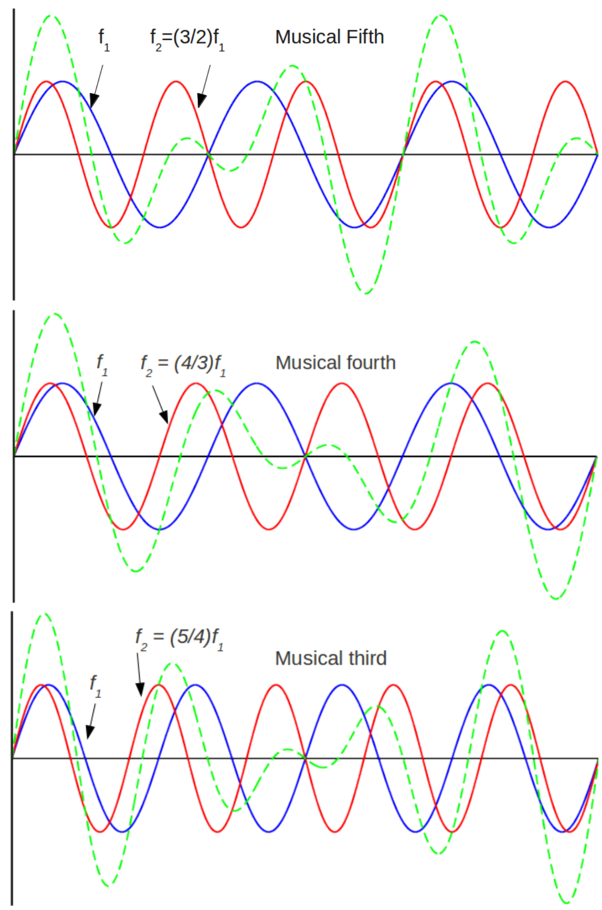


Figure 6.8: (***For color version see Appendix E***) Superposition curves for three musical intervals, the fifth, fourth, and third.

The frequencies of each harmonic are located along the x-axis, and their amplitudes along the y-axis. The simplest sound wave corresponds to that of a pure tone, a single sinusoidal wave. Figure 6.9 shows the sound spectrum of a pure, simple harmonic tone. It consists of a single line, since the sinusoidal wave has a particular frequency (location on the x-axis) with a specific amplitude (height along the y-axis).

If we change the frequency of a sinewave, the horizontal location of its spike in the sound spectrum will change. Likewise, if we change the amplitude of a sinewave, the height of its spike in the sound spectrum will also change. Figure 6.10 shows how the location of the sound spectrum peak changes as the frequency of a sinewave is doubled and as the amplitude is reduced to half.

The sound spectra for two individual sine wave tones with the same amplitude but different frequency ($f_2 = \frac{3}{2}f_1$) are shown in the first two panes of figure 6.11. The sound spectrum of their superposition (the result of being played simultaneously) is shown in the third pane. These two sounds played together will be perceived as two separate pitches forming the interval of a musical fifth. The oscillation of the eardrum (chapter 7) under the influence of this sound wave follows the same pattern as their superposition curve.

The sound spectra of real musical instruments, as measured by electronic spectrum analyzers, do not appear as "clean" as those depicted in figure 6.11, where the sound spectra peaks are thin lines with no

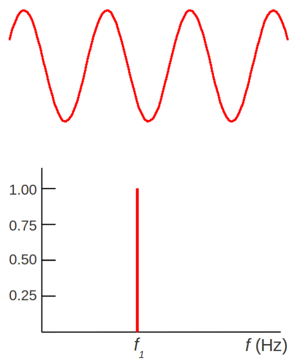


Figure 6.9: A simple sinewave of frequency f_1 and amplitude 1 (above) and its sound spectrum (below).

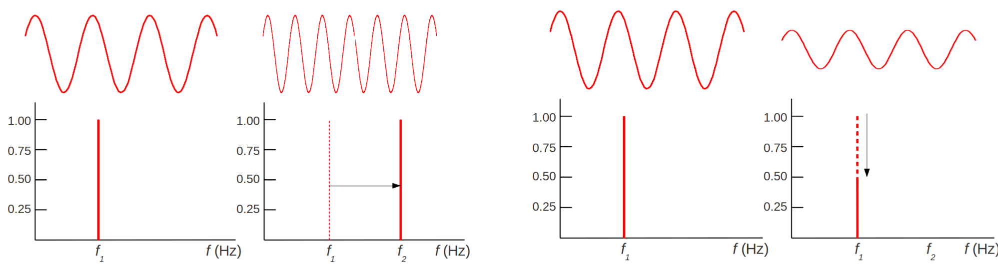


Figure 6.10: When the frequency of a sinewave doubles, the location of its sound spectrum spike moves up the frequency axis by a factor of 2. Likewise when the amplitude is reduced, the height of its sound spectrum spike reduces.

background noise. Figure 6.12 shows waveforms and sound spectra for four instruments, as measured by a spectrum analyzer. Notice the variation in amplitude for the harmonics (in this figure the *sound level*, which is the logarithm of the intensity, is plotted on the y-axis) and the regular spacing of the harmonic peaks along the x-axis, typical of *tonal* instruments.

Each spectrum contain several peaks, corresponding to the instruments' individual harmonics. Note that the relative heights of the individual harmonics differ between the four spectra. The unique collection of frequencies and amplitudes of an instrument's harmonics gives rise to the distinct voice, or timbre, of the instrument. Note in particular that the harmonics for each instrument are equally spaced along the x-axis. This is because their frequencies are integral multiples of the fundamental harmonic. Unlike the two simultaneous sinewave vibrations of figure 6.11, where the ear could hear two distinct pitches, for each of these instruments the ear hears only a single pitch (corresponding to the fundamental) with a distinct timbre (arising from the specific collection of harmonic frequencies and amplitudes that superimpose with the fundamental to produce its complex waveform). We might ask why it is that in the first figure we hear two distinct tones, while in the second figure for the three instruments we hear only a single tone with timbre. We'll take up this mystery in later chapters.

6.4 Sound Waves

So far we've established that the periodic oscillation pattern for any point on a musical instrument undergoing vibration can be represented by a plot of vibrational amplitude (y axis) vs. time (x-axis). This vibrational motion interacts with the surrounding air molecules and sets them into vibrational motion, which can also be represented by a plot of amplitude vs. time (which will look very similar to that of the instrument). A wave is then produced that moves away from the instrument as sound, and all molecules in the room inherit the same vibrational pattern as those adjacent to the instrument

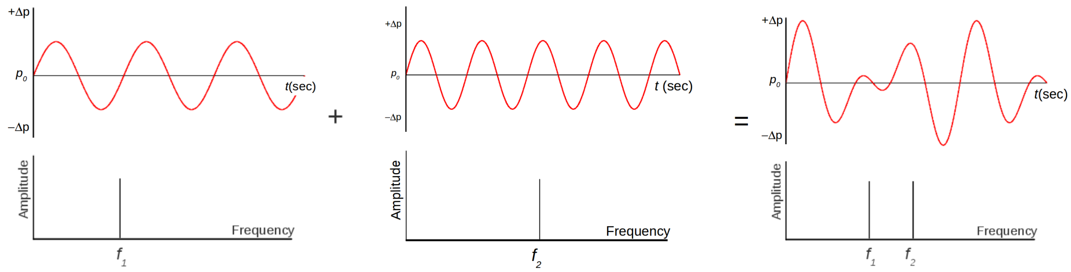


Figure 6.11: Waveforms for sinewaves with frequency f_1 (left), $f_2 = \frac{3}{2}f_1$ (middle), and their superposition (right) along with their sound spectra (below each).

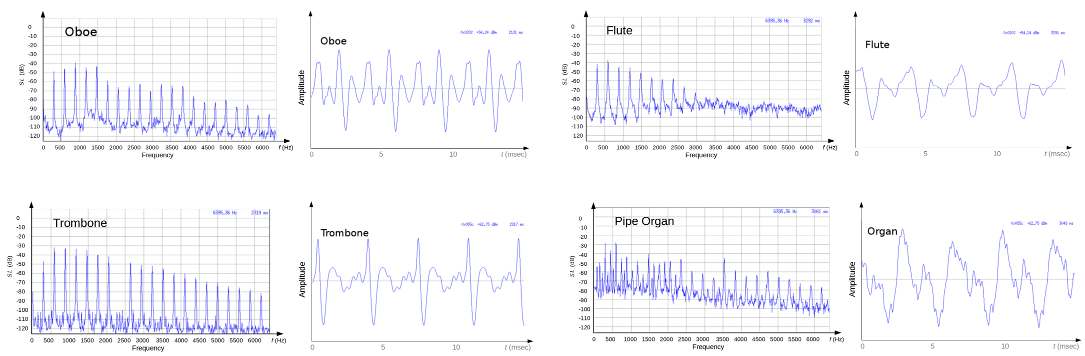


Figure 6.12: Sound spectra for four musical instruments (on the right side of each pair), along with their sound waveform (on the left), as measured by a spectrum analyzer. Note that the sound spectra contain spikes that have shape and width, and that they are positioned on top of a broad base of low-level noise.

body. The sound wave in the air has variation in its shape *both in space and in time.*

6.4.1 Temporal Variation in the Sound Wave

In order to analyze the wave's variation in time, imagine standing to the side of the sound wave as it passes by. Fix your attention on a small patch of air in front of you through which the wave moves, and look at the motion of one of the molecules in that patch. As the wave moves through, the molecule moves left and right in periodic motion about its equilibrium point. (This is similar to the way a cork floats up and down on the surface of a lake as a wave passes by.) At times the molecule is very close to its neighbors, and at other times it is far from them, corresponding to variations in pressure. As a result of the motion of many molecules, the air pressure in this small patch of air has a time variation which can be plotted as pressure (y-axis) vs. time (x-axis). The waveform for this pressure variation will look *very similar* to the waveform for the original instrument vibration.

6.4.2 Spatial Variation in the Sound Wave

Alternatively, imagine standing further back and taking a "picture" of the entire wave as it is spread out along its direction of motion, thereby "freezing" its image in time. Notice the variations in pressure in the wave along this direction of motion. We could plot the variations in pressure of the air as a function of distance from the instrument, in which case our x axis becomes *distance* instead of time.

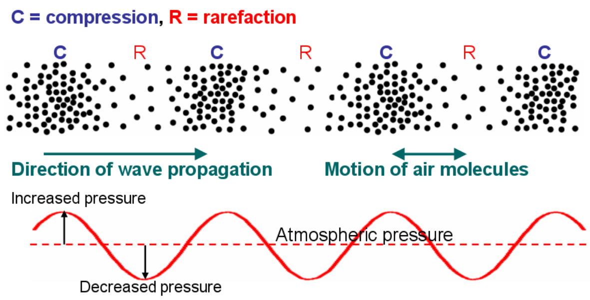


Figure 6.13: A "picture" of a sound wave frozen in time. The physical distance between neighboring compression zones corresponds to the wavelength. At the bottom is a graphical representation of the pressure variations in the medium, which has a similar shape to the vibrational pattern in the vibrational source of the wave.

An important point to note is that the plot of pressure variations as a function of distance x would have a *similar* shape to them plotted in time. For the pattern in space, the distance between successive peaks corresponds to the *wavelength* of the sound. For the pattern in time, the time between successive peaks corresponds to the *period* of the sound.

6.4.3 Superposition of Waves

Consider once again the two sine waves depicted in figure 6.11. If we were to feed an electrical signal of the first of these two sine waves (with frequency $f_1 = f$) into an acoustical speaker, the speaker cone would move in and out with sinusoidal motion and act on the local air to produce pressure variations that move out from the speaker cone as a sound wave. This wave would have the same frequency as the speaker cone's motion. If we were to take a picture of this wave in order to freeze it at a particular moment in time, we would see the sinusoidal variation of air pressure frozen in time and extending out in space across the picture. The distance between neighboring crests would be the wavelength of the wave, which we could calculate using the frequency and the speed of sound. If instead of taking the picture we focus our attention on one patch of air as the wave passes by and look at the pressure variations in that patch as a function of time, we would see the same sinusoidal variation.

If we were instead to feed the acoustical speaker with an electrical signal of the second sine wave (middle pane, with frequency $f_2 = \frac{3}{2}f$), the spatial and temporal patterns of the propagating wave would have correspondingly shorter wavelengths and higher frequencies.

Now let's feed *both* electrical sine wave signals (of frequencies f and $\frac{3}{2}f$) to the acoustical speaker simultaneously. The shape of the cone's motion will be complex now, corresponding to the superposition of the two sine waves as shown in the third pane of figure 6.11. The wave that propagates away from the speaker cone would also have pressure variations of this more complex shape, corresponding to the superposition of the two sources.

Figure 6.14 shows what the three waves (the first two individual waves and the third superposition wave) would look like as a function of distance x. Most sounds in nature are *much* more complex than this example, and yet they can all be envisioned as the superposition (albeit a very complicated one) of many simple sine waves adding together from multiple sources.

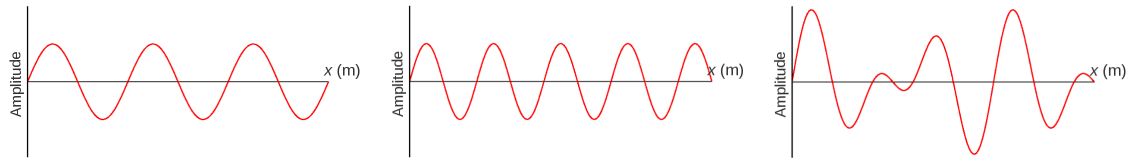


Figure 6.14: Two sinusoidal waves propagating through space, first individually (left and middle) and then simultaneously (right). When they propagate together through space, the pressure variations in space are represented by the superposition of the two waves.

6.4.4 Beats

When two tones with nearly equal frequency are sounded together, they produce *beats*. Figure 6.15 shows two such waves. The two tones might originate, for example, from two tuning forks that differ in frequency output by a very small amount. As sinusoidal waves from each tuning fork move out into the room, the air molecules move in a pattern dictated by the superposition of the two waves. In the upper half of the figure, both individual waves are shown; notice that they have very similar but slightly different frequencies. If you look carefully you will see that there are regions where the two waves are completely *in phase* with one another and therefore add *constructively* (at times t_B and t_D), and regions where they are completely *out of phase* and add *destructively*, that is, cancel one another out (at times t_A, t_C, and t_E). Their superposition curve is shown by the solid-curve waveform in the lower pane, which has a distinctive "pulse" pattern in its amplitude. The dashed curve follows this amplitude variation, which we hear as a quick pulsing in the volume of the sound.

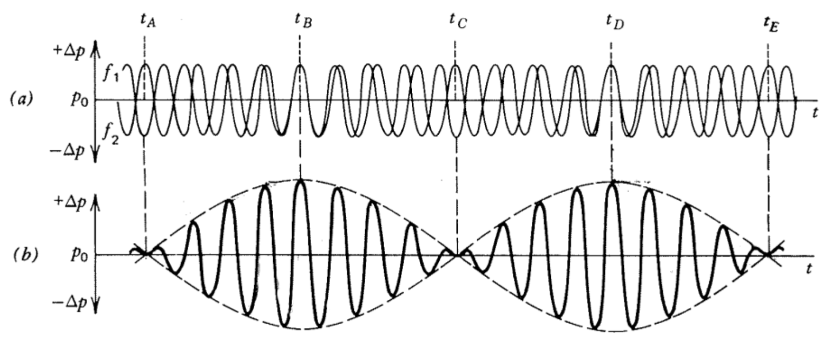


Figure 6.15: Beat wave – produced by two simultaneous sinusoidal vibrations very close to one another in frequency. The superposition curve consists of the single sine wave (solid curve) inside a much lower frequency "envelope" waveform (dashed curve) that shows the variation in the combined wave amplitude, or sound volume. [1]

When the two waves differ in frequency by less than about 15 Hz, we hear their sound combination as a pulsating *single tone* or single pitch (even though it is made of two separate tones!). We hear it as a single tone because the human ear cannot discern such a small difference in frequency between the two separate tones. (We will come to understand this limitation better in chapter 8 when we learn how the human ear functions.) For this reason, we call it a *fused tone*. The two individual tones add by superposition (figure 6.15) to form what our ear perceives as a single fused tone of frequency f_{ft}, which is equal to the *average* of the two separate tones,

$$\boxed{f_{ft} = f_{avg} = \frac{f_1 + f_2}{2}}. \tag{6.1}$$

The frequency of the dashed sine wave (that outlines the variation in amplitude of the fused tone curve) is called the amplitude curve and has frequency f_{amp}, which is

$$f_{amp} = \frac{f_2 - f_1}{2}. \tag{6.2}$$

Notice that the figure shows exactly one full period of the dashed amplitude curve, consisting of two "halves" that each contain a rise and fall in amplitude of the superposition curve. This rise and fall occurs very quickly, and is heard by the ear as a pulse, or "beat." Each full period of the amplitude curve contains exactly two beats, so that the *beat frequency* for this tone is exactly twice f_{amp}:

$$\boxed{f_{beat} = f_2 - f_1}. \tag{6.3}$$

The resulting sound we hear is that of a single tone, whose frequency (and therefore pitch) is exactly midway between the two tones, and whose amplitude (or loudness) is rapidly modulated from zero to maximum and back to zero at the beat frequency. The sensation we therefore have is of a rapid sequence of pulses of sound with a well-defined pitch. Note that the closer the two component tones are in frequency to one another, the lower the beat frequency (equation 6.3), and the farther apart two tones are in frequency, the more rapid the beats. As long as the two are separated in frequency by less than about 15 Hz the ear will hear a single fused tone with beats. When the separation between the two tones becomes larger than 15 Hz, the sensation of beats turns into the sensation of a "roughness" or "buzzing" of the fused tone. Chapter 7 will address this perception in much greater detail as we consider the way the ear processes the sensation of two separated tones over a wide range of frequency difference.

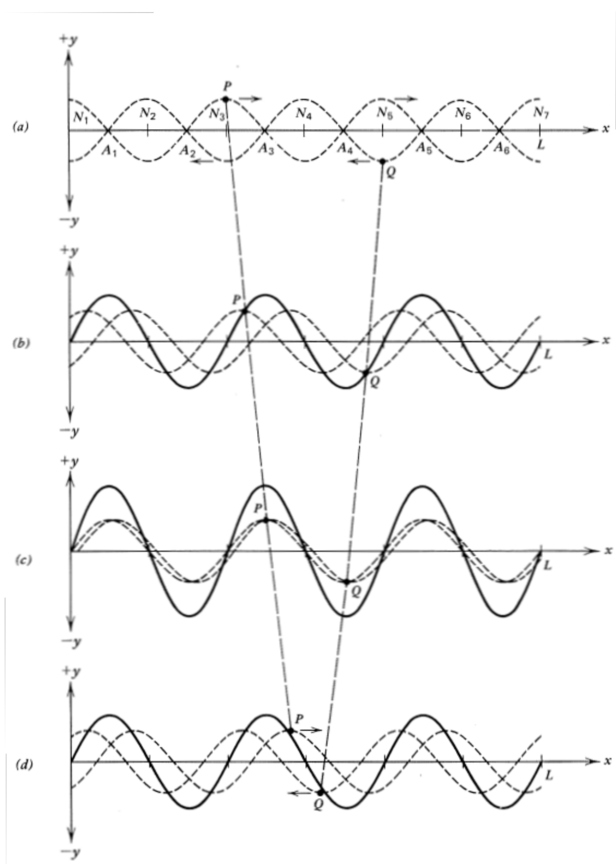


Figure 6.16: A standing wave can be understood as arising from two oppositely traveling waves. The superposition of these two waves is the standing wave.[2]

6.4.5 Standing Waves

We've already seen that standing waves can exist on a vibrating system with well-defined boundaries, such as for a string tied down at its ends, or the air column in a tube of fixed length. Consider the example of the string. When the string is plucked near one end, the pulse from that pluck passes down the length of the string in both directions, bounces off the ends, and returns. This continues in rapid succession, as the two counter-propagating waves move back and forth on the string. They travel at exactly the same speed. Figure 6.16 shows two counter-propagating waves in the top pane, and the pair is depicted as being completely out of phase with one another. The superposition at this instant is zero everywhere along their length. This time therefore corresponds to a time when the string is flat and moving through its equilibrium point. A very short time later (second pane down), the two waves have progressed in opposite directions, so that they add to form a superposition wave with intermediate amplitude. A short time after that they add constructively to form the maximum superposition amplitude. The net result of this back and forth motion of the two traveling waves is to produce a superposition wave that oscillates between zero and maximum amplitude. It appears to be standing still on the string since both component waves move in opposite directions at the *exact same speed*. The regions corresponding to nodes of the standing wave are places where the two component waves *always* cancel, and the antinodes correspond to points where the two component waves move between cancellation and complete constructive interference. We'll talk more about standing waves and instruments in future chapters.

6.5 Chapter Summary

Key Notes

- A periodic waveform, no matter how complex in shape, can always be represented as a superposition of individual sinewaves with different frequencies and amplitudes. The more complex the tone, the larger number of sinewaves necessary to add together to make that waveform.
- Musical instruments emit sound consisting of complex waveforms.
- **Harmonics and Overtones:** For many musical sounds, the constituent tones making up a periodic waveform are related to one another. When the frequencies f_2, f_3, f_4, *etc.* are whole number, or integral, multiples of the fundamental f_1, we call them *harmonics*. When this integral multiple relationship doesn't exist, for example when they are not whole number multiples of the fundamental, they are referred to as *overtones*.
- **Sound Spectrum:** A complex waveform consists of the superposition of numerous individual sinewaves. The graphical representation of the waveform in terms of its component waves is called its *sound spectrum*. We plot the frequencies of the component waves along the x-axis and their relative intensity along the y-axis. Figure 6.9 shows the sound spectrum for a pure tone of frequency f_1, and figure 6.10 shows the sound spectrum for a complex tone consisting of two component tones. Figure 6.12 shows the much more complicated sound spectra for a few common instruments.
- The phase of a wave represents where it is in terms of its full cycle, in angular terms.
- If two identical masses hanging on identical springs are each released at the same time, their oscillations are said to be in phase with each other.
- The sine wave curves that we would use to represent each of these oscillations (plotting the distance from equilibrium on the vertical axis and time on the horizontal axis) will completely overlap.

- Another way of saying that they are in phase is that there is a 0 phase difference between them.
- On the other hand, if one is released at the exact moment that the other is at the opposite end of its cycle, the two are completely out of phase, or 180° apart. The relative phase between two oscillations is a measure of how far apart their two sine waves are "shifted" along the x-axis.
- Sometimes the phase difference is stated as the fraction of a cycle.
- Vibrations can add by the process of *superposition*. Two oscillations that overlap one another add together so that their amplitudes *add* to give the resultant oscillation.
- If two harmonic (that is, sine) oscillations having different frequency are added by superposition, the resultant oscillation can have a complex shape
- An example of superposition might be a speaker cone receiving two separate harmonic signals of different frequency. Since the cone can only move in one direction at a time, it will move according to the oscillation resulting from the superposition of the two sine oscillations. This resultant wave will not be harmonic since the two waves adding together have different frequencies.
- The shape of a superposition waveform will depend on the relative phase of the two component waves
- A sound spectrum consists of a plot of all frequency components that make up a complex waveform, with frequency on the x-axis and amplitude on the y-axis
- Waves superimpose in the same fashion as vibrations superimpose
- Beat waves result from the superposition of two waveforms that are separated in frequency by about 15 Hz or less.
- Standing waves result from the superposition of two counter-propagating waves of the same frequency and speed (and often amplitude).

Exercises

Questions

1) What is meant by superposition of waves? What is it useful for? Explain briefly.
2) What is the relationship between a complex waveform and its sinusoidal component waves?
3) What is the relationship between a complex waveform and its sound spectrum?
4) Can any complex periodic waveform be represented as a superposition of sinusoidal waves of different frequencies and amplitudes? Explain briefly.
5) When two waves of different frequency are added by superposition, does the resulting waveform depend on their relative phase? Explain briefly.
6) When an air molecule moves under the periodic influence of a complex traveling sound wave, how is its physical motion related to each of the individual component sinusoidal waves that make up the complex wave?

Problems

1. When two oscillators have a relative phase of $\frac{1}{6}$ cycle, what is the relative phase angle?

2. Two masses are suspended from identical springs. They both oscillate with equal amplitudes and frequencies. However, they do not oscillate in phase, but rather have a non-zero phase difference between them. Both masses are observed and the times are recorded when each reaches its uppermost position. These times are:

   mass 1: . . . 9.6, 10.8, 12.0, 13.2, . . . (seconds)

   mass 2: . . . 10.5, 11.7, 12.9, 14.1, . . . (seconds)

   What is the period for each of the oscillators? What is the relative phase between them?

3. When two identical frequencies f_1 and f_2 are superimposed, how does the pitch vary with the phase difference between the two oscillations? How about the amplitude?

4. Two tones with frequencies f_1, and $f_2 = \frac{3}{2}f_1$ together form the interval of a musical fifth. If two oscillators are being driven at these two frequencies, f_1, and $f_2 = \frac{3}{2}f_1$, each with amplitude of 0.2 mm,

   a) Show the two waveforms on a y (mm) versus t (s) graph when there is an initial phase difference of 180°.

   b) Carefully draw the superposition curve.

5. From the following data, construct a sound spectrum.

   f_1 =200 Hz I_1 =0.001 W/m^2

   f_2 =400 Hz I_1 =0.0005 W/m^2

   f_3 =600 Hz I_1 =0.0002 W/m^2

   f_4 =800 Hz I_1 =0 W/m^2

   f_5 =1000 Hz I_1 =0.0001 W/m^2

6. Two pure tones of equal amplitude are played, one with a wavelength in air $\lambda_1 =$ 1.40 m, and the other $\lambda_2 = 1.36$ m. The air temperature is 30°C. When these two tones are played at the same time, name and describe in detail (qualitatively and quantitatively) what you hear. Why is it important for the two tones to have equal amplitude in order for this to occur?

7. Two organ pipes emit sound at frequencies 523.0 Hz and 520.6 Hz. What is the frequency of the tone that is perceived when both are played simultaneously? What is the frequency of the beats they produce?

8. If three instruments play together, with frequencies 440 Hz, 438 Hz, and 443 Hz, what beat frequencies will result?

9. For a semiclosed tube length of $L = 43$ cm, what is the wavelength and frequency for the first mode of vibration (n = 1)? For the fifth mode? Sketch the standing wave pattern for the pressure variations in this tube.

10. Consider two adjacent strings on a guitar. The length between the fixed ends for each string is $L = 65.0$ cm and each has linear mass density $\mu = 3.2$ g/m. String 1 is tuned to a fundamental frequency of $f_1 = 220$ Hz, and has the lower frequency of the two strings. When played together, they produce beats at a rate of 6 Hz.

    a) What are the tensions in strings 1 and 2?

    b) The tension in string 2 is now *increased* so that its fundamental mode of vibration produces beats with the *second harmonic* of string 1. If this new beat frequency is 8 Hz, what are the two possible tensions for string 2?

11. Consider an open tube of length 0.65 m and diameter 3.5 cm. When set into its 5^{th} mode of vibration, what is the distance between adjacent antinodes in its standing wave pattern?

12. You observe a standing wave vibrating on a long string. You don't know the length of the string, but you measure the distance between adjacent nodes to be 28 cm. If the string vibrates at a frequency of 60 Hz, with what speed do the component waves move along the string?

References

[1] Rigden, Physics and the Sound of Music, 2nd Edition, pg. 74 (1985)

[2] Rigden, Physics and the Sound of Music, 2nd Edition, pg. 76 (1985)

Chapter 7

Perception of Sound - Part 1

The experience of music has its final and full realization in the ear-brain system. Perhaps it's safe to say that sound that does not end up being *heard* by a listener does not constitute "music" in the strict sense of the word. When a piece of music is stored on a CD or in an mp3 player, what really resides in the storage is *information* represented by a long series of digital bits (as we shall learn in chapter 18). It is not music *per se* but the "promise" of music in stored encoded form. It does not become music until it is decoded electronically, transformed into sound waves, and ultimately listened to by a human being. This might bring up an interesting question ... if a piece of music is played in the forest but there is no one around to hear it, does it constitute music? What do you think? We will now consider the end point of music, the human auditory system, starting with the ear.

7.1 The Human Ear

The ear consists of primarily three sub-systems: the outer, the middle, and the inner ear. Figure 7.1 depicts these three components.

7.1.1 The Outer Ear

The outer ear consists of the auricle (or pinna), the ear canal (or auditory canal), and the eardrum (or tympanic membrane). The function of the auricle, which is the externally visible cartilaginous structure, is to collect sound and direct it into the auditory canal. It has considerably larger surface area than the opening of the canal, and thus serves as a "funnel" to allow more sound to be collected (amplified) than would be possible with just the canal. In order to demonstrate this sound collecting capability of the auricle, try the following. Cup your hands around your ears to form a sort of curved backstop. With your thumbs behind each ear, pull the auricles forward so that they also become part of the cup you are forming. These new hand cups increase the collecting area of the outer ear, and you will notice an amplification in the sound level that you receive. As you adjust the shape and direction of the hand cups you will notice a change in the volume and character of the sound. Notice also that you can hear your own voice with greater volume and clarity in this fashion.

The auricle also, by virtue of its unusual shape, helps us to detect the location of sound sources. Sound originating from a specific location in our local environment reflects from the auricle and is directed into the auditory canal. Very subtle differences in the character of the sound entering the canal occur which depend on how specifically they reflect from portions of the auricle. Sound originating from different locations in the room will reflect differently from portions of the auricle, and the brain is trained over time to detect these subtle differences and associate them with the locations of sound sources. Location detection is also assisted by our having two ears, one on each side of the head, and we will discuss these "binaural effects" in chapter 8.

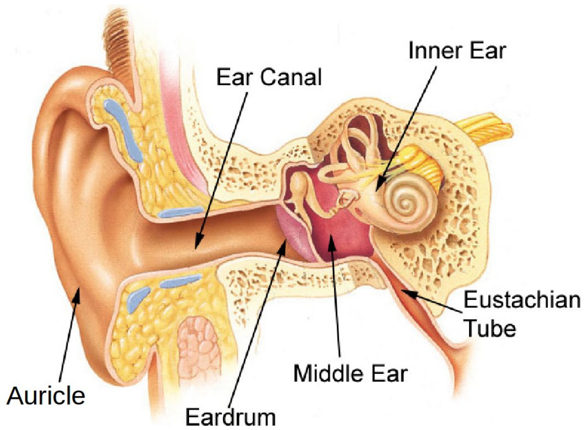


Figure 7.1: Schematic of the human ear, consisting of the outer ear (auricle and ear canal), the middle ear (consisting of the chain of three bones, or ossicles, the opening of the Eustachian tube), and the inner ear (consisting of the cochlea, which passes information on to the brain via the auditory nerve, and the semicircular canals, used for maintaining balance).[1]

7.1.2 The Middle Ear

The middle ear includes the three "ossicles," or the three very small bones (the hammer, the anvil, and the stirrup, or by their Latin names the malleus, incus, and stapes), and the opening of the Eustachian tube, which connects the middle ear with the back of the throat and helps equalize any pressure differences between the middle ear and the outside world.

The eardrum, or tympanic membrane, is a thin piece of tissue between the outer and middle ear. It can be damaged by physical contact or excessively loud sound. Surgical procedures are available today to repair or replace the eardrum. Air is in contact with both sides of the eardrum and is typically at normal atmospheric pressure. Sound waves entering the auditory canal set the eardrum into vibration. The malleus is connected to the eardrum on its inner side, and when set into vibration, acts like a lever as it passes on vibrations to the incus. The incus in turn passes these vibrations on to the stapes, whose faceplate rests against the window of the cochlea, a small coiled tube filled with fluid, membranes, and nerve cells. Figure 7.2 shows a detailed view of each of these small bones.

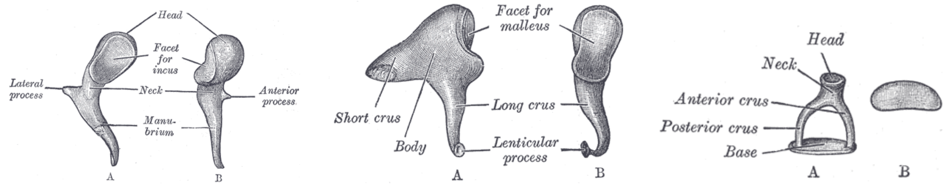


Figure 7.2: (***For color version see Appendix E***) Expanded views of the three bones in the middle ear, the malleus, incus, and stapes.[3]

On a related note . . .

Eustachian Tube and Colds

The Eustachian is normally closed at the point where it meets the middle ear, sealing off and isolating the middle ear from the throat and therefore the outside world. When atmospheric pressure changes occur outside of the ear but remain constant in the middle ear, a pressure difference builds up across the eardrum, causing mild discomfort. This pressure difference can be relieved by temporarily opening the Eustachian and allowing air to flow in such a way as to equalize the two pressures. This is typically accomplished by yawning or swallowing. When you contract a cold, swelling can occur near where the Eustachian meets the middle ear, preventing it from opening. This is why it can be difficult for us to ease the sense discomfort even if we swallow.

7.1.3 Amplification of Sound in the Middle Ear

The pressure vibrations in a traveling sound wave are eventually transferred, via the eardrum and ossicles, into pressure variations in the fluid of the cochlea. The fluid has considerably higher inertia than does air, making it harder to set into vibrational motion, resulting in lower amplitude pressure variations. Fortunately the ossicles serve not only to transmit the vibrations to the cochlear fluid, but to *amplify* them along the way in order to increase the ear's overall sensitivity. This process of amplification happens primarily in two ways [2].

First, the eardrum has an area of approximately 55 mm^2, while the cochlear window has an area of only about 3.2 mm^2. Pressure, as you will recall from chapter 3, is defined as force per unit area, $P = \frac{F}{A}$. Since the ratio of the eardrum area to that of the cochlear window is approximately $\frac{55}{3.2} = 17$, the pressure imparted to the window of the cochlea is approximately 17 times larger, since the force is the same but the area is so much smaller. An additional amplification derives from the different lengths of the ossicles. Because the malleus is longer than the incus, the two form a basic lever. The same work done by the eardrum on the malleus is passed on to the incus. Recall that work is defined as $W = F \times d$. Since the malleus moves farther, owing to its longer length, the incus moves less far, meaning that it passes on greater force to the stapes and therefore to the Cochlear window. As a result of these two basic amplification mechanisms, the pressure induced at the opening of the cochlea is approximately 22 times the pressure induced by the eardrum - a very effective method to increase the sensitivity of the ear!

7.1.4 The Inner Ear

The inner ear consists of the cochlea and three semicircular canals. The cochlea is responsible for decoding the frequency information of a sound wave, and passing electrical nerve signals on to the brain. The semicircular canals are important for helping us maintain physical balance.

The cochlea is the most important and complex component of the ear. The main purpose of the cochlea is to receive the complex vibrations of sound, partially decode the information into its frequency components, and send the information on to the brain in the form of electrical signals. The cochlea is coiled up, but for purposes of understanding its function, it is helpful to envision it stretched out into a straight tube, as depicted in figure 7.3. The cochlea actually consists of two main chambers separated by a very thin membrane called the basilar membrane, which extends along the entire length of the cochlea. When the stapes vibrates against the window of the cochlea, its faceplate acts like a piston that moves in and out at the oval window, producing pressure waves that travel through the fluid and down along the length of the basilar membrane, much as waves move along a whipped bed sheet.

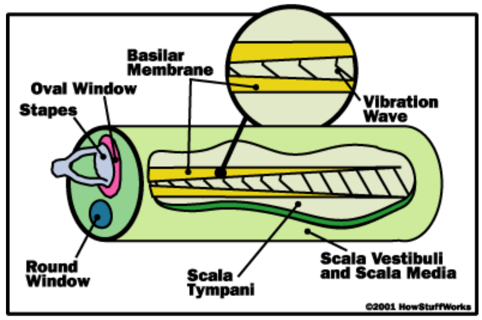


Figure 7.3: Schematic view of the cochlea and its interior. The movement of the stapes at the oval window causes pressure variations to move down the cochlea along the basilar membrane, which activates nerve cells that enable electrical signals to be passed on to the brain.[2]

On a related note . . .

Semicircular Canals, Balance, and Colds

The semicircular canals are looped tubes partially filled with fluid and oriented at specific angles relative to one another. Hair cells lining the inside walls of the canals sense the specific level of the fluid and communicate this information to the brain. As your body tilts in three-dimensional space, the canals rotate as well, while the fluid in the canals remain level with the surface of the earth. The hair cells communicate the movement of the fluid in the canals, sending this information on to the brain which then creates in us the sensation of tilt that we can sense even with our eyes closed. An infection from a cold can create swelling in the semicircular canals, making them less sensitive to the motion of the fluid, which can cause a loss of balance and sensation of dizziness.

To get a sense of how complicated is the function of the inner ear auditory system, even though it is confined in such a tiny space, let's take a closer look at what happens when we listen to music. The basilar membrane contains 20,000-30,000 thin fibers extending across the width of the cochlea with varying lengths and stiffnesses (short and stiff at the oval window and slowly transitioning to longer and more flexible toward the far end). The fibers along the membrane vary in length, and therefore have different resonant frequencies. When a sound wave propagates through the fluid of the cochlea along the basilar membrane, some of the fibers are activated into resonant vibration by the harmonic components of the sound wave. Wherever these basilar membrane fibers vibrate resonantly, they release sufficient energy to set tiny cilia, or hair cells on the "organ of corti" adjacent to the basilar membrane, into vibration. These cilia brush against the tectorial membrane, whose stimulation causes action potentials (basically tiny voltage differences) in auditory nerve fibers, which extend from the cochlea to the brain. The stimulated neurons in this nerve bundle send signals to other neurons, and in this way the signals propagate through a number of regions in the brain (including many nerve fibers crossing over to the hemisphere on the opposite side of the brain to the stimulated ear!) before reaching the auditory cortex.

The neural processing of sound is "tonotopically" organized, meaning that the fibers in the basilar membrane respond maximally to certain frequency ranges, and these frequency range values start from the highest audible frequencies near the oval window all the way down to the lowest audible frequencies at the very end of the cochlea. This tonotopic organization along the basilar membrane is *preserved* in the auditory cortex in the brain. This means that neurons in the auditory cortex that respond to similar frequencies are located adjacent to each other, and the overall geometric placement of the neurons in the auditory cortex are arranged from high to low frequency sensitivity in similar fashion to the organization of the fibers along the basilar membrane[6]. The brain then processes and interprets the sound into the familiar sensation of sound and music. Simply Amazing!

7.1.5 Frequency Range of Hearing

The range of human hearing extends from sounds with frequency as low as around 15 Hz to as high as 20,000 Hz. The sensitivity of the ear ranges widely over this frequency range, as we will see shortly. Sounds lower than 15 Hz, if loud enough, can be "felt" by the whole body more than heard.

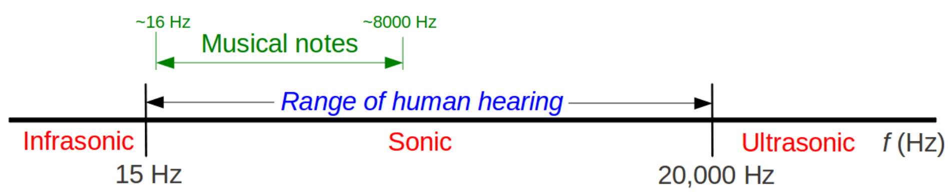


Figure 7.4: (***For color version see Appendix E***) The range of human hearing extends from about 15 to 20,000 Hz. The range of musical notes produced by instruments extends to as high as around 8000 Hz, but harmonics up to 20,000 Hz are very important for timbre and fidelity.

7.2 Pitch vs. Frequency

Frequency is something we can *measure* with instruments that are specially suited for counting cycles and measuring time very precisely. A frequency meter simply counts the number of cycles for a particular waveform over a predetermined period of time. As long as the frequency meter is able to measure time accurately, it can therefore measure the frequency of a tone accurately. On the other hand, *pitch* is a subjective, humanly perceived aspect of sound. We associate pitch with musical tones. The most common musical scale in western music is divided up into 12 semitones of pitch, each with a particular note designation. On a piano, "middle C" is near the center of the keyboard and forms a "home" pitch for music played in the key of C.

The pitch of a tone is closely related to its frequency – that is, we can effectively think of them as being "proportional" to one another without too much loss of generalization. Pitch is a *psychological* attribute of musical sound, measured in "note values," whereas frequency is an objective, *numerical* attribute of the sound wave, measured in Hz. The connection between pitch and frequency is a bit of a mystery. While frequency is absolute, in that it can be determined unambiguously with a sufficiently accurate timing device, pitch is somewhat relative, in that it can be perceived differently by the ear depending on the musical context. There are even "audio illusions" that can fool the ear into thinking it's hearing a pitch that it's not – more on this soon . . . .

We've established in earlier chapters that sound comes about from vibrations. When the body of an instrument, such as a guitar, vibrates in sympathetic response to the vibrating strings, it gives rise to vibration waves that propagate through the air as sound. This traveling wave will have exactly the same frequency as the vibrating guitar string and body. This means that the frequency at which molecules in air oscillate back and forth about their equilibrium positions is the same as the frequency of the guitar vibrations that produced the sound wave. The sound wave moves through the air at a speed determined by properties of the air, including the ambient temperature, but *not* dependent on frequency of the wave. This means that all frequency sound waves will have the same speed. When the sound reaches the human ear, the eardrum and bones of the middle ear are set into vibration at exactly the same frequency. When the inner ear passes this vibrational energy on to the nervous system and brain, the sound is transformed into that magical quality we call "pitch." We don't *hear* "frequency," but instead we *perceive* "pitch." The relationship between the two turns out to be complex, and not yet fully understood. Whereas frequency is absolute, in that it has one particular value that persists, pitch can be subjective, and therefore perceived differently in different circumstances.

7.2.1 Of Beats and Pitch

When the frequency of a vibration is below the range of human pitch sensitivity, we can hear sound, as individual "pulses" rather than continuous tones with identifiable pitch. For example, if two tones close in frequency produce beats at a frequency of 5 Hz, we will perceive the individual beats easily. If the beat frequency is increased to around 15 or 20 Hz, we lose the ability to distinguish individual beats and instead hear a "buzzing." The frequency at which this transition in our auditory system from sensing beats to buzz is located somewhere in the vicinity of 15 - 25 Hz. As the frequency increases to 20 to 30 to 40 Hz and beyond, we begin to experience the emergence of a low pitch note that rises in value as the frequency increases.

On a related note . . .

Hummingbirds . . .

Hummingbirds flap their wings at a frequency around 80 beats/sec. Disturbance of the air at this frequency creates a tone well within the human ear's sensitivity range, giving rise to a very recognizable tone, or "hum" (and hence the bird's name).

Frequencies on both sides of this pitch transition region are critical to music. The low frequency side (~ 1 to ~ 20 Hz) is the domain of music's rhythmic and melodic time structure (*e.g.* from a slow beat to note sequences to a harp glissando), and the high frequency side is the domain of pitch (*e.g.* from bass notes to the piccolo and crash symbol).

7.3 Loudness Perception

The greater the intensity of sound entering the ear canal, the larger the vibrational amplitudes on the eardrum, the larger amplitude of the pressure waves in the cochlea, and therefore the higher the perceived level of loudness to the listener. But does the ear hear all frequencies with equal sensitivity? The answer is no.

We are now ready to consider how a listener judges the loudness of a tone played at different frequencies. So far, the two measures of sound intensity that we've considered are the *intensity*, measured in units of W/m^2, and the *sound level*, SL, measured in db. Note that both of these are "objective" measures of sound, since they can be measured using calibrated scientific instruments. The intensity is a measure of the power emitted by the instrument and how it spreads out and diminishes with distance from the source. The sound level, which is a relative measure of sound between two separate sources, is calculated as the logarithm of intensity ratios, and is still therefore an objective measure. We are now ready to consider two *subjective* measures of sound loudness – subjective in that they depend on the listener's *impression* of loudness, and will therefore vary somewhat from person to person.

7.3.1 Loudness Level

The ear does not respond to all frequencies with equal sensitivity. You might think that two different tones with identical intensity would have the same perceived loudness by a listener, but if they differ in frequency, there is a very good chance that they will be perceived to have different loudness levels. Consider figure 7.5, which shows a curve corresponding to equal perceived loudness over a wide range of frequency. By "equal-loudness" curve we mean that every point on this curve has the same level of perceived loudness to the human ear. Note that in moving from low to high frequency (x-axis values), the curve moves through a variety of sound levels (y-axis values) in order to maintain the same perceived loudness.

The unit of loudness level is the *phon*. The phon is defined in such a way that at the standard frequency of 1000 Hz (a frequency at which the ear is particularly sensitive), the sound level in db is equal to the loudness level in phons:

$$\boxed{\text{For tones at 1000 Hz:} \quad LL \text{ in phons} = SL \text{ in db}}.$$

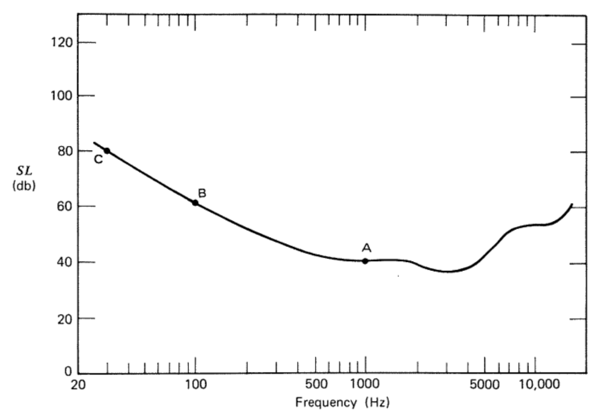


Figure 7.5: The equal loudness curve for loudness level of 40 phons. The perception of loudness is constant at every point on the curve. Note that at low frequencies, the y-axis values are quite high, meaning that for the ear to perceive these low frequency tones to have the same loudness as mid-frequency tones, the sound level must be increased. [4]

The equal loudness curve in figure 7.5 corresponds to a loudness level of 40 phons. At a frequency of 1000 Hz, the sound level that gives this 40 phons of loudness level is 40 db (by definition of the phon). Then, in order to see how the ear responds to other frequencies, we need only follow the equal loudness curve and see what sound level a particular frequency needs to be adjusted to in order to have the same perceived loudness level. For example, point A on the curve indicates the point at 1000 Hz where 40 db of sound level has, by definition, 40 phons of loudness level. Point B, at a frequency of 100 Hz, also has a loudness level of 40 phones (since it also lies on the equal-loudness curve). But note that in order to provide the listener with the same sense of loudness level at this lower frequency, the sound level of the tone needs to be increased to approximately 60 db. Thus in going from frequency 1000 Hz to 100 Hz, the tone needs to be *increased* by 20 db in order for the listener to have the impression that the sound has the same loudness level. Point C indicates that in order for a tone at 30 Hz to have a loudness level of 40 phons, we need to increase its sound level to 80 db. The ear is significantly less sensitive to these lower frequencies, as evidenced by the fact that we need to increase the sound level much more in order for the listener to judge the tone to have constant loudness level.

Figure 7.6 shows loudness curves spanning the entire range of sensitivity for the human ear. Note how much more curved the lines are for low sound intensity levels than for high. The equal loudness curves tend to flatten across the frequency range as the overall sound level increases. The largest overall loudness level difference for low *vs.* high frequencies occurs near the threshold of human hearing.

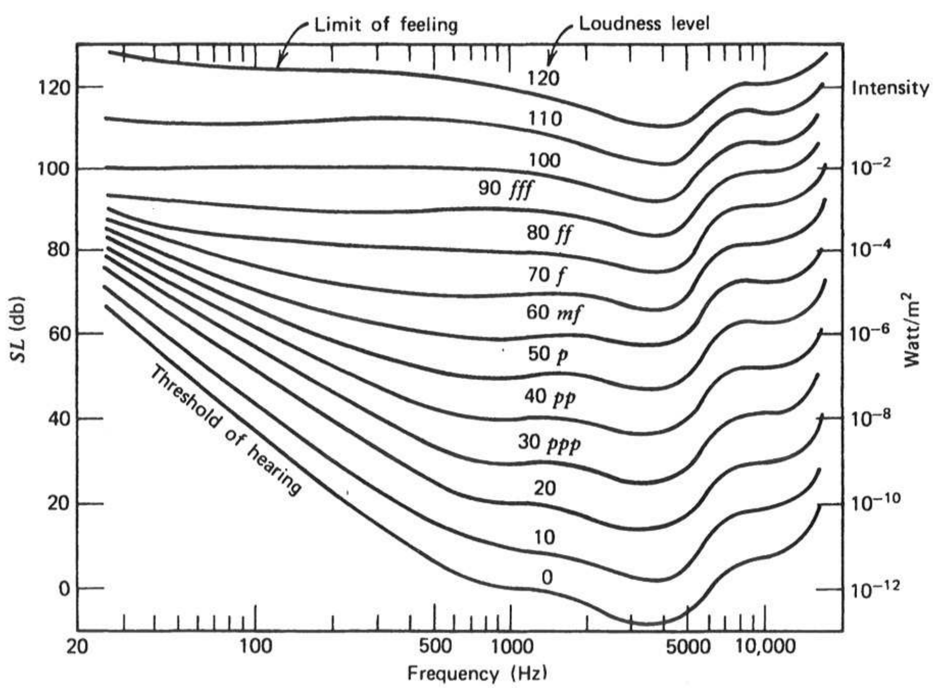


Figure 7.6: The equal loudness curves covering the entire range of human hearing in sound level and in frequency. The loudness levels denoted by the curves are specified in *phons.* [4]

Example 7.1

Sound Level and Loudness Level *What is the loudness level of a 7000 Hz tone whose $SL = 60$ db?*

Solution: Using figure 7.6, we can locate the point at the intersection of horizontal line, traced from 60 db on the y-axis and the vertical line traced from 7000 Hz on the x-axis, to find the point of interest. If you use a rule to locate this point, you should see that the intersection of these two lines occurs very near to the 50 phon curve. Therefore, a sound with $SL = 60$ db at 7000 Hz frequency will produce the sensation of approximately 50 phons of loudness level.

7.3.2 Loudness

The second subjective measure we want to consider is the "loudness" (L), expressed in units of "sones." This quantity measures the ear's sense of different levels of loudness for a particular frequency. A 1000 Hz pure tone with $SL = 40$ db is defined to have a loudness of 1 sone. A 1000 Hz tone with 2 sones of loudness is judged to have twice the loudness of 1 sone. Figure 7.7 shows the variation of loudness (y-axis) with sound level (x-axis). Note that the y-axis is logarithmic.

Note also the difference between the loudness level and the loudness. The **loudness level**, the first of the two subjective measures of the ear's sensitivity we considered, concerns the ear's judgment of the relative loudness of *different frequency* tones. In order to make tones of different frequencies sound the same to the ear, the intensities of each need to be adjusted in intensity or sound level, and figure 7.6 gives a graphical representation of how this is to be done. The **loudness**, the second subjective measure we considered, concerns how the intensity or sound level of a *single frequency* tone needs to be adjusted in order for the ear to hear it as 2 times as loud, 4 times as loud, *etc.*

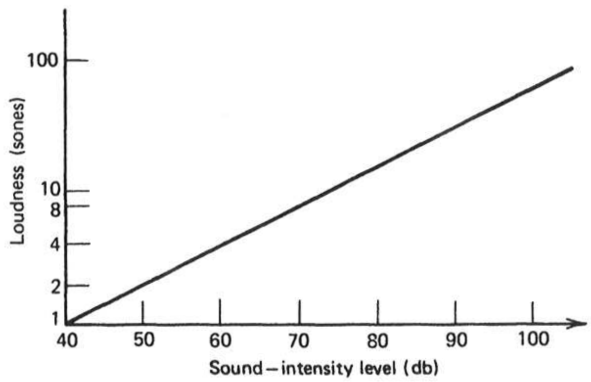


Figure 7.7: The experimental relation between loudness and sound intensity for a pure tone at frequency 1000 Hz[5]

One particularly interesting feature of the ear's ability to discern loudness differences is illustrated in the following example.

Example 7.2

Trumpets and Loudness *How many trumpets playing together at a frequency of 1000 Hz are required to sound twice as loud as 1 trumpet?*

Solution: The "intuitive" answer might be 2! But this is far from the case. See the line in figure 7.7 for the relationship between loudness L (subjective) and sound intensity SL (objective). To double the loudness L of a tone from 2 sones to 4 sones (along x-axis), a change in the sound level SL of approximately 10 db is required (along y-axis). According to the definition of SL, this corresponds to a factor of 10 in the intensity (W/m^2). If each trumpet puts out the same intensity (which is a measure of it power output), then in order to increase the intensity of 1 trumpet by a factor of 10, we would need 10 trumpets in order to sound twice as loud as 1! This result may not be very intuitive, but underscores the remarkable nature of our auditory system. Doubling the intensity of a sound does *not* double our sense of the loudness. The auditory system effectively "compresses" changes of intensity down to a smaller difference in sensation, and this is the main reason behind why we can hear over such a large range of sounds – 12 orders of magnitude in intensity, as we learned in chapter 4.

7.4 Chapter Summary

Key Notes

- The human ear consists of the outer, the middle, and the inner ears. The outer ear includes the pinna, the auditory canal and the auditory membrane (eardrum). The middle ear consists of the three ossicles (bones) and the Eustachian tube. The inner ear consists of the cochlea and the semi-circular canals.
- The middle ear bones serve to amplify the vibrations from the eardrum and transmit them to the cochlea.
- The cochlea serves to turn fluid vibrations in its interior into nerve impulses that are sent to the brain via the auditory nerve.
- Frequency range of the ear extends from about 15 Hz to 20,000 Hz for a young person.
- The frequency of a tone is absolute and can be measured with precision. Pitch is a subjective, psychological attribute of sound and is generally proportional to frequency.
- The *loudness level* of a tone measures how loud tones of different frequencies sound to the human ear. How loud a tone is perceived to be can vary substantially with frequency. When tones of different frequency are played with same intensity level, the ear perceives a difference in their loudness levels, as shown in figure 7.6.
- The *loudness* of a tone is measured in sones and corresponds to the perceived level differences for a single frequency tone played at different intensities. Figure 7.7 plots the relationship for this subjective measure of loudness.

♪ ♪ ♪

Exercises

Questions

1) Briefly describe the function of the three bones in the middle ear.

2) Describe the role of the basilar membrane in the process of hearing.

3) Where and how does a complex waveform become deconstructed into its component waves in the ear?

4) In what two ways do the vibrations of the eardrum become amplified in passing through the middle ear?

5) What is the role of the tiny hair cells of the organ of corti in the auditory system?

6) What is the relationship between pitch and frequency? Which is objective and which is subjective?

7) What happens to our perception of beats when the beat frequency exceeds 15 Hz?

8) What is an equal loudness curve useful for?

9) How is loudness level (measured in phons) related to the sound level (measured in db) for a particular sound?

10) How is the unit of loudness level, the phon, defined?

11) What is the loudness? How is it related to the sound level?

Problems

1. Assume that the outer ear canal is a cylindrical tube 3 cm long, closed at one end by the eardrum. Calculate the resonance frequency of this tube (see figure 5.8). Our hearing should therefore be especially sensitive for frequencies near this resonance.

2. At what frequency does the wavelength of sound equal the distance between your ears, which you can take to be approximately 15 cm? What is the significance of this with respect to your ability to localize sound?

3. The effective area of the eardrum is estimated to be approximately 0.55 cm^2. During normal conversation, the sound pressure variations of about 10^{-2} N/m^2 reach the eardrum. What force is exerted on the eardrum?

4. Suppose a plucked guitar vibrates as if it were a damped simple harmonic oscillator. In terms of the properties of simple harmonic motion, how do you explain the constancy of pitch of the plucked string?

5. a) What is the loudness level of a 1000 Hz tone whose $SL = 60$ db?
   b) What is the loudness level of a 7000 Hz tone whose $SL = 60$ db?
   c) What is the loudness level of a 100 Hz tone whose intensity $I = 10^{-2}$ W/m^2?

6. One flute is playing a tone of frequency 300 Hz with a SL of 40 db.
   a) What is the LL?
   b) If three more flutes join the first and play the same note, each flute contributing a $SL = 40$ db, what will be the LL of the combination?

7. What is the loudness of a 1000 Hz tone whose LL is 60 phons?

8. The loudness level of a 1000 Hz tone is 80 phons at a distance of 3 m from the source. Assuming the validity of the inverse-square law, what would what would be the LL at a distance of 6 m from the source?

9. A 1000 Hz tone and a 100 Hz tone are judged to be equally loud at a distance of 1 m from their sources. The SL of the 1000 Hz tone is determined to be 90 db. At the distance of 1 m,
   a) What is the intensity of the 1000 Hz tone?
   b) What is the sound level of the 100 Hz tone?
   c) What is the intensity of the 100 Hz tone?

   At a distance of 100 m,

   d) What is the intensity of the 1000 Hz tone?
   e) Which of the two tones sounds louder?

10. The total amplification provided by the three bones of the middle ear is approximately 22, meaning that the pressure induced on the window of the cochlea is approximately 22 times the pressure induced by the air wave on the ear drum. The intensity of a tone is proportional to the square of the pressure amplitude. What increase of sound level (in db) for a typical sound with intensity I is provided by the middle ear bones?

11. A large number of pure tones are played in sequence, in order of increasing frequency, starting at 30 Hz and ending at 15,000 Hz. The absolute sound level of each tone is very soft, only 20 db.

    a) What are the lowest and highest frequencies for these tones that would be heard by the human ear?

    b) What is the range of frequencies that would be heard with a loudness level above 10 phons? above 20 phons?

    c) What approximate frequency would be heard with the highest loudness level?

    d) If the sound level of all the tones is increased to 40 db, what is the range of tones that would be heard with a loudness level above 30 phons?

References

[1] http://tmfpff.blogspot.com/2011/06/ear.html

[2] http://science.howstuffworks.com/environmental/life/human-biology/

[3] http://www.learnbones.com/middle-ear-bones-anatomy

[4] Harvey Fletcher and W.A. Munson, "Equal-Loudness-Curves," *Journal of the Acoustical Society of America*, 1933

[5] Rigden, Physics and the Sound of Music, 2nd Edition, pg. 53 (1985)

[6] (Shiffman, H. R. (2001). Sensation and Perception: An Integrated Approach, 5th ed. John Wiley & Sons, Inc)

Chapter 8

Perception of Sound - Part 2

8.1 Two Pure Tones

We are now ready to consider the nature of the ear's response to hearing multiple tones simultaneously. We begin with a perception of two pure tones together, and then will later consider what happens when these tones become complex. Even with so simple a combination as that of two pure tones, the auditory system's response reveals some interesting surprises, depending on the specific frequencies of the two tones, and the difference between them.

8.1.1 Critical Band Plot

Let's now seek to understand and characterize the response of the auditory system to a pair of simultaneous pure tones. Two important features we will be concerned with are

1. the average frequency, $\bar{f}$ of the tones, and
2. their difference in frequency, Δf.

These are the same two quantities that were important for characterizing beats that are produced by two tones with very similar frequency. When we hear two simultaneous pure tones with frequencies that differ by less than 15 Hz, we perceive only a single pitch tone (the fused tone) whose amplitude "pulses" at the beat frequency. You'll recall from chapter 6 that the frequency of the fused tone is the *average* frequency of the two tones, and the beat frequency is the *difference* in frequency. When the separation in the frequencies exceeds 15 Hz, the beats begin to blur into what is called "roughness of tone." As the frequency difference between them is increased further, the two tones are eventually heard as two separate tones when the ear has sufficient resolution to separate them.

Figure 8.1 shows a summary of the ear's response. On the x-axis is plotted the *average* frequency between the two pure tones, and on the y-axis is plotted their *difference* in frequency (the absolute value). A very important point to note is that any pair of pure tones we consider are represented on this plot by a *single point*, whose location ends up in one of the 4 regions shown.

First consider two pure tones with the same frequency of 1000 Hz. These tones are said to be in "unison." Where would they be placed on the plot? Since the difference in their frequencies is $|f_2 - f_1| = 0$ Hz, their y-value would be zero. Their average frequency is $\frac{f_1+f_2}{2} = 1000$ Hz, so we would locate them on the x-axis at 1000 Hz, and $y = 0$. What would we hear? It should be clear that since they share the same frequency, they are the same tone and we would hear just one pitch. Now imagine that we separate their frequencies slightly so that one of them is at $f_1 = 998$ Hz and the other is at $f_2 = 1002$ Hz. Their difference is now $f_2 - f_1 = 4$ Hz. The x-axis point for this pair would still be located at their average of 1000 Hz, but now the y-axis point would be moved up to $|f_2 - f_1| = 4$ Hz, locating the point in **region 1**, corresponding to the sensation of a single pitch tone with beats.

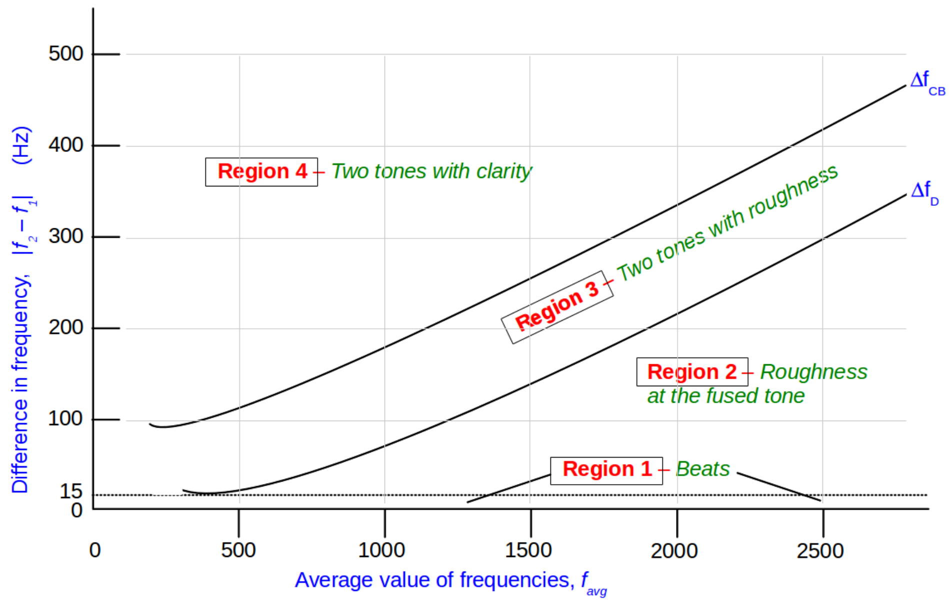


Figure 8.1: (***For color version see Appendix E***) Critical band plot, summarizing the auditory response from perception of two simultaneous pure tones.

With such a small frequency difference between the two, the human auditory system cannot discern the two tones separately, but perceives a single tone, called the *fused tone*. The ear will continue to perceive beats for separations between the two tones up to approximately 15 Hz. Note that this is true regardless of the frequency of the two tones over the entire auditory range – as long as their difference is 15 Hz or less, we perceive beats.

When the frequency difference between the two tones is larger than 15 Hz, what the ear perceives next depends the two tones' average frequency.

When $f_{avg} < 500$ Hz For two tones with average frequency f_{avg} below about 500 Hz and frequency difference $|f_2 - f_1|$ above 15 Hz, the point on the plot moves into **region 3**, corresponding to "two tones with roughness." The ear no longer hears beats and instead begins to perceive two separate pitches with roughness of tone. This perception will continue with increasing frequency difference between the two tones until the difference reaches about 100 Hz, at which point they move into **region 4**, where they are perceived as two tones with clarity.

When $f_{avg} > 500$ Hz For a pair of pure tones with average frequencies *above* 500 Hz and frequency difference $|f_2 - f_1|$ above 15 Hz, , the point moves into **region 2**, "roughness at the fused tone." The ear continues to hear a single fused tone pitch, unable to discern separate pitches. With increasing separation in frequency the point then moves vertically into **region 3** where the tones are discerned as "two tones with roughness," and with increasing difference between them the point eventually moves into **region 4**, where "two tones with clarity" are perceived. Note that for the highest frequencies shown on the chart, the separation between the two tones must exceed about 400 Hz before the ear perceives two tones with clarity.

8.1.2 Discrimination and Critical Bands

The two curves that separate **regions 2** and **3** (labeled Δf_D, the "discrimination band") and **regions 3** and **4** (labeled Δf_{CB}, the "critical band") denote locations of transition in the auditory system's perception. The discrimination band Δf_D marks the transition between when the ear perceives two tones as a single tone or two separate tones. If two tones have a frequency separation *less* than the f_D value, the auditory system will hear their combination as only one pitch (the "fused" pitch). When their separation is greater than the f_D value the ear will hear two distinct pitches.

The critical band marks the transition point where the ear moves from perceiving roughness of tone to perceiving clarity of tone. The frequency separation for the two tones at which these transitions occur depend on the frequencies of the two tones, and as the average frequency increases the separation values for these two bands become larger. Note that the bands are depicted as *sharp lines* in the plot, whereas the actual transitions perceived by the ear are somewhat *subjective* and therefore occur smoothly over a small range of frequency separation. Thus the lines should be understood as a bit "blurry," only to be used as rough guides.

Example 8.1

Perception of Tone Pairs *Two pure tones are played simultaneously. Briefly describe what the human ear perceives when the two tones are, respectively, a) 900 and 1000 Hz, b) 400 and 550 Hz, and c) 1500 and 1600 Hz.*

Solution: a) The average of 900 and 1000 Hz is $f_{avg} = \frac{f_1+f_2}{2} = 950$ Hz, and their separation $|f_2 - f_1| = 100$ Hz. Referring to figure 8.1 and locating the point 950 Hz on the x-axis, we draw a vertical line starting from this point. We then draw a horizontal line starting from the y-axis point at 100 Hz, and find the point where the two lines intercept. This point lies in **region 3**, meaning that the ear would hear two tones with roughness.

b) The average of 400 and 550 Hz is 475 Hz, and their difference is 150 Hz. Again locating the x-axis point at 475 Hz and following the vertical up to the y-axis point at 150 Hz, we see the point falls in **region 4**, so that the perception is two tones with clarity.

c) An average of 1550 Hz and difference of 100 Hz places this last point in **region 1**, so that the perception is roughness at the fused tone. The ear would hear the average frequency of 1550 Hz with a rough tone quality.

8.1.3 Combination Tones

Consider the interval of the musical fifth. The two pure tones that make up this interval have a frequency ratio of $\frac{3}{2}$, *i.e.* $f_2 = \frac{3}{2} f_1$. Note in figure 8.2 that the distance over which the curve for frequency f_1 moves through two complete cycles is the same distance that it takes tone f_2 to move through three complete cycles. The resulting superposition curve has a periodicity *longer* than each of the component tones. Over this same distance it moves through one complete cycle. Therefore the frequency for the superposition curve is $f_s = \frac{1}{2} f_1$. The human auditory system is sensitive to

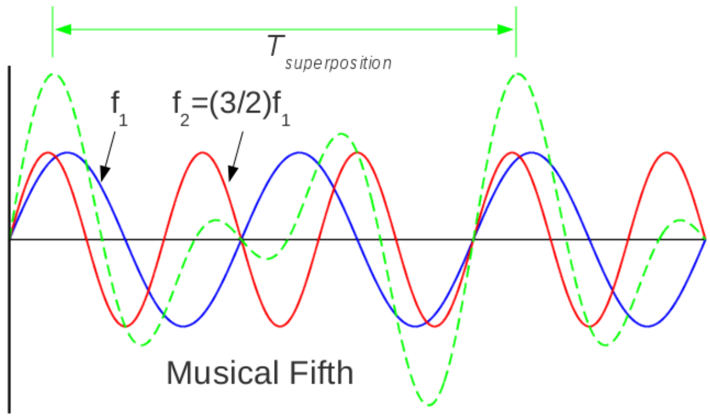


Figure 8.2: (***For color version see Appendix E***) Superposition of two tones constituting the interval of the musical fifth. The periodicity of the superposition curve is longer than both component curves, having period twice that of f_1.

the periodicity of this superposition curve. Even though there is no separate tone present with its periodicity, it is faintly perceived as being present.

The human auditory system is sensitive to and picks up on the periodicity of this superposition curve. The listener hears not only the two tones that are present, but senses the additional presence of a weak lower-pitched tone which is not actually present in the sound. We will now explore the origin of this third tone.

8.1.4 The Origin of Combination Tones

As we previously discussed in section 8.1.1, when two pure tones are very close together in frequency, we perceive beats at frequency $f_b = f_2 - f_1$. As the separation in their frequency is increased beyond 15 Hz, we begin to perceive the beats as a "buzz," since their frequency difference produces beats that are too high in frequency to be heard as separate. The buzz that accompanies the two tones is the origin of roughness of tone that we identified in chapter 7. The buzz has frequency $f_2 - f_1$. When the frequency difference between the two tones increases further, the buzz gets higher in frequency and begins to be perceived (very faintly) as an *additional tone*. Thus what begins as beats when the frequency difference is small, evolves with increased frequency difference into a faint buzzing sound, and with even higher frequency difference into a very faint tone with a well-defined pitch, once the ear can discern it as such.

Thus when two tones with increasing frequency difference move vertically across the critical band from **region 3** to **region 4**, the sound evolves from that of two tones with roughness to that of two tones with clarity. This transition happens when with increasing frequency the buzzing evolves into a faint tone with pitch. The perceived tone is not actually physically present in the original two tones, but is sensed up by the auditory system as a third tone. Since it arises from the combination of the two pure tones, it is called a *combination tone*.

In the case of the musical fifth considered above, the superposition curve has frequency $f_s = f_2 - f_1 = \frac{3}{2}f_1 - f_1 = \frac{1}{2}f_1$. The periodicity resulting from this superposition curve is perceived by the ear as a combination tone with frequency $\frac{1}{2}f_1$, corresponding to one musical octave below f_1. The combination tone for the interval of a musical fourth occurs at frequency $f_2 - f_1 = \frac{4}{3}f_1 - f_1 = \frac{1}{3}f_1$ (which corresponds to an octave and a fifth below f_1), and that of the musical (major) third is $f_2 - f_1 = \frac{5}{4}f_1 - f_1 = \frac{1}{4}f_1$ (which occurs two octaves below f_1).

Here is an interesting and important point. The tones arising from the musical fifth, *i.e.* the two component tones of frequencies f_1 and $\frac{3}{2}f_1$ and the combination tone of frequency $f_s = \frac{1}{2}f_1$, constitute the first three terms in a harmonic series (recall this from chapter 5), with f_s serving as the "fundamental," and f_1 and f_2 as the second and third harmonics, respectively. Thus when two tones constituting a musical fifth are played, the periodicity of their superposition curve enables the auditory system to sense the presence of a fundamental tone that is not actually present, and therefore called a "missing fundamental."

This leads us naturally, then, to the concept of fundamental tracking. When the ear perceives a complex tone consisting of an entire family of harmonics, the tone is decomposed, as we've learned, inside the cochlea into its component frequencies. The complex vibration that moves through the fluid along the basilar membrane sets specific groups of fibers into resonant vibration. These localized vibrations activate specific neurons located in the organ of corti, which then send electrical signals to the brain via the auditory nerve. When this musical fifth is heard, the superposition curve of the two tones causes a resonance along the basilar membrane at the location where the fundamental would normally be located, even though it does not correspond to either of the tones being sounded. The auditory system "senses" the presence of the fundamental tone and "tracks" it as if it also is (weakly) present along with the sound of the two tones. As we will soon learn, *all* musical intervals produce combination tones, some of which are able to be perceived and some of which are quite faint.

8.2 Perception of Pitch

Given the multitude of vibrations corresponding to the complex tone harmonics along the basilar membrane, the question naturally arises as to how we end up associating the sound with well-defined, single pitch. The answer lies partly in the relationship between neighboring harmonics. For a typical sequence of harmonics consisting of even and odd integral multiples of the fundamental, each pair of successive harmonics creates a combination tone with periodicity equal to the that of the fundamental (which is quite remarkable). Figure 8.3 shows a "C" chord whose notes correspond to the first 6 harmonics of a note with pitch C. Note that the difference in frequency between neighboring harmonics is always f, the frequency of the fundamental. It is believed that the human auditory system uses this commonly recurring periodicity in a complex tone to identify its pitch as associated with the fundamental. We can therefore think of successive harmonics as "supporting" and "strengthening" the vibration of the fundamental tone through the combination tone they all have in common. It even happens that when the the fundamental harmonic is removed from the sound, the ear still perceives its presence through the many combination tones produced by the higher harmonics, continuing to give us a strong sense of pitch, even though the fundamental is missing.

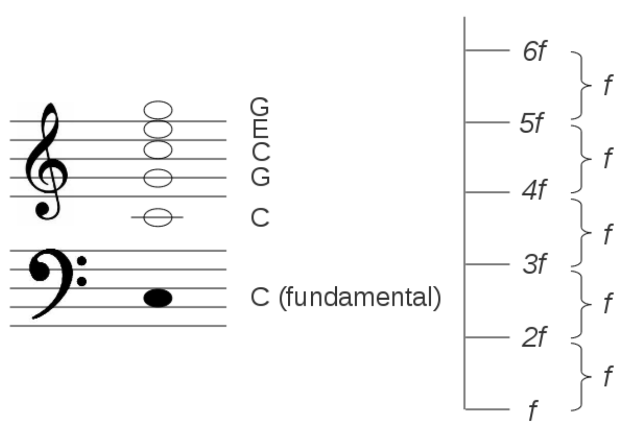


Figure 8.3: Notes corresponding to the first 6 harmonics of a complex tone with fundamental pitch "C". The difference in frequency between all neighboring harmonics, corresponding to the combination tone for each pair, is f, the frequency of the fundamental. In this way the harmonics above the fundamental pitch all serve to support and strengthen the sense of pitch of the complex tone, which is associated with the fundamental.

8.3 Auditory System Non-linearity

The human auditory system is *linear* when processing tones with low to moderate amplitude. To be linear means that in response to being driven by a sound, whether simple or complex, the auditory system passes on to the brain a *faithful record* of that sound, in all its fine detail. Think of the word "linear" as referring to a one-to-one correspondence between the sound source and what is perceived. If processing a pure tone, the auditory system passes on to the brain a sound spectrum consisting of just the one pure tone component without any additional structure added. If processing a complex sound, it passes on to the brain all amplitudes and frequencies faithfully reflective of the original source sound spectrum. In each of these cases, the listener experiences the sound with *high fidelity.* Components of the ear critical to providing this linear response include the specific oscillations of the eardrum and middle ear bones, the faithful translation of these oscillations into fluid waves in the cochlea at the oval window, and finally the translation of these waves into resonant responses along the basilar membrane and conversion of these movements to electrical signals to the brain via the auditory nerve. Linearity translates into high fidelity.

8.3.1 Aural Harmonics

When the eardrum is set into oscillation by low to medium amplitude sounds, it vibrates well within its "elastic limit," and as such passes on with great fidelity the original waveform to the middle and inner ears. (Think of a rubber band: it responds to a gentle pull with a smooth springy resistance, as long as it is not pulled beyond its elastic limit.) If the eardrum is driven by a pure tone, its oscillation will be at the single, well defined frequency of the pure tone. The listener experiences a faithful rendering of the pure tone sound. When the eardrum is driven into oscillation by much higher amplitude sounds, it can be pushed beyond its elastic limit and become stiffer, and its shape distorted. (The rubber band, when pulled beyond its elastic limit, loses its springiness and becomes stiff and unyielding.) The result of the shape distortion of the eardrum is that it no longer oscillates at the single, well defined frequency of the pure tone, but begins to "wobble" in a more complicated fashion. Extra components of oscillation with frequencies corresponding to harmonics of the fundamental begin to appear in the eardrum. Note that these extra harmonic components are *not part of the original sound* but are *added* by the auditory system.

Figure 8.4 depicts the shape that the eardrum might take on as it is driven by a high amplitude pure tone beyond its elastic limit. As its shape becomes distorted, the eardrum oscillates with increased complexity and passes on to the middle ear a sound spectrum which includes additional harmonics not found in the original tone. For a loud tone of frequency f_1, additional frequencies $2f_1$, $3f_1$, $4f_1$, *etc.* begin to appear in the sound spectrum, and the "timbre" of the tone is thus altered. The addition of these extra harmonics to the original waveform constitutes a "distortion" of the sound. The listener experiences a *decrease* in fidelity of the original sound, since the extra components added by the ear alter the perceived timbre of the original sound. We will see in chapter 16 that sound system components and loudspeakers can also add distortion to recorded sound when driven with levels of sound high enough to push them beyond their linear response limits.

The extra tones added by the auditory system to a loud tone at frequency f are called *aural harmonics*, since they are produced at frequencies $2f$, $3f$, *etc.*. It is important to remember that they do not belong to the original sound, but are added by the overstretched eardrum. Other components of the auditory system, including the middle ear bones and the inner ear cochlea, can also exhibit non-linear behavior and add distortion to the sound.

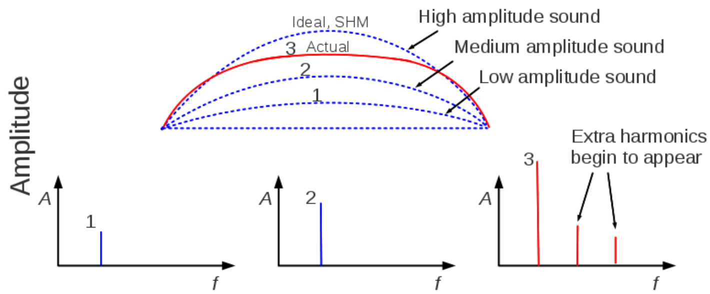


Figure 8.4: (***For color version see Appendix E***) Illustration of the eardrum shape when being driven into oscillation by low (1), medium (2), and high (3) amplitude pure tones. When driven beyond its elastic limit (3), its shape becomes distorted from sinusoidal, and as a result extra harmonics begin to appear in the sound spectrum which are passed on to ossicles of the middle ear. These extra harmonics are not part of the original tone driving the eardrum into oscillation, but are rather added by the eardrum as a result of being stretched beyond its elastic limit.

8.3.2 Aural Harmonics and Combination Tones

When two pure tones with frequencies f_1 and f_2 are played loudly together, the non-linear response of the ear will add several harmonics to the sound passed on to the auditory cortex. Harmonics with frequencies $2f_1, 3f_1, 4f_1, \ldots$ will be added to the perception of tone f_1, and harmonics with frequencies $2f_2, 3f_2, 4f_2, \ldots$ will be added to the perception of tone f_2. The presence of these harmonics constitutes distortion, since the complex timbre they add to the perceived the sound is not actually present in the original pure tones. The presence of harmonics also creates new combination tone possibilities. The loudest combination tone arises from the two pure tones, occurring at frequency $|f_2 - f_1|$. The next two most noticeable combination tones will have frequencies $|f_2 - 2f_1|$ and $|3f_2 - f_1|$. Additional combination tones from the higher harmonics are technically present but can be much more difficult to perceive.

If the two tones form a musical fifth, so that $f_2 = \frac{3}{2}f_1$, the most prominent combination tones will therefore be $f_2 - f_1 = \frac{1}{2}f_1$, $2f_1 - f_2 = \frac{1}{2}f_1$, and $3f_1 - 2f_2 = 0$. The first two of these have the same frequency value, while the the third will not be perceived since has zero frequency. Thus for this loud musical fifth a tone corresponding to the octave below f_1 will be the perceived. Table 8.1 lists the combination tones for common musical intervals.

Table 8.1: Combination tones for several common musical intervals

Interval	f_1	f_2	$f_2 - f_1$	$2f_1 - f_2$	$3f_1 - 2f_2$
Octave	f_1	$2f_1$	f_1	–	–
Fifth	f_1	$3/2f_1$	$1/2f_1$	$1/2f_1$	–
Fourth	f_1	$4/3f_1$	$1/3f_1$	$2/3f_1$	$1/3f_1$
Major Third	f_1	$5/4f_1$	$1/4f_1$	$3/4f_1$	$1/2f_1$
Major Sixth	f_1	$5/3f_1$	$2/3f_1$	$1/3f_1$	–
Minor Third	f_1	$6/5f_1$	$1/5f_1$	$4/5f_1$	$3/5f_1$

8.4 Deconstruction of a Complex Tone

We've already learned that deconstruction of a complex tone occurs predominantly in the cochlea, via localized resonances along the basilar membrane that "map out" the components of a sound spectrum. Electrical signals of these frequency components are then passed on to the brain via the auditory nerve where the actual processing of the sound occurs and the sensation of sound becomes manifest to the listener. Several open questions remain as to how this miracle of sensation occurs. For example, how is it that the brain hears a note from one instrument as a fundamental with timbre, while it hears two separate instruments playing the same note (and therefore sharing the same harmonics, albeit with different amplitudes) as two separate instruments? How is it that the incredibly complex tone arising from a full orchestra with numerous spectrum components spanning a wide frequency range can be decomposed into the sensation of individual numerous instruments? The answers to these questions are still uncertain, and lie outside the scope of traditional physics (being primarily neurological in nature).

Some of the features that can help the auditory system distinguish one instrument from another include the distinct timbres, the relative phasing of the sound arising from each, and the attack and decay character of the sounds from each (a topic we will take up in a later chapter). Because of the strong association we place on the attack and decay characteristics of some instruments, it can be hard to identify the instrument if a recording of one is played backwards, even though the timbre is still present. Another remarkable feature of music is that when a listener is fully engaged in the enjoyment of music, much more of the brain is active in the processing of the sound than the auditory cortex. Other parts of the brain which are active during the experience of music include those portions responsible for processing emotion, memory, learning and plasticity, attention, motor control, pattern perception, imagery, and more.

8.5 Binaural Effects

Our ability to pinpoint the location of sounds in our environment is based on the perception of relative loudness as well as of relative phase of the sound waves reaching both ears. The head plays a large role in this perception. When sound originates from a source located in front of and to the right of a listener, for example, both ears will receive the same sound spectrum, but with slightly different loudness and slightly different phasing. The wavelength of a typical sound wave (say, 1000 Hz) is about 35 cm, or about the size of the human head. Therefore the phase difference in the arrival of the sounds to each ear might have a relative phase difference of approximately 1/2 cycle, or 180°. Additionally, the head will shadow one of the sounds more than the other, creating additional difference in the nature of sound each ear receives. When a listener turns his or her head while listening to the sound, more precise location can be sensed as the phasing and shadowing is varied. All of this is processed in the brain to produce the sense of sound location.

For sounds located in the vertical plane and equidistant from each ear (for example, located below, directly ahead of, or directly overhead of us), there is no loudness or phasing difference between what the two ears hear that might help identify the location of the sound. The shape of the pinna, or outer ear, plays a large role in these cases. The specific shape of the earlobe causes the capture and reflection of sound into the ear canal to have subtle differences depending on the sound's location, and the brain picks up on these subtle cues to locate the sound in the vertical plane. Our ability to locate the position of sounds in our local environment, even when little difference of sound reaches the two ears, underscores the remarkable capability of our whole auditory system, which includes the ear, the auditory nerve, and the brain.

The most faithful recording of the 3-D nature of sound can be accomplished using *binaural* recording, where two microphones are placed in a “dummy head” that has the same size, shape, and absorption characteristics of a typical human head, including soft molded cups representing the outer ears. The microphones, positioned where the two ear openings are located, pick up on subtle phasing and shadowing clues from the contribution of the fake head that would be present for a real human listener. Each microphone signal is recorded in a separate track, and when played back with the use of headphones they render the most realistic reproduction of sound for the listener. The 3-D rendering of sound via binaural recording is sometimes called audio holography, since it shares some of the features of light-based holography for the 3-D rendering of visual fields.

8.6 Hearing loss

The middle ear can be damaged in a few different ways, some of which can lead to significant hearing loss. Small tears in the tympanic membrane can heal automatically, but major tears require surgical attention or even replacement. If the ossicles become displaced by trauma (for example by insertion of foreign objects deep into the ear canal), they can become less effective at communicating the vibrations from the eardrum to the inner ear. If the base of the stapes is pushed into the cochlea, permanent neural-related hearing loss can occur even if the bones are placed back in their proper positions.

The inner ear can also be damaged by exposure to excessive sound levels. High intensity sound waves communicated by the base of the stapes into the oval window produce significant vibration wave amplitudes at the front end of the cochlea. Since high frequency sensation occurs at this end, initial hearing loss is often in the high frequency range. The use of certain drugs and other substances can also lead to neural damage in the organ of corti.

8.7 Chapter Summary

Key Notes

- When a wave with a somewhat complicated amplitude variation enters the ear (perhaps the results of superposition of two simple harmonic waves), the eardrum moves in a fashion directly related to the amplitude variation of the wave. This oscillation is passed on to the bone chain and is eventually passed into the inner ear.
- It is thought that in the liquid inside the inner ear cochlea, each pure tone component is propagated independently, resulting in our perception of two pure tones instead of the complicated superposition.
- Consider the frequency separation of these two pure tones. If $f_2 = f_1$, of course, the two are in unison, and we hear just one tone. If the two tones have different frequency, and if $f_2 - f_1$ is less than 15 Hz, the ear hears beats, as discussed earlier, whose frequency is given by

  $$\Delta f_B = f_2 - f_1$$

  which is twice f_{amp}.
- If the beat frequency exceeds about 15 Hz, the ear no longer hears beats but rather a “roughness” quality to the fused tone.
- Region I of figure 8.1 corresponds to the beat frequency region, where $f_2 - f_1 < 15Hz$. Note that region I occurs between 0 and 15 Hz for all tone frequencies.

- Region II corresponds to the frequency differences that correspond to the perception of roughness of tone (since the individual beats cannot be distinguished) at the fused-tone pitch. Hence, the ear still hears one pitch, but very roughly. Note that this band varies in width over the frequency range.
- Region III corresponds to the perception of two separate tones with roughness of tone quality. It is bounded on the low frequency separation side by Δf_D, the limit of frequency discrimination, above which the ear hears two distinct tones, and on the upper side by Δf_{CB}, the critical band frequency, below which the ear still hears roughness of tone.
- Region IV corresponds to the region above Δf_{CB}, where the perception is of two tones with clarity of tone quality.
- Combination tones: If we hear two tones with different pitch, the difference between them, (*i.e.* $\Delta f = f_2 - f_1$) can create the perception of an additional tone. For example, for $f_2 = 3/2 f_1$ (a musical "fifth" interval), the difference tone between these two fundamental tones is $\Delta f = 1/2 f_1$, an octave below the lower pitch. Thus when this fifth is played loudly, a tone will be perceived one octave below the lower frequency.
- In addition, the ear will add aural harmonics to each of these two fundamental tones, creating $2f_2$, $3f_2$, $2f_1$, $2f_2$, etc. The addition of these aural harmonics creates several other combination tones. T
- A careful analysis of figure 8.2 shows that a complete cycle of the superposition curve is longer than the cycle length of either component sine wave (f_1 or f_2). As a result of this longer period, and thus lower frequency, the ear will hear a tone that is lower in frequency than either component wave.
- The auditory system places significance on this "missing fundamental," and the process of perceiving it is called fundamental tracking. Different combinations of frequencies will produce different missing fundamentals, and the figures on page 86 depict several of these. Be warned, however, that this effect is quite subtle and can be difficult to hear.
- Binaural effects, which arise from our having two ears separated from one another, allow us to detect the location for the origins of sounds.

Exercises

Questions

1) Define the critical band for the hearing process. Is it exact or approximate? Objective or subjective? Explain briefly.

2) What is meant by "fundamental tracking"? Explain briefly.

3) Explain briefly how the harmonics of a complex tone support the vibration of the fundamental to increase our sense of definite pitch.

4) What are combination tones? Are they real or perceived?

5) What are aural harmonics? How do they come about?

6) How are aural harmonics and the sensation of distortion linked?

7) How does the human auditory system use relative phase to pinpoint the location of a

sound?

8) How does the human auditory system discern whether a sound is directly above or directly ahead of us?

9) Describe briefly two types of hearing loss, one associated with the outer/middle ear, and the other associated with the inner ear.

10) Why is it that loss of sensitivity to high frequencies is commonly the first stage in hearing loss?

Problems

1. A pure tone with frequency 1000 Hz is played. A second pure tone, higher in pitch, is played along with the first. If two distinct tones are perceived, what is the smallest possible frequency of the second tone?

2. Two pure tones having frequencies 1000 and 1040 Hz are played together softly. a) What is perceived? b) How about for two tones with frequencies 1500 and 1550 Hz? c) for two tones with frequencies 500 and 650 Hz?

3. At an average frequency of 750 Hz, what is the value of the discrimination band, a) Δf_D? and b) the critical band Δf_{CB}?

4. Two tones have periods of $T_1 = 0.00455$ s and $T_2 = 0.00303$ s. If these two tones are played together, what is the period of the missing fundamental?

5. Two pure tones with frequencies 440 Hz and 587 Hz are played together loudly. What are the frequencies of the combination tones most likely to be perceived?

6. Pure tones with frequencies of 440 and 448 Hz are sounded together. Describe what is heard (pitch of the fused tone, beat frequency). Do the same for tones with frequencies of 440 and 432 Hz.

7. Calculate the first three most prominent combination tones from $f_1 = 900$ Hz and $f_2 = 1000$ Hz.

8. Suppose we have primary tones with frequencies 600 and 760 Hz, each with strong first, second, and third harmonics. List the frequencies of the nine most prominent combination tones that could be formed by these components. Do any of them form a pattern that might enhance the perception of a particular pitch?

9. Four primary tones form the major chord $C_4E_4G_4C_5$ (of frequencies 261.63 Hz, 329.63 Hz, 392.00 Hz, and 523.25 Hz, respectively). What additional notes may arise as combination tones? (List only the most prominent.)

10. Using frequencies of the three notes $B^\flat_2$, $B^\flat_5$, and C_6, verify that the lowest note nearly corresponds to the difference tone produced by the upper two.

11. Using information from figure 8.1 describe the pitch and associated perceptions for an 880 Hz sine wave mixed with another tone of equal strength at a) 882 Hz, b) 890 Hz, c) 920 Hz, and d) 1100 Hz. You can ignore the contribution from combination tones.

12. Measure the distance between your ears. Divide this distance by the speed of sound (assume room temperature) to find the maximum difference in arrival time $\Delta t = (L_2 - L_1)/v$ that occurs when a sound comes directly from the side.

13. Calculate the difference in arrival time at the two ears for a sound that comes from a distant source located 45° from direct north for a person facing due north. Assume that the sound reaches the ears along paths that are parallel to one another (and therefore both along the 45° line), which is a good approximation for a very distant source. Assume that the ears are separated by 15 cm.

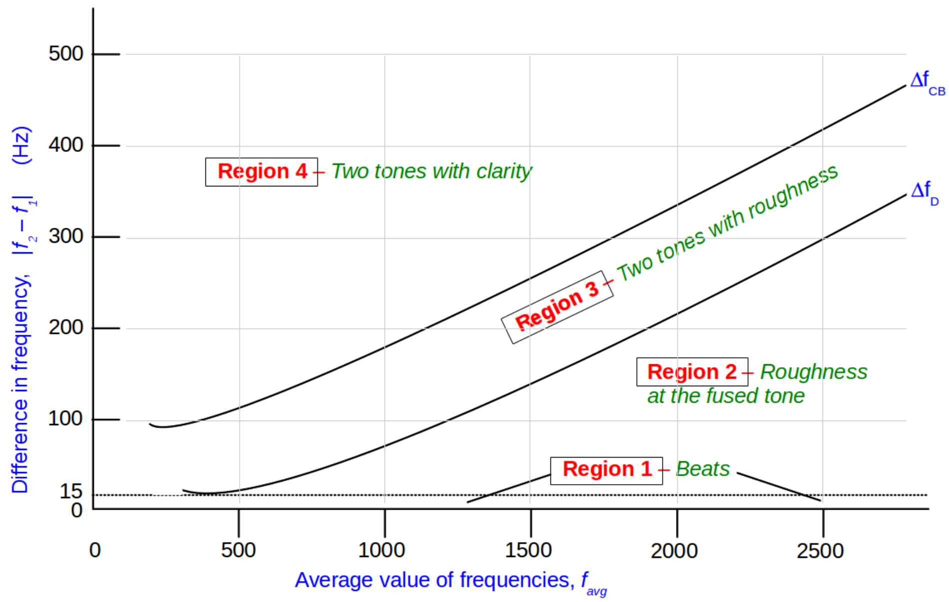


Figure 8.1: Critical band plot, summarizing the auditory response from perception of two simultaneous pure tones.

References

Chapter 9

Stringed Instruments

Stringed instruments are basically as old as music. Drawings exist of the lyre and the harp dating back to B.C. times. The Greeks (particularly Pythagoras) were interested in the vibrational resonances of strings, and developed theories of consonance and dissonance. In the Medieval era, the harp and the lyre were associated with the heavens, and it was believed that angels played them. The earliest stringed instruments used gut strings and animal shells, skulls, and other natural containers as resonators. Modern instruments typically use wooden bodies and steel strings.

9.1 Instruments - The Stringed Family

Most orchestras include a string section, consisting of violins, violas, cellos, and double basses. These orchestral stringed instruments are bowed instruments, meaning that the source of sound from each of them derives primarily from pulling a bow across the string setting it into vibration. Their strings can also be plucked with the finger or tapped with the bow for different sounds. Each of these stringed family members has four strings and a similarly shaped body, with increasingly larger size from violin to bass. Of the four, the violin has the widest range of pitch. Owing to its smaller size, the violin is more versatile than its other family members, agile, and able to play very rapid and complex rhythmic passages. The four strings of the violin are traditionally tuned a musical fifth apart to the values (from lowest to highest) G_3 (196.00 Hz), D_4 (293.66 Hz), A_4 (440.00 Hz) and E_5 (659.26 Hz). The viola is tuned one musical fifth below the violin, the cello one musical octave below the viola, and the double bass not quite one octave below the cello. The lowest string on the double bass is tuned to E_1 (41.203 Hz), and neighboring strings are tuned a musical fourth apart, instead of a fifth as for the other string family members. Given that the limit of human frequency sensitivity drops below about 20 Hz, this is quite a low note! Both the cello and the double bass are too large to hold like the violin or viola, and are therefore have a large peg at the bottom that is placed on the floor. Figure 9.1 depicts the frequency ranges for each of the four family members.

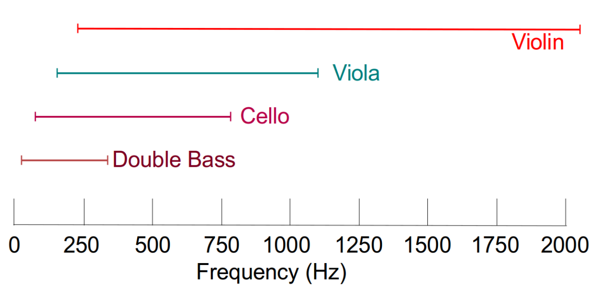


Figure 9.1: Frequency ranges of the four orchestral stringed family instruments

Each of the four family members has four strings, and a wooden box consisting of a top and bottom plate, with sides connecting these two plates. This box serves to amplify and enrich the sound coming from the strings. The interplay between the strings and the box is very complex, and the quality of the sound depends very much on the artisanship of the instrument's construction. The box forms an enclosed volume of air, and together with the properties of the wooden enclosure, form a system that can respond resonantly to a wide variety of frequencies. The strings are stretched from the base of the instrument, where they are mounted to pegs just below the bridge, and over a fingerboard to the top of the instrument, where they are pulled down over a "nut" and wound around the tuning pegs (see figure 9.2). Unlike other stringed instruments, such as the guitar, mandolin, ukulele, and banjo, the orchestral stringed family members' finger boards have no frets. The musician places a finger of her left hand down on the string at a particular position on the finger board in order to shorten the length of the vibrating string, to change its resonant frequency and therefore its pitch. The bow is pulled across the strings just above the bridge, and variations of the bow position relative to this bridge and the strength of the draw on the strings varies the resulting timbre of the instrument.

9.2 The Violin

We'll focus on the violin as the main example of the stringed family. Much of what is described here also applies to the viola, cello, and bass. The strings can be made from metal or gut, or a combination of the two. The box is hollow, delicate and fragile. The total tension of the four strings is about 50 pounds, or about 222 N, amounting to a total downward force of about 20 pounds on the top plate through the wooden bridge. Some mechanical support is required to keep the arched top from being crushed from this force, so a sound post is placed between the top and bottom plates, held in place simply by friction below the right bridge foot – that is, the post is not glued or secured in any other way. See figures 9.2 and 9.3.

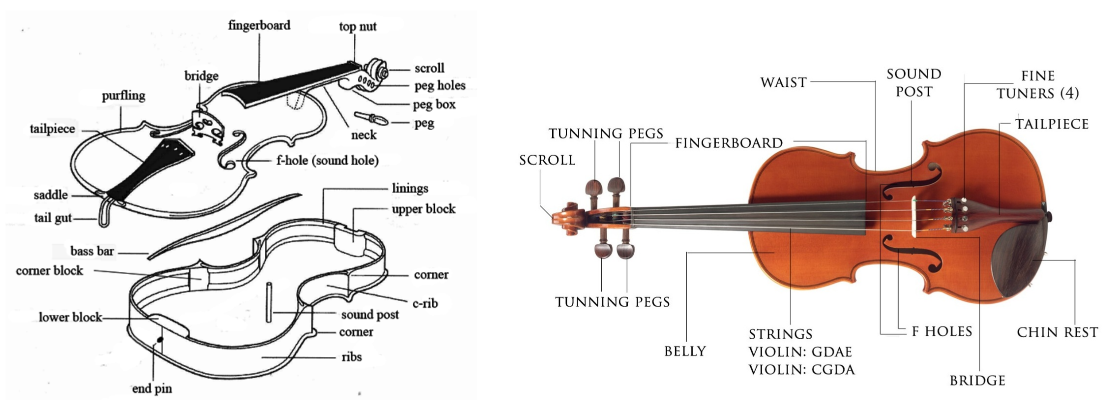


Figure 9.2: Anatomy of a violin.[1]

A long strip of shaped spruce wood is mounted to the underside of the top plate along the length of the violin body underneath the left bridge foot. It is called the "bass bar" since it is located under the lowest-pitched string. The role of both the sound post and the bass bar are essential to the acoustical sound and character of the violin.

9.2.1 Strings and Bridge

The vibrating strings cause the bridge to vibrate, which (as seen through high-speed video) rocks back and forth on the top plate, causing the plate to vibrate. Since the right foot of the bridge is located above the top of the sound post, it does not move much as a result. However, the left bridge foot is free to rock up and down on the flexible plate where no post is positioned, thus transmitting its vibrations to the plate – see figure 9.3. The top plate vibrations are transmitted to the bottom plate by the sound post (hence its name!). Thus the entire body, including the enclosed air volume, vibrates as a result, transmitting sound into the surrounding environment.

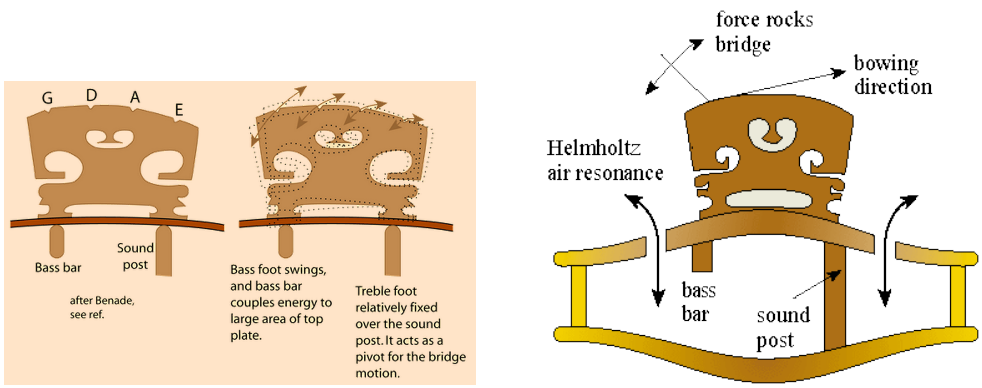


Figure 9.3: Left: The violin bridge, bass bar, and sound post. The predominant motion of the bridge is a rocking motion, which transmits the vibrations of the strings to the body of the violin via the left foot, which moves the plate up and down. Right: The f-holes allow the movement of air as the body of the violin is alternately compressed and expanded in volume by action of the vibration of the left bridge foot.[2]

9.2.2 Resonance Properties

A violin maker can greatly influence the resonant properties of the violin by recognizing the wood quality, including the direction and shape of the grain, and through very careful shaving and shaping of the components. In particular, small variations in the wood quality and shaping and positioning of the bass bar and sound post, and the tightness of the overall construction, affect the resulting tone greatly. The "f-holes" located to the left and right of the bridge are cut into the top plate, which allow the movement of air into and out of the violin body. The body of the violin actively shapes the spectrum of the emitted sound, since there are many frequencies at which it vibrates naturally. The resonant response of the body is quite large and gives the violin its tremendous dynamic range. The volume of air defined by the top, bottom, and sides of the violin also participates in the sound-shaping process. Even though the violin body can resonate at many frequencies over a broad range, there are two resonances at relatively low frequency that are of particular interest to us, having to do with 1) the structure of the wooden box (called the "main wood resonance," MWR) and 2) the air volume enclosed by the box (called the "main air resonance," MAR).

9.2.3 Helmholtz Resonator

The resonance frequency for the second of these two important resonances can be calculated, but only to rough precision. The resonant frequency of a *Helmholtz resonator* (see figure 9.4) is given by

$$\boxed{\text{Helmholtz Resonator: } f_H = \frac{v}{2\pi}\sqrt{\frac{A}{VL}}} \tag{9.1}$$

where v is the velocity of sound in air, A is the area of the hole opening, V is the volume of the resonator, and L is the length of the neck. When these quantities are expressed in MKS units, the frequency is in Hz.

Figure 9.4: Helmholtz resonators (left), useful in understanding the main air resonance for a violin body. The shape is quite different than that of a violin, so that the calculation of a violin body main air resonance will only be approximate, but is useful for getting a feel for the typical frequency range for such a resonance. Variables used in the calculation are depicted in the schematic (right). [3]

9.2.4 MWR and MAR

When the body of the violin is bowed as loudly as possible, the strength of its power output can be measured as a function of frequency, in order to assess the quality of the instrument. The amplitude of the response of the violin body over its entire pitch range, one semitone at a time, can be plotted as shown in figure 9.5. Two of the peaks in the lower frequency range of this response curve correspond to the main wood resonance, MWR and the main air resonance, MAR. The main wood resonance is the lowest resonance of the wood body itself (think if it as a "fundamental" of the wood box structure) and the main air resonance is the lowest resonance of the air volume (again, as if it were a "fundamental" of the air enclosure). An additional resonance, called the "wood prime resonance," has its frequency located one octave below the MWR and arises as a result of the MWR interaction with string harmonics.

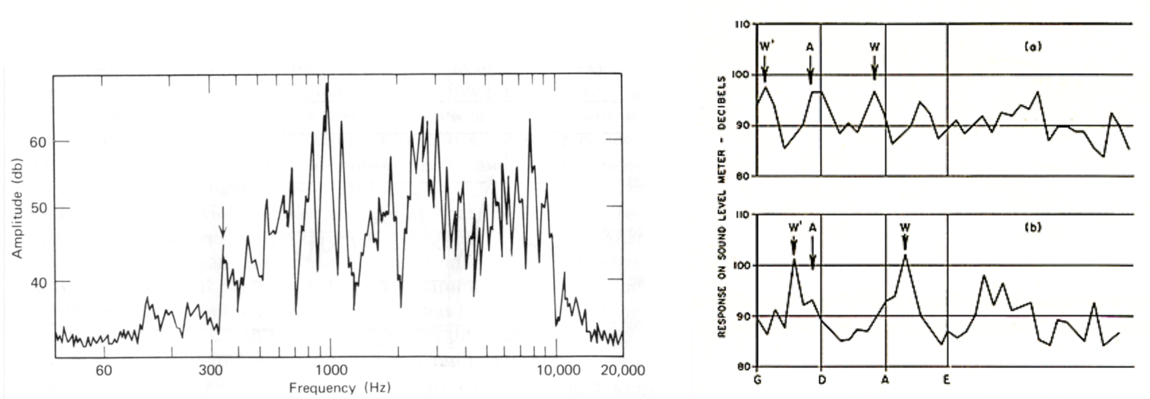


Figure 9.5: The loudness response for a violin which is driven into vibration externally. On the left is a full spectrum over the entire range of the violin's output, and on the right is a closeup of the region corresponding to the frequencies of the 4 open strings. The "W" locates the main wood resonance (MWR), the "A" locates the main air resonance MAR, and the "W'" locates the wood-prime resonance, otherwise known as the "wolf" tone.[5]

Through extensive studies, it has been determined that violins with the best tone quality have their MWR and MAR occur at particular frequencies relative to the open string tunings. The exact frequencies at which the MAR and the MWR occur are important and determine the quality of the violin. As shown in figure 9.5, the good violin (top – a 1713 Stradivarius[5]) locates the MAR (open circle) and the MWR (closed circle) just below, and within a whole tone of the D and the A strings, respectively. Additionally, they should be separated from each other by a musical fifth in pitch. The lower curve in the figure is for a lower quality violin (a 250 year-old violin of unknown origin[5]), where the separation between the MAR and the MWR is larger than a fifth and they are individually located away from their optimal positions relative to the string tunings. Also, we can learn a little more about the violin's overall output by observing some detail in the broader output spectrum on the left in figure 9.5. Note the dip in its output between about 1000 Hz and 1100 Hz, and that the output peaks are much more dense for higher frequencies. This means that we would expect a more uniform and spread-out resonance response from the violin at higher frequencies, in the range of about 1100 – 10,000 Hz. At lower frequencies, where resonances are a little more distinct and separated, we would expect the output to be more non-uniform and somewhat frequency dependent. This is one of the reasons why the placement of the MAR and MWR are critical to the quality of the instrument, especially in its lower frequency range. We'll come back to this point when we discuss the other members of the string family, and why they are not able to be constructed as optimally as the violin in the placement of these resonances.

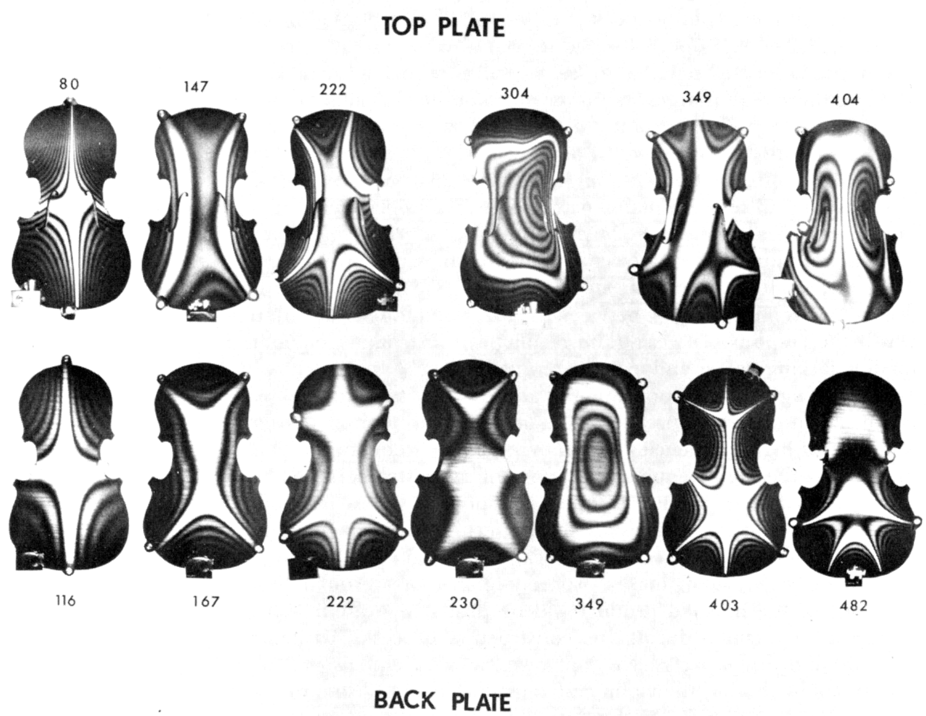


Figure 9.6: Vibrational modes in the top and bottom plates of a violin via a holographic method. The top row images (for the top plate) correspond to frequencies 80, 147, 222, 304, 349, and 404 Hz. The bottom row images (bottom plate) correspond to frequencies 116, 167, 222, 230, 349, 403, and 482 Hz. [4]

Every piece of wood is unique. Considerable skill and experience is required to know how to fashion the top and bottom plates of the violin to maximize the musical quality of the final instrument. Each of the top and bottom plates has "tap tones," which the violin maker adjusts to yield just the right values of tap tone pitch, to give them a "bell-like" tone. Again, as a result of extensive studies, it has been determined that the pitch of the back plate should not be the same as the front plate, but should not differ in their tap tones by more more than a whole tone. When the plates are put together, the body of the resulting instrument also has its own characteristic tap tone. The tap tone for the assembled violin whose response curve is shown in figure 9.5 is marked in the left figure with an arrow.

The tap tones are determined by the shape and character of the wood grain in the plates. The violin maker needs skill and experience working with the wood to obtain clear tap tones at the right frequencies. Using modern techniques of holographic interferometry, the resonance structure of the plates can be assessed over a broad range of frequencies. Several vibrational modes are shown in figure 9.6. As we also saw in the case of resonances in a two-dimensional drumhead, the resonance patterns in the violin plates are quite complicated.

Other ways in which the quality of the instrument is determined involves the care with which the f-holes are carved, which influence the MAR frequency, and the mass and specific positioning of the sound post and the bass bar. The final tone of the instrument depends critically on the wood quality, positioning, and tightness of the sound post. The mass, shape, and positioning of the bass bar are likewise critical in determining the frequency and shape of the tap tone resonance.

It is important to note that *both* the timbre and power of the violin's output are highly dependent on the quality of the instrument's body. A look at figure 9.5 shows that the resonant response of the body is quite broad. When the strings are set into motion, they contain many modes of vibration typical of strings. When these vibrations are communicated to the body, the resonant properties of the body shape and add to the sound in critical ways. When we hear music from a violin, we are hearing the sound character of both the strings and the body, in terms of the power of the output and the timbre.

We've been able to cover only a small fraction of the issues involved in forming a high quality instrument. Clearly there is considerable artisanship in the construction of string instruments. Violin making is an old and established art. While much can be learned from modern scientific investigations of high quality instruments, there is nevertheless really no "scientific" "replacement" for the considerable experience and intuition developed over years of violin making.

9.3 Larger String Family Members

One of the first items to note with regard to the other three members of the string family is that their MAR and MWR cannot be optimally positioned. Recall that for the violin, the optimal positioning of these two resonances is for each to be just below and within a whole tone of the natural frequencies of the two open middle strings. This placement is critical in providing support for the low frequency output of the violin. We can imagine what might be necessary in order to design the viola, cello, and double bass to have similar optimization in the MAR and MWR.

Consider the following. The strings of the viola are each tuned one musical fifth below the violin, and those of the cello are tuned an octave and a fifth below the violin (*i.e.* one octave below the viola). Hence, the longest wavelength of the viola is 1.5 times larger than that of the violin. One could imagine, therefore, that optimization of the viola could be achieved by designing it such that its strings were 1.5 times the length of the violin's (and by the same reasoning, the cello could be designed to have strings 3 times the length of the violin's). If we were to scale all the other dimensions of the viola and cello accordingly, we would end up with dimensions that were 1.5 and 3 times those of the violin, respectively. This would make the viola unable to be held, and the cello would end up the size of the double bass! The actual dimensions of these instruments are only 1.17 and 2.13 times those of the violin, meaning that their MAR and MWR are located much higher in frequency than would be optimal. Thus the lower frequency support for these instruments is reduced, compared with the violin. Figure 9.7 shows typical loudness curves for the viola and cello. Note that both the MAR and the MWR are placed high with respect to the two middle strings. The wood prime peak is higher as well.

The "wolf tone" results from an interaction between the string vibrations and the wood prime resonance. The sound of the wolf tone is very unpleasant. It is therefore very undesirable for the wolf tone

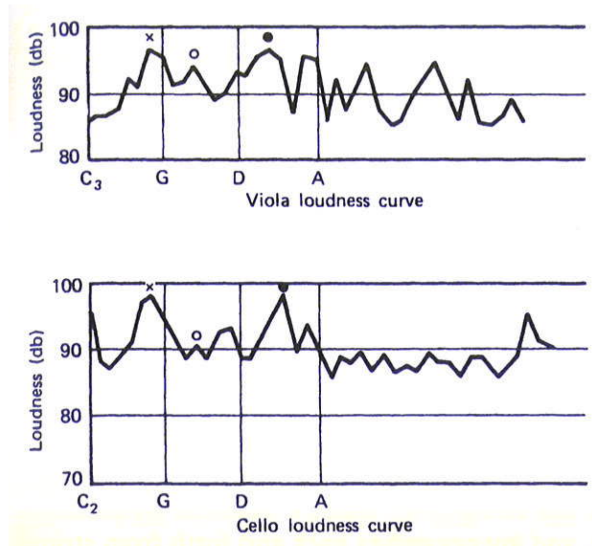


Figure 9.7: Loudness curves for the viola and the cello. The placement of the MAR and the MWR are not optimal for these instruments, as they are more typically for the violin.[7]

to occur during a musical performance. In order to minimize this possibility, violins are designed in such a way as to place the wolf tone midway between two semitones of the instrument.

9.4 Chapter Summary

Key Notes

- **The stringed instruments** – These are principally bowed instruments. The string family consists of the violin, viola, cello, and the double bass. Each has four strings and a similarly shaped body.
- **The violin**: The total tension of the four strings is about 50 pounds, amounting to a total downward force of about 20 pounds on the top plate. Some mechanical support is required to keep the arched top from being crushed from this force. The sound post is held between the top and bottom plates by friction.
- The vibrating strings cause the bridge to vibrate, which causes the top plate to vibrate, which then transmits the vibrations to the bottom plate. The entire body of the violin vibrates to produce the sound, and the patterns of vibration in the top and bottom plates for various notes are shown in the chapter.
- The resonant modes of vibration are very complicated. The "f-holes" are cut into the top plate on each side of the bridge.
- The body of the violin actively shapes the spectrum of the emitted sound, since there are many frequencies at which it vibrates naturally. The resonant response of the body is quite large, and gives the violin its tremendous dynamic range. The volume of air defined by the top, bottom, and sides of the violin also participates in the sound-shaping process.
- See figure 9.5. Two peaks in the response curve of the violin have names: the main wood resonance (MWR) and the main air resonance (MAR). The exact frequencies at which the MAR and the MWR occur are very important and determine the quality of the violin.

- A good violin (top) locates the MAR (open circle) and the MWR (closed circle) just below the D and the A strings, respectively, and they should be separated from each other by a musical fifth in pitch.
- A poorly constructed violin is one where the separation between the MAR and the MWR is larger than a fifth and they are individually located away from their optimal places.
- Each of the top and bottom plates also have tap tones, which the violin maker adjusts to yield just the right values of tap tone pitch. The pitch of the back plate should be about one semitone higher in frequency than the front plate. When the plates are put together, the body of the resulting instrument also has its own characteristic tap tone.
- The violin maker can influence the resonant properties of the violin by design of the bass bar (located underneath the top plate near the lowest pitch string) and the sound post. Small changes in wood quality, tightness of construction, and exact position of the sound post result in a greatly enhanced tone.
- **The larger string instruments**: For practical reasons, the viola and cello are constructed with larger bodies not scaled exactly as the pitch requirements would dictate. The viola is tuned an entire fifth below the violin, and for this reason should be made with dimensions 1.5 times those of the violin. This scaling would make it impossible to play it as it is normally held. Likewise, the cello should be 3 times as large as the violin, but cannot because of playing constraints. Thus, the MAR and MWR of the viola and cello are not placed optimally in relation to their two middle strings. The double bass would be twice as large as it presently is if it were designed with the same optimal MAR and MWR locations as the violin's. Thus, the viola, cello, and bass are constructed with compromises that are made between instrument size and playing convenience.
- The "wolf tone" results from an interaction between the string vibrations and a wood resonance. The sound is unpleasant, and therefore it is undesirable for the wolf tone to occur at a particular note that is often played. Violins are designed in such a way as to place the wolf tone between two semitones.

Exercises

Questions

1) What is the basic function of the sound post in stringed instruments?
2) What is the basic function of the bridge in stringed instruments?
3) What is the main air resonance? How is it important to the function of stringed instruments?
4) How are the pitches of the top and bottom plate tap tones related to each other in a high quality string instrument?
5) What is the optimal placement for the MAR and MWR for a high quality violin?
6) Why is it that the placement of the MAR and MWR cannot be located optimally for a viola and cello?
7) What is a wolf tone? Is it useful for producing music? Explain briefly.

Problems

1. Imagine removing the sound post from an inexpensive violin. If you were then to play the instrument, describe your impression of the difference in sound before and after the post is removed, and why this should be the case.
2. The open D-string on a violin is tuned to a frequency of 294 Hz.
   a) What would be the desirable frequency of the MAR (hint: see appendix C)?
   b) Let's use this frequency to estimate the volume of the enclosed air in a good violin. To carry out this estimation, assume that the thickness of the top plate at the f-holes is 0.25 cm and that the area of each f-hole is 6.5 cm^2. What is the calculated volume? Does your answer seem reasonable?
3. Imagine that all the dimensions of the cello were exactly twice those of the violin. This means that each of the length, height, and width for the cello would be twice the values of the violin. Using the Helmholtz resonator formula, show that the air resonance frequency of a cello would be 1/2 that of the violin.
4. What difference should it make to a violin's sound if it had no f-holes? Explain how the vibrational characteristics of the body would change.

References

[1] , http://www.violinsonly.info/, http://suzukiviolin.com/

[2] http://hyperphysics.phy-astr.gsu.edu/

[3] http://www.phys.cwru.edu/ccpi/Helmholtz_resonator.html

[4] The Acoustics of Violin Plates, by Carleen M. Hutchins, Copyright 1981

[5] Acoustical Measurements of Violins, Hutchins and Fielding, Physics Today, 1968, The Physics of Violins, by Carleen Maley Hutchins. Published by Scientific American, Inc. by Scientific American

[6] Rigden, Physics of Music, 2nd Edition, Wiley, pg. 160

[7] Rigden, Physics of Music, 2nd Edition, Wiley, pg. 167

References

Chapter 10

Woodwind Instruments

The instruments of the woodwind family come in a large variety of sizes and shapes. They are constructed using primarily cylindrical or conical tube geometries. Their resonant properties, including the variety and amplitude of the harmonics produced, are determined not only by their tube shapes, but also by the sizes and placement of the key holes and the shape of the mouthpiece. Sound waves are generated by edge-tones (as in the flute), or by single or double reeds (as in the clarinet, oboe, and bassoon).

10.1 The Woodwind Family of Instruments

The most common members of the woodwind family of instruments are the flute, the clarinet, and the oboe. Examples of other common varieties are the bassoon, English horn, and the piccolo.

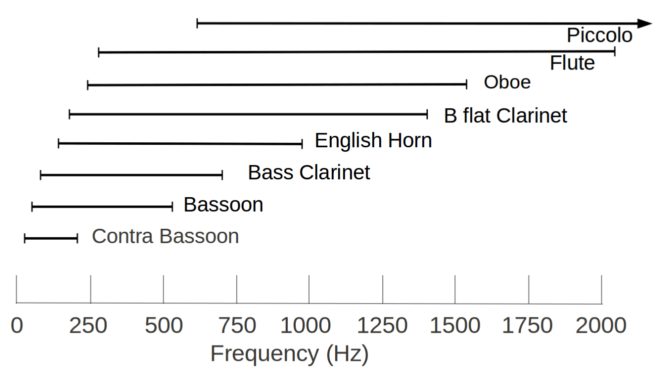


Figure 10.1: Approximate frequency ranges for a number of the woodwind instruments

10.2 Basic Function of Woodwinds

Before we get into a more detailed discussion of individual members of the woodwind family, let's consider some features they all share. We studied the resonance structure of cylindrical and conical tubes in chapter 5. We determined the harmonic structure of each of these geometries, and concluded that open tube instruments support all even and odd harmonics, and that semiclosed instruments only support the odd harmonics:

$$\boxed{\text{Open tube instruments:} \quad \lambda_n = \frac{2L}{n} \qquad f_n = \frac{n}{2L} 20.1\sqrt{T_A} \qquad n = 1, 2, 3 \ldots} \tag{10.1}$$

$$\boxed{\text{Semiclosed tube instruments:} \qquad \lambda_n = \frac{4L}{n} \qquad f_n = \frac{n}{4L} 20.1\sqrt{T_A} \qquad n = 1, 3, 5 \ldots} \tag{10.2}$$

The woodwind instruments are based primarily on these two geometries, but as you can well imagine, these geometries represent "ideal" cases, and the actual resonant properties of the instruments are much more complex and nuanced. At the most basic level, the function of a woodwind instrument is to produce changeable vibrational resonances in an air column. The actual way these resonances come about depends on a lot more features than are present in the ideal open and semiclosed tube cases, which are limited to column length and radius. The variety of timbres arising from members of the woodwind family is sufficient testimony to their more complicated makeup, compared with the simpler tube types on which they are based!

10.3 The Flute

The flute is an open-open cylindrical tube, which supports the vibration of all the whole note multiples of the fundamental. By closing all of the holes of the flute, a performer can play the very lowest note available, which is the fundamental mode of vibration of the entire air column. By "overblowing" (which basically means blowing harder with a tighter pursing of the lips), he or she can promote instead the excitation of the second harmonic, one octave above the fundamental. This is analogous to activating the second harmonic on a guitar string by lightly touching the exact center part of the string and plucking it. The first harmonic, the fundamental, is suppressed, since it would have an antinode at the center, except that the string is being held at that point. By overblowing properly on the flute, the first harmonic can be suppressed in a similar fashion, so that the second harmonic is the primary note being sounded. Additional overblowing can produce the fourth harmonic, or two octaves above the lowest note.

Figure 10.2: A modern silver flute.[1]

Let's explore the function of the key holes in the flute. In order to play all 8 notes of the major diatonic scale (you may remember it as the scale sung to "do-re-mi-fa-sol-la-ti-do'," otherwise known as the solfège), we need 6 keys. Figure 10.3 shows the fingerings required to play all eight notes of the major diatonic scale. First, with all the holes closed (top of the figure) the lowest note sounds, which we call *do*. Opening the key hole at the end shortens the tube somewhat, enabling the note *re* to be produced. Opening the last two holes enables *mi*, etc., until all the holes are opened, at which point the note *ti* is sounded. Finally, with all the holes closed once again, the note *do'*, the octave above *do*, is sounded by overblowing.

Figure 10.4 shows a representation of the sound spectrum, along with the key hole configuration and the standing wave shape for each of the notes. The standing wave corresponding to the lowest fundamental resonant frequency has a node located at each end of the instrument, an antinode at the center, and a wavelength equal to twice the effective length of the tube (first panel in figure 10.4). Note that the sound spectrum peaks occur at integral multiples of the fundamental. When the hole nearest the far end of the instrument is opened, the effective length of the tube is shortened, so that its new end is located at the open hole (the actual effective length end is just beyond the hole, as was the case for

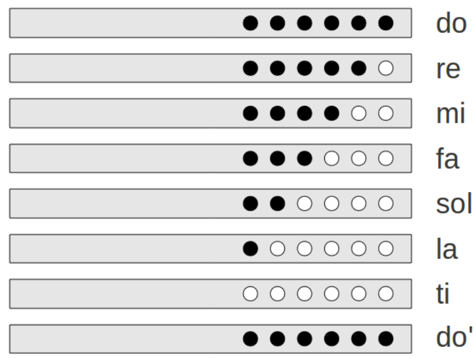


Figure 10.3: Key hole configurations for the flute to play all notes in the major diatonic scale. Note that *do'* is played using the same fingering configuration as for *do* (*i.e.* with all key holes closed) by overblowing.

the earlier effective length discussion for open tube ends). The upper right panel of the figure shows this second note. The actual location of the key holes will be discussed shortly.

Note an important point. Since the flute is an open-tube instrument, its notes contain all harmonics, even and odd. Hence when the fundamental is overblown, the note that sounds is the octave. This is why 6 keys are needed to span the entire diatonic scale. The number of notes is what ever is needed to span the distance between the first and second harmonic of the instrument. When the key sequence for the scale is followed up to the point where overblowing can provide the last note, then from that point onward the same key sequence can be used to span the next octave of notes. In contrast, the clarinet, which is a semiclosed cylindrical instrument, has primarily only odd resonances in its spectrum. The distance between the $n = 1$ and the $n = 3$ resonance in the clarinet, *i.e.* the distance between the first and second supported resonances, is an octave and a fifth. Let's explore the implication of this different sound spectrum and the number of keys required in order to play the entire scale.

10.4 The Clarinet

As a semiclosed cylindrical instrument, the clarinet does not support the vibration of even harmonics. Close inspection of its sound spectrum does reveal the presence of some weaker even harmonics, arising from some of the more complex features not found in the simpler archetypal semiclosed tube, such as keyholes, the mouthpiece, and the bell, to name a few. For the purposes of its operation in playing notes, we need to recognize that the *primary* resonances it supports are the odd harmonics. The second supported harmonic above the fundamental is the $n = 3$ harmonic, which is located a full 12th (*i.e.* an octave and a fifth) above the fundamental. Thus, 10 keys are needed to bridge the gap in pitch between the fundamental and the first note that can be played by over-blowing that same key configuration (see figure 10.6). It is difficult to overblow the clarinet to produce this next higher harmonic, and so the clarinet contains a "register" key, which operates a very small hole near the mouthpiece in order to enable this resonance to be sounded well. The clarinet performer can also overblow to produce the third supported harmonic, *i.e.* the $n = 5$ harmonic resonance, which extends the range of the clarinet considerably.

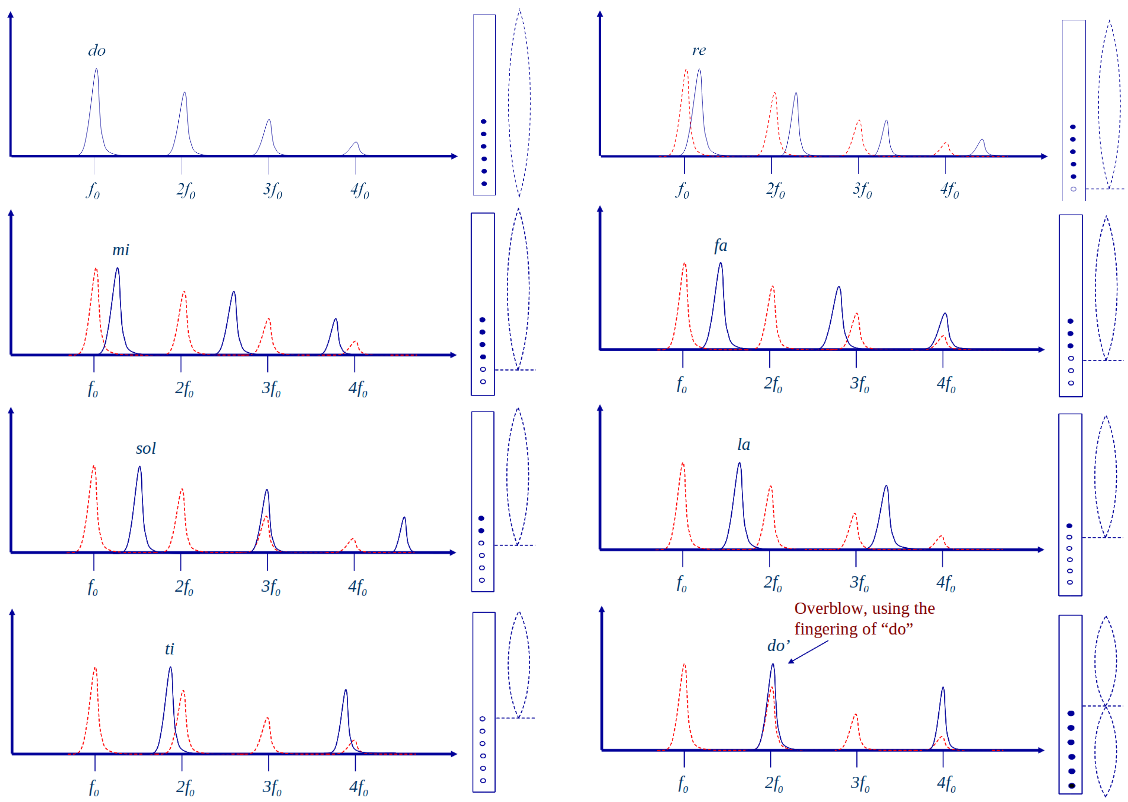


Figure 10.4: (***For color version see Appendix E***) Resonances in the flute as a function of key hole coverage. In each of the eight panes is shown a sound spectrum on the left, which includes the fundamental and several harmonics, and a schematic of the flute with open and/or closed holes on the right, with a representative standing wave showing the resonance for that particular closed hole configuration.

10.5 The Oboe

The oboe has a conical bore. As in the case of the flute, the oboe, which supports both even and odd harmonics, requires only six keys to play the major diatonic scale, since over-blowing activates the second harmonic and the same key sequence can be used to move up through the second octave. To encourage the formation of the second harmonic, the oboe has a small vent hole strategically located on the bore.

It should be noted that for all woodwind instruments, additional holes besides these basic ones necessary for the major scale are present so that sharps and flats (the chromatic scale) can also be played. The holes are not equally spaced along the tube: the spacing between the holes increases down the length of the instrument. For some woodwinds the hole diameters increase as we move down the length, such as in the saxophone, which has a conical bore. The hole sizes increase as the bore diameter increases, so that the effective pressure node is positioned in similar relation to the key location.

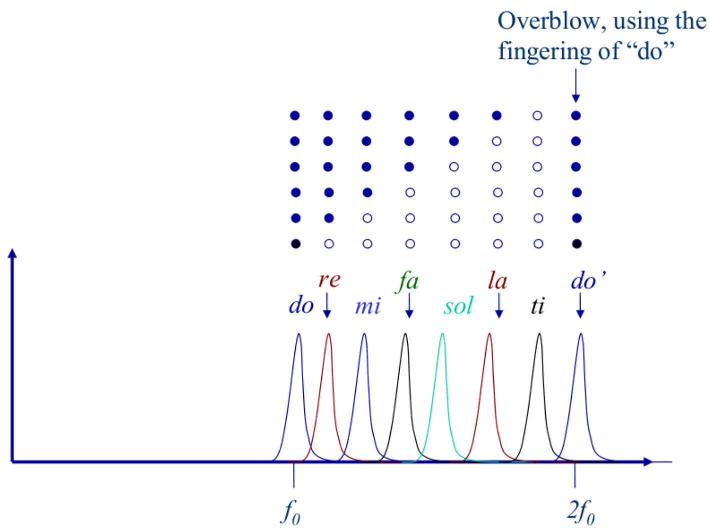


Figure 10.5: (***For color version see Appendix E***) The notes and key hole closings for the flute, starting from all holes closed and back again with use of the register key

10.6 Beyond the Archetypal Forms

Sound produced by the woodwind instruments is quite different than that which would be predicted on the basis of their primary geometry – as open or semiclosed cylindrical or conical tubes. Items that contribute to this difference include the covered and open key holes, the mouthpiece, and the flare (or bell) at the end (for some instruments).

10.6.1 Effect of the Key Holes

Let's now consider the additional effect of key holes on the resonant properties of an instrument, a feature shared by all woodwinds. As we've already learned, the effective length of the instrument is changed by opening and closing the key holes, thereby changing the resonant frequency of the vibrating air column. The exact wavelength of the fundamental resonance depends, as it did for the archetypal forms, on the length of the column and the size of the key hole relative to the tube diameter (which determines the effective length for the location of the pressure node). The fact that the key holes exist already represents a departure from the open and semiclosed tube geometries in the following way. The tube has a wall thickness, and wherever a key hole exists, it adds extra volume to the enclosure by taking away some of that wall material - see figure 10.7. When we take into account all of the key holes, the total volume added can be appreciable. In addition to altering the interior air space (compared with the open and semiclosed archetypal forms), another appreciable effect of these cavity wall holes is in their specific placement, and how this affects the amplitudes of the harmonics produced by the instrument, and hence its ultimate timbre.

You'll recall that we were able to determine the location of the effective end of open and semiclosed tubes by using a formula that includes its radius. We do not have an equivalent formula for locating the effective end of a tube near an open hole. However, we can observe the following: the larger the key hole (relative to the tube diameter), the closer the effective end is to the hole location, and the smaller the key hole, the farther down the line the effective end is located (see figure 10.8). This explains, in part, why the key holes become larger in size in proportion to the increasing tube diameter for the saxophone.

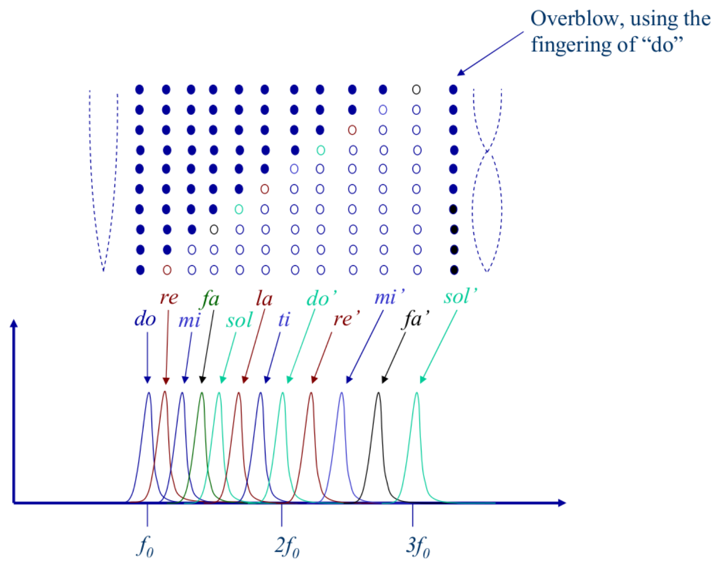


Figure 10.6: (***For color version see Appendix E***) The notes and key hole closings for the clarinet, starting from all holes closed and back again with use of the register key.

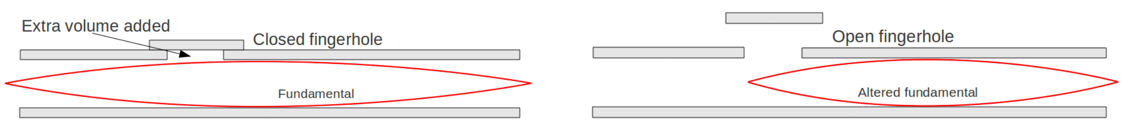


Figure 10.7: Simplified view of the keyhole in a wind instrument. When closed, the hole in the column wall at the location of the keyhole represents added volume to the tube geometry, and when opened, represents a new tube end, thus changing the fundamental vibrational mode wavelength.

10.6.2 Effect of Mouth Piece

The presence of a mouthpiece on any wind instrument changes the resonant frequencies from what they would be for the tube alone. For cylindrical tube instruments (the clarinet, for example), the mouthpiece adds a constant length to the instrument. Since it has a different shape than the cylindrical tube, the length it adds to the air column is not simply its physical length, but an *effective* length. For conical tube instruments, interestingly, the effective length added by the mouthpiece is different for different frequencies. Its effective length is longer for high frequencies than for low.

For the reed instruments, the mouthpiece provides the closed end of the semiclosed tube. As you will recall, the resonant frequencies for a semiclosed tube are

$$\boxed{f_n = \frac{n}{4L} 20.1 \sqrt{T_A} \qquad n = 1, 3, 5 \ldots}. \tag{10.3}$$

Mouthpieces have very different shapes than the tubes to which they are attached. The mouthpiece can share the same radius at the connection point to the instrument tube, but has a more complicated interior shape. How then, can we understand how this mouthpiece affects the resonance properties of the instrument? Let's consider first the case of a semiclosed cylindrical instrument. We can imagine "replacing" this mouthpiece with one that is much simpler – one with a cylindrical interior volume that shares the same radius as the instrument tube, and the same volume as the actual mouthpiece.

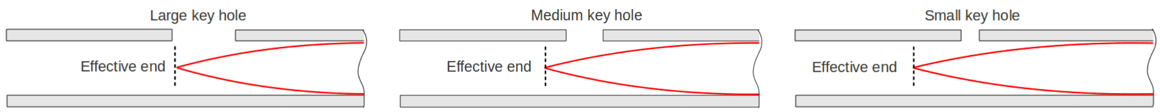


Figure 10.8: The key hole size determines where the effective end of the tube is located when opened.

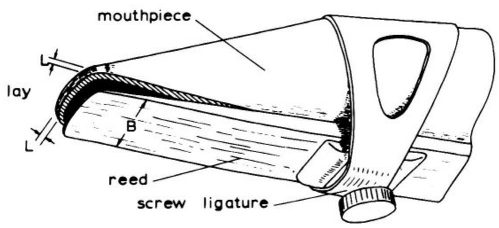


Figure 10.9: Clarinet mouthpiece. Its shape is not that of the tube to which it is attached, and so in order to understand the effective length it adds to the instrument we can use its volume to compute its *effective* length.[2]

This cylindrical mouthpiece affects the resonance properties of the instrument by extending the length of the instrument tube by its own length. We can compute the mouthpiece length since we know its radius (the same as the instrument's) and its volume (the same as the actual mouthpiece we are modeling). Its length will then be used as the *effective* length of the actual mouthpiece.

The volume of a cylinder is $V = (\pi r_{tube}^2)L$. By setting this volume to that of the actual mouthpiece, its length can then be computed using

$$\text{Cylinder:} \qquad L = \frac{V}{\pi r_{tube}^2}. \tag{10.4}$$

This, then, is the effective length of the mouthpiece, which we will call L_{mp}. The addition of the mouthpiece makes the instrument look longer by the amount L_{mp}. The resonant properties of our semiclosed instrument now become, with the addition of the mouthpiece,

$$\boxed{f_n = \frac{n}{4(L_{tube} + L_{mp})}\sqrt{T_A} = \frac{n}{4(L_{tube} + V/\pi r_{tube}^2)}\sqrt{T_A} \qquad n = 1, 3, 5, \ldots}. \tag{10.5}$$

The story for conical instruments is a bit more complicated. The effective length added by the mouthpiece is a function of the frequency and is given by

$$\boxed{L_{mp} \approx \frac{\pi V}{r_{tube}^2} n^2} \tag{10.6}$$

where n is the number of the harmonic. Thus if a tone of frequency f is played, which includes the higher harmonics $2f$, $3f$, *etc.*, the effect of this length dependence on harmonic number is that the instrument looks increasingly "longer" for the higher harmonics. This has the effect of *flattening* the higher harmonics relative to the lower, an effect that ultimately helps shape the unique timbre of the instrument.

10.7 Sound Radiation

When a woodwind instrument is played, energy is supplied by the performer, and this energy is transformed primarily into sound and heat. The production of sound, of course, is the primary purpose of the instrument, and it fills the room in the form of traveling waves of sound energy. Heat is also produced by friction between the air column and the instrument walls. Energy conservation dictates that whatever energy is put into the instrument by the performer will be put out into the environment in these two forms of energy. In order to play a long note, the performer must provide continual energy input to sustain the standing wave in the instrument column. The amount of energy that actually gets converted into sound is only about 2% for woodwinds. They are rather inefficient in this respect.

You might be wondering where the sound is radiated from woodwind instruments. Primary sound emission is through the first open hole below the mouthpiece. The next largest sound emission comes from the next hole down. Not until all the holes are closed does sound leave the end of the instrument. Therefore those woodwinds that have a bell on the end only use the bell when all holes are closed. It turns out that high frequency sound is radiated more efficiently through the bell and through the key holes than low frequency sound. This behavior affects the timbre of the instrument, since higher harmonics are typically enhanced with respect to the lower harmonics giving the instrument a brightness of tone.

10.8 Chapter Summary

Key Notes

- Woodwinds consist of cylindrical and conical instruments. The finger holes modify the resonant frequencies of the tube and thus allow many pitches to be played. The mouthpieces also influence the timbre of the instrument.
- **The flute**: is an open-open cylindrical tube, which allows the air column to vibrate at all the whole note multiples of the fundamental. Since the second harmonic is one octave above the fundamental, 6 holes are needed to play the diatonic scale (7 pitches). The open cylindrical instrument is played by stopping various holes, and the octave is played by over-blowing the fingering of the fundamental. The flute can hit the third and sometimes the fourth harmonics. In this fashion, six keys provide a range of three to four octaves.
- **The clarinet**: is a semiclosed cylindrical instrument, and for this reason is missing its second harmonic. The first harmonic above the fundamental is the third, which is located a full 12^{th} above the fundamental. Thus, 11 keys are needed to bridge the gap in pitch between the fundamental and the first note that can be played by over-blowing. To assist with hitting the third harmonic by over-blowing, the clarinet contains a register key, which operates a very small hole near the mouthpiece. The clarinet can over-blow the fifth harmonic as well which extends the range considerably.
- **The oboe**: has a cone-shaped bore. Like the flute, the oboe, which has both even and odd harmonics, requires only six keys to play the basic scale since over-blowing activates the second harmonic. To encourage the formation of the second harmonic, the oboe has a small vent hole strategically located on the bore.
- For all these instruments, additional holes are present so that sharps and flats (a chromatic scale) can be played. The holes are not equally spaced along the tube: the spacing between the holes increases down the length of the instrument. For some woodwinds the hole diameters

increase down the length, such as in the saxophone, which has a conical bore. The hole sizes increase as the bore diameter increases.

- The mouthpiece adds additional volume to the instrument, and adds to the character of the instrument's timbre.
- The sound emanates from the holes, primarily the first open hole. Energy is lost through the interaction of the air column and the walls of the instrument, and through the radiation of sound.
- Note that the higher harmonics fall successively further and further below the exact odd multiples, so that by the 9th harmonic, the actual location of the peak is closer to the eighth harmonic frequency, and this is where the strongest amplitude exists for the harmonics. This feature gives the clarinet much of its distinct timbre.

Exercises

Questions

1) Name two features of real woodwind instruments that change the character of their supported resonances from the simpler archetypal forms.
2) What is the purpose of the keys on woodwind instruments? Explain in terms of the fundamental operation of the instrument.
3) How many keys are required on a flute in order to play all eight notes of the major diatonic scale? Explain your reasoning briefly.
4) How many keys are required on a clarinet in order to play all the notes of the major diatonic scale over two full octaves? Explain your reasoning briefly.
5) What is the purpose of the small vent hole in an oboe?
6) Why is it that the presence of the keys, even if all of them are closed, still produce a difference in harmonic structure of a wind instrument compared to the that of the archetypal form on which it is based?
7) What effect does the mouthpiece have on the resonant structure of woodwinds?
8) Where on the instrument is the majority of sound radiated from woodwinds? Can you suggest why?

Problems

1. Assume the $B^\flat$ clarinet has an overall length of 67 cm.
   a) Based on this actual length, what should be the lowest frequency of the clarinet?
   b) The lowest note played by the $B^\flat$ clarinet is D_3. What is the effective length of the clarinet?
   c) How do you account for the difference in the answers to a) and b)?
2. It is possible to fit a flute-type head joint to a clarinet so that it plays in the manner of a flute. What would you expect the lowest note to be? Additionally, explain whether you would expect the timbre of the instrument to change significantly, and why.

3. The length of a flute (from the embouchure hole to the open end) is about 60 cm. What do you expect the frequency of the lowest note to be?

4. Explain why the flutist can accomplish with six holes what a clarinetist requires seven holes to achieve, namely to play the sequence of notes making up the diatonic scale.

5. How many side holes must a flute-like instrument have if *two* octaves are to be played? How many side holes must a clarinet-like instrument have if *two* octaves are to be played? Explain briefly why.

6. Assume that the mouthpiece of a clarinet has a volume of 8.0 cm^3. The mouthpiece and the bore onto which it slides have an inner radius of 0.75 cm. Calculate the effective length that the mouthpiece adds to the length of the instrument. In general terms, how would its addition alter the frequencies of the tube to which it is attached?

References

[1] http://oz.plymouth.edu/∼ asjacques/CSDI1200/woodwinds.html

[2] http://www.speech.kth.se/music/publications/leofuks/thesis/instruments.html

Chapter 11

Brass Instruments

So far we've seen that for the basic instruments of the string family, shape and physical characteristics are very similar. And the woodwind instruments have essentially the same shape, being based on cylindrical and conical bore geometries. For the woodwinds in particular, we saw that a few important features caused the resonant structure to differ significantly from that of the simpler plain tubes, namely the key holes, mouthpieces, and flares at the end of a few. What makes the brass instruments different than the woodwinds is that they combine several tube types into one instrument; one instrument often employs a conical, a cylindrical, and a bell section all working together. Of course, this brings extra complication to the task of understanding the resonant structure of such instruments, but that's also what brings about the unique timbre of the brass instruments.

As with the woodwinds, so with the brass ... there are extra features that bring about a more rich and complicated resonant structure than simple tubes would predict. Brass instruments employ valves, mouthpieces, and bells.

The main instruments of the brass family consist of the trumpet, trombone, French horn, and the tuba. Figure 11.1 depicts the frequency ranges for these instruments. The trumpet, trombone, and French horn have long cylindrical sections with a short bell section and a short tapered mouthpiece section on either end. The tuba, on the other hand, is conical.

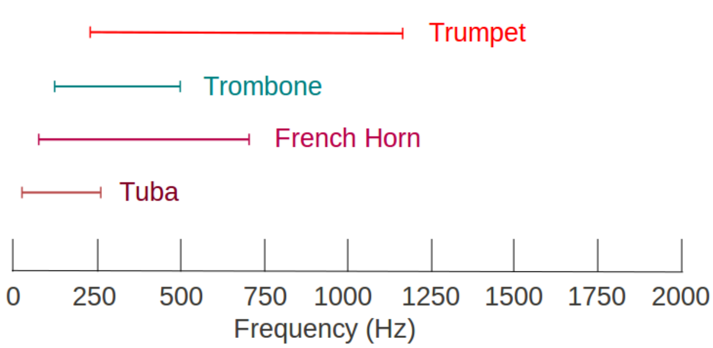


Figure 11.1: Frequency ranges of the various brass instruments

Figure 11.2 shows the three main sections of a typical brass instrument, and table 11.1 shows the typical lengths for these sections for the three instruments containing these sections.

A full understanding of the intonation and resonant characteristics of brass instruments is very complicated, and so we shall cover briefly some of their main features, to get a sense of how their intonation derives from the specific construction and section lengths. We'll begin with a discussion of the trumpet.

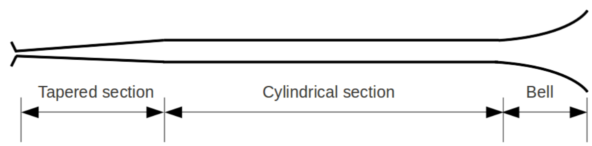


Figure 11.2: Three main sections of the trumpet, trombone, and French horns, a tapered mouthpiece section, a long cylindrical section, and a short bell section.

11.1 The Trumpet

The lip end of the instrument forms a closed end. While the end is actually "open" so that performer can blow into it, for resonant vibration purposes it acts like a closed end when the lips activate the resonant vibration in the column, in that it corresponds to an antinode of vibration. As the performer blows into the end, there is a "feedback" relationship between the frequency of the lips and the frequency of the standing wave resonance in the column. That is, as the blowing lips vibrate, they set the air column into vibration, which establishes its own resonant frequency based on the column geometry, which guides the vibrating lips into adopting the same frequency so that they vibrate in unison. In this way the performer's lips are locked into the right vibrational frequencies governed by the column resonance. As the pursing of the lips and air pressure are changed (the equivalent of overblowing for the wind instruments), different resonances supported by the air column can be activated by a similar "locking in" procedure. The performer can also bend the pitch somewhat by "lipping up" or "lipping down" a note by forcing the air column to vary its vibrational frequency somewhat and therefore vary the pitch. This is made possible by the fact that the mass, inertia, and muscles of the lips can exert a large influence on the air column. In woodwind instruments this lipping up or down to bend the intonation of a note is much more difficult to achieve, since it is the reed, and not the lips, that is setting the air column into vibration. The mass of the reed is small and therefore less able to influence the more massive air column, and therefore intonation of the instrument.

Rather than *predicting* the resonant structure of the trumpet, which is beyond the scope of our current ability, given the combination of cylindrical, conical, and bell sections, let's start by looking at the actual resonant structure to see if we can make sense of it. The family of resonant frequencies able to be supported in a B$^\flat$ trumpet (with all the vales open) is shown in table 11.2. First a brief word about valves. All valves open means that the air column is as short as it can be, and therefore the family of resonant vibrational frequencies is at its highest. When a trumpet valve is closed, a small section of tubing is added to the instrument (see the right side of figure 11.3), thus making it longer, and therefore resonant at a lower frequency. Depressing trumpet valves therefore lowers the pitch of the instrument.

Table 11.2 lists the frequencies of the blown resonances in the trumpet. Each of the frequencies looks as if it belongs to a family of harmonics, whose fundamental tones has frequency 115 Hz. However,

Table 11.1: Section lengths (in %) and total length (in cm) for three brass instruments. These numbers are only approximate since they vary by manufacturer.

Section	Trumpet (%)	Trombone (%)	French Horn (%)
Mouthpiece section	21	9	11
Cylindrical section	29	52	61
Bell section	50	39	28
Total length (cm)	140	264	528

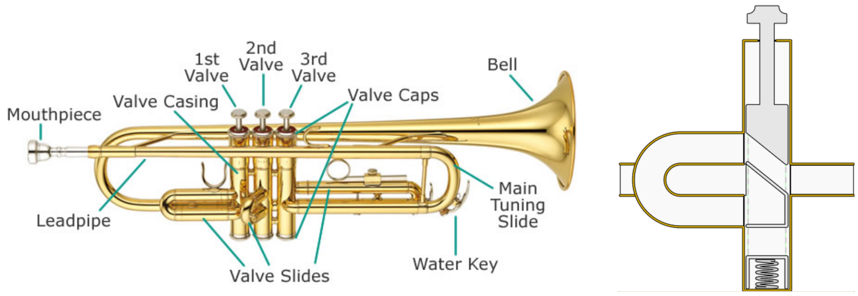


Figure 11.3: Schematic of a trumpet (left) and detail of a valve (right, shown depressed). When the valve is open (button up), the small section of extra tubing is bypassed. When the valve is depressed (button down) the extra tubing is added to the length of the overall tube length, increasing the wavelength and therefore decreasing the frequency and dropping the pitch.[1]

this fundamental tone is not supported by the trumpet. Since all of the harmonics that *would* belong to this fundamental are present, but the fundamental is not, it is referred to as a *missing* fundamental. How does this harmonic structure come about?

11.1.1 Effect of the Bell

Let's consider the geometry of the instrument to see if we can make some sense of this harmonic structure. The total length of the trumpet is approximately 140 cm. Given that the tube is closed on one end, we would expect a fundamental wavelength $\lambda_1 = 4(1.40 \text{ m}) = 5.6$ m, corresponding to a frequency of about 62 Hz. We would expect odd harmonics also to be present, at frequencies 186, 310, 434, 558, 682 Hz, and so on. Now consider the addition of the bell section, which has the effect of altering these frequencies. The bell is at the end of the instrument, and in total represents a rather large percentage of the total instrument length (see table 11.1). The bell defines the end of the instrument, and unlike a conical section that increases its radius linearly along its length, the bell increases its radius in more of an "exponential" fashion. The question naturally arises as to where, exactly, the effective end of the instrument lies, that is, where the pressure node for the supported resonances lies.

The physical explanation for where the effective end of the instrument is located is rather complicated, but the result is that it very much depends on frequency. The effective end of the instrument for low frequencies is deep within the instrument, near where the bell begins to flare outward. For the very lowest frequency, it is as if the bell has no effect, and the end corresponds to the end of the cylindrical section. For medium frequencies, the effective end is further down into the bell section, and for highest frequencies the effective end is much further out toward the end of the bell where it flares outward quickly (see figure 11.4). In essence, the higher the frequency, the more the standing wave penetrates further into the bell so that the effective length of the instrument *increases with increasing*

Table 11.2: Blown natural column frequencies of a B$^\flat$ trumpet, and their relationship to a "missing" fundamental at 115 Hz.

Blown Frequencies (Hz)	Breakdown of Frequencies
231	$231 \approx 2 \times 115$
346	$346 \approx 3 \times 115$
455	$455 \approx 4 \times 115$
570	$570 \approx 5 \times 115$
685	$685 \approx 6 \times 115$

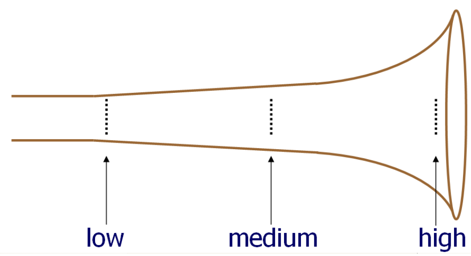


Figure 11.4: The effect of the bell on defining the end of the instrument for various frequencies. Note that higher frequencies penetrate further into the bell so that the effective length of the instrument increases with increasing frequency. The net result of this is that the higher frequencies are progressively flattened.

frequency. The net result of this is that the higher and higher frequencies experience a longer and longer instrument, and are progressively *flattened,* or lowered, in pitch. The family of harmonics which "began" as a typical harmonic series of a semiclosed tube becomes, with the addition of the bell, more compressed, with higher frequencies flattened in pitch considerably.

11.1.2 Effect of the Mouthpiece

The trumpet mouthpiece is fixed to its end, and consists of a small tapered section with a cup-like opening, and adds extra length to the instrument. As in the case of the bell, lower frequencies are changed very little, while higher frequencies experience a longer instrument. When the wavelength of the resonance is long compared to the physical length of the mouthpiece, the added length to the instrument corresponds to the effective length of a cylindrical mouthpiece sharing the same volume as the trumpet mouthpiece (as was the case for the woodwind instruments). For higher resonances with wavelengths comparable to the size of the mouthpiece, it begins to have its own resonances as well, which make the total tube length look longer.

11.1.3 Sound Spectrum Compression

The result of the harmonic series compression that results from the addition of the bell and the mouthpiece is shown in 11.5. Notice now that the usable resonances (starting with the second peak in the figure) are located at what look like integral multiples $2f_1$, $3f_1$, $4f_1$, *etc.* of a fundamental $f_1 = 115.5$ Hz that is not actually present. Hence the first usable resonances occur at do, sol, do', mi', *etc.* This sequence corresponds to the lowest note, a musical fifth above it, an octave, and octave and a third, *etc.*, giving rise to the familiar harmonic sequence used in "reveille" and "taps," and able to be played on instruments without valves. The pedal tone (the lowest peak in the figure) is located nearly a musical fifth below the missing fundamental and is out of tune with the other harmonics, and is thus rarely used in music.

11.1.4 Valves

Thus, the first usable resonance in the trumpet is the second harmonic of the missing fundamental, activated by setting the air into vibration by the lips through the mouthpiece. The next usable resonance is the third harmonic, activated by changing the vibration of the lips to strike this resonance. How are all of the semitones between these two harmonics played? As with all other instruments employing air column resonances, the harmonic frequencies of a brass instrument are determined by

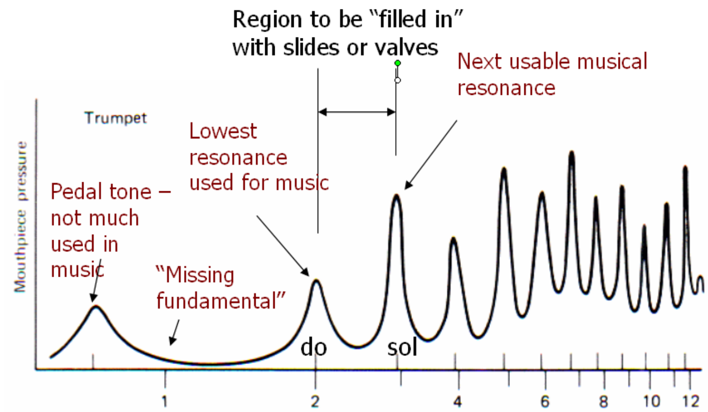


Figure 11.5: The sound spectrum for the trumpet. Usable resonances appear as the second, third, fourth *etc.* resonances of a fundamental that is not present.[2]

the length of the air column. Thus, by changing the length, different notes can be activated. Since the second and third harmonics are separated by a musical fifth, which corresponds to 7 semitones, the trumpet needs an additional 6 "lengths" to bridge the gap between these resonances in order to play all the notes of the musical scale.

Figure 11.6: The valves on a trumpet. Activation of the valves in various combinations enables all semitones of the musical scale between the first two resonances to be played.[3]

As we will learn in chapter 14, a semitone change in pitch corresponds to a change in frequency of approximately 6%. Therefore in order to lower a note by one semitone, we would need to increase the length of the air column by a factor of 1.06. With all possible combinations of the three valves on the trumpet, the pitch can be changed one semitone at a time by adding extra small sections of length to the instrument (see figure 11.6) to bridge the gap between the second and third harmonics. With each valve combination, and thus each new note, a whole new family of harmonics becomes available (accessible by the proper lipping), allowing all notes to be played over the entire pitch range of the instrument. However, not all valve combinations produce exactly the right pitches, and some compensation on the part of the player's lipping is required to achieve good intonation.

For example, depressing the middle valve increases the length of the instrument by 6%, lowering the pitch by one semitone. Depressing the first valve increases the length of the instrument by an *additional* 6% over the middle valve length, lowering the pitch by two semitones, or one full tone. Pressing both the first and the second valves *should* therefore lower the pitch by three semitones. As the following example illustrates, the action of these two valves together does *not,* in fact, produce the exact 3 semitone drop.

Example 11.1

Trumpet Valves and Intonation *Given a trumpet length of 140 cm with all valves opened, what added length should be provided by the second valve in order to lower the pitch by one semitone? What length should be added by the first valve to lower the pitch by a full tone? When both the first and second valves are depressed, what length results, and how does this compare to what would be needed for exactly three semitones?*

Solution: When valve 2 is depressed, the length is increased by 6%, or 0.06(140 cm $\approx$ 8.4 cm. Thus the length for one semitone below the first note is 148.4 cm. To lower the pitch by another semitone, we now need to compute the addition to be added to this new length, which is 0.06(148.4 cm) $\approx$ 17.3 cm. Thus the length for one full tone below the first is 157.3 cm. With both the first and the second valves depressed, the total added length to the instrument is $8.4 \text{ cm} + 17.3 \text{ cm} = 25.7$ cm. The question is, does this produce three semitones below the first note? For a drop of three semitones, the length of the instrument should be increased by $0.06(140) + 0.06(148.4) + 0.06(157.3) \approx 26.74$ cm, which is longer than the length actually added by the first two valves. This makes the result slightly sharp, requiring the performer to lip down on the note to achieve proper intonation.

11.2 The Trombone

The length of the trombone is approximately 270 cm. The trombone changes pitch by operation of a slide piece that produces a change in length of the air column. See figure 11.7 for a depiction of the slide positions and the notes they produce. With each slide position, a whole new family of harmonics becomes available, as in the case for the trumpet. Thus by operating the slide at 7 seven different positions, all eight semitones between the first note and the fifth above it can be played. After that, the same slide positions are used while overblowing to play higher notes. Each time the pitch is dropped by a semitone, the length of the column must be lengthened by 6%. As the pitch is lowered by *successive* semitones, the length of the instrument increases, so that 6% of an increasing length also itself increases. This is why the spacing between the slide positions increases as the pitch is lowered.

Example 11.2

Trombone Slide Positions *The trombone length is 270 cm. How far out must the slide be moved in order to drop the pitch by one semitone? How far must the slide be moved out in order to drop the pitch by one full tone?*

Solution: In order to drop the pitch by one semitone, the increase in the length of the trombone air column must be 0.06(270 cm) $\approx$ 16 cm. This can be accomplished by moving the slide out 8 cm, since it has both an upper and a lower section that move out from the main tube. In order to decrease the pitch by one full tone, the length must be increased by an additional $0.06(286) = 17$ cm. This means moving the slide out an additional 8.5 cm. Thus for a drop in pitch of one semitone, the slide must be moved out a total distance of 8 cm + 8.5 cm = 16.5 cm.

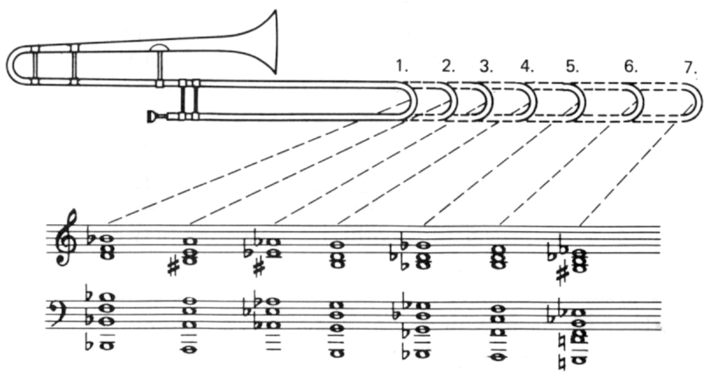


Figure 11.7: Schematic of the trombone, with slide positions shown and their corresponding notes.[4]

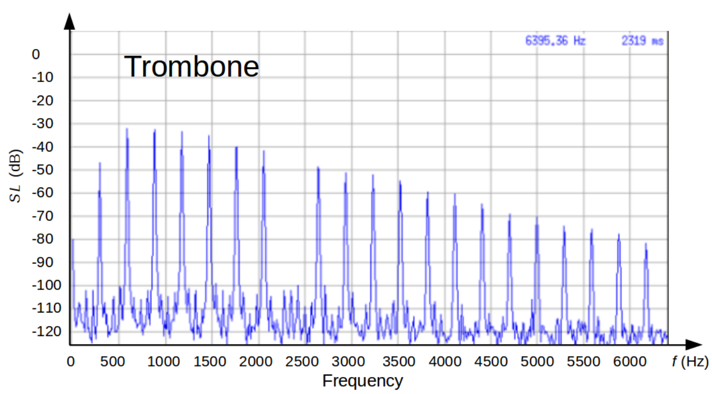


Figure 11.8: Spectrum of trombone resonances. [2]

11.3 The French Horn and Tuba

The French horn utilizes approximately sixteen resonances, and in the upper pitch range these resonances are very closely positioned (see figure 11.9). This means that as a performer is hitting notes in the upper register of the instrument, he or she needs to apply just the right lipping to strike the appropriate resonance. Since the resonances are so close to one another, it is easy to jump from one to another with a very subtle change of the lipping. This explains the very “slippery” intonation of the French horn, and what makes it such a difficult instrument to play in its upper frequency range.

The tuba is the low-voiced member of the brass family. It is a valve instrument which comes in 3, 4, or 5 valve varieties.

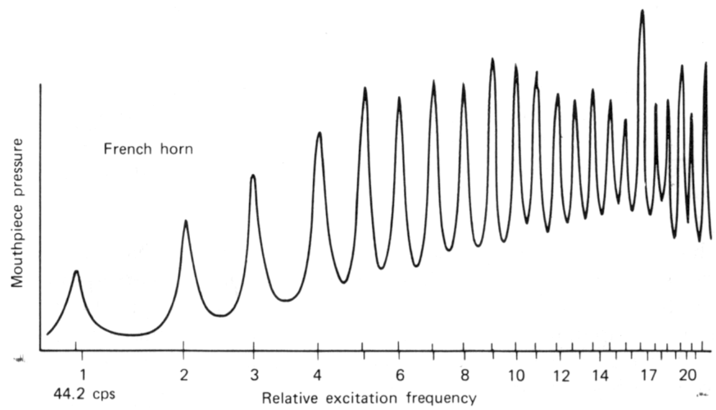


Figure 11.9: Spectrum of French horn resonances. [2]

11.4 Sound Radiation

All brass instruments radiate all the sound from the bell, so that the sound emanating from the instrument is very directional. Mutes are hollow devices that have resonant properties of their own, which can be placed in the bell to significantly alter the timbre of the instrument by being played at or near their resonance.

11.5 Chapter Summary

Key Notes

- Several of the brass instruments include cylindrical and conical sections, as well as a mouthpiece and a bell-shaped end. The lips activate the air column and exert considerable control over the frequencies it vibrates at. This combination of shapes gives the brass instrument very complicated resonant frequency structure.
- **The trumpet**: In figure 11.5 the sound spectrum of the trumpet is shown, where the fundamental is actually missing, and the first usable resonance is located at the second harmonic. (The peak approximately a fifth below the fundamental is the "pedal tone" and is rarely used in performance.)
- Thus, the first usable resonance is the second harmonic, activated by setting the air into vibration by the lips through the mouthpiece. The next usable resonance is the third harmonic, activated by changing the vibration of the lips to strike this resonance. How are the notes between these two harmonics played? The harmonic frequencies of a brass instrument are determined by the length of the air column.
- Thus, by changing the length, different notes can be activated. Since the second and third harmonics are separated by a musical fifth, which corresponds to 7 semitones, the trumpet needs an additional 6 "lengths" to bridge the gap between these resonances and thus to hit all the notes of the musical scale.
- With combinations of the three valves, the pitch can be changed one semitone at a time to bridge the gap between the second and third harmonics. With each valve combination, and thus

change of length, a whole new family of harmonics becomes available, allowing all notes to be played over the entire pitch range of the instrument. However, not all valve combinations strike exactly the right pitches, and some compensation on the part of the player is required to achieve good intonation.

- **Trombone**: In a trombone, the slide piece allows a change in length of the air column. With each slide position, a whole new family of harmonics becomes available, as in the case for the trumpet. For the trumpet and trombone, approximately 8 or 9 resonances can be used.
- **French horn**: The French horn, on the other hand, utilizes approximately sixteen resonances, and in the upper pitch range these resonances are very closely positioned. This explains the very "slippery" intonation of the French horn, and what makes it such a difficult instrument to play in its upper frequency range.
- **Tuba**: The tuba is the low-voiced member of the brass family. It is a valve instrument which comes in 3, 4, or 5 valve varieties.
- All brass instruments radiate all the sound from their bell, so that the sound emanating from the instrument is very directional. Mutes are hollow devices that have resonant properties of their own. They can significantly alter the timbre of the instrument by being played at or near their resonance.

Exercises

Questions

1) How does the bell of the trumpet affect the resonance structure of high frequencies versus low frequencies? What is the basic consequence of this for the trumpet's sound spectrum?
2) Why is it that the bell on the end of woodwind instruments is relatively unimportant in determining the sound of the instrument for most of the notes that it plays, whereas for brass instruments the bell is important for all notes it plays?
3) Explain why the French horn is difficult to play in its upper range where it is easy to slip from one note to another.
4) Why is it easier on a brass instrument to "lip down" or "lip up" a note than on a woodwind instrument?
5) Briefly, why does pressing the middle and first valves of the trumpet not accomplish a drop in pitch of exactly three semitones when the middle valve drops the pitch one semitone and the first valve drops it two semitones?
6) Why are only three valves necessary to play all the semitones between the first and second tube resonances in the trumpet when 6 holes are necessary for the same function in the flute?

Problems

1. Why does the bell section of a trumpet affect high frequencies more than low frequencies?
2. Assume that the length of the trombone is 270 cm when the slide is in the first position. How far would the slide need to be

moved out to lower the pitch by six semitones?

3. The trombone's low $B_2^\flat$ is played with the slide in the first position (i.e. shortest length). a) What is its effective length – that is, the length of an idealized semi-closed tube with the same second-mode frequency? (Remember that the "fundamental" is missing for brass instruments). b) Estimate how far the slide should be moved to reach the second position, in order to play A_2, remembering that each cm of slide motion adds 2 cm to the tube length. c) If already in the sixth position (F_2), how much and in what direction must you move the slide in to reach the seventh position (E_2). How about to reach the second position (A_2)?

4. The first valve depressed alone on the trumpet drops the pitch by a full tone, the second drops it by a semitone, and the third drops it by three semitones. With all three valves depressed, what discrepancy would you predict between the actual and ideal lengths added to the air column?

5. From figure 11.9, explain why beginning French horn players often have trouble hitting the correct frequency.

References

[1] http://www.musicshowcaseonline.com/resources_instrumentCareBrass.asp

[2] The Acoustical Foundations of Music, John Backus, W.W. Norton & Company, 1969

[3] http://www.dreamstime.com/stock-photo-trumpet-valves-image2649960

[4] The Trumpet and Trombone, Philip Bate, W.W. Norton & Company, Inc., 1966

Chapter 12

The Piano

12.1 The Modern Piano

The majority of instruments that derive sound from strings are either plucked or bowed. The modern piano is one of the few exceptions to this. It is classified as a percussive instrument, since sound derives from the striking of strings by felt hammers. Forerunners of the piano include the clavichord (where the strings are also hammered, as in the modern piano) and the harpsichord (where the strings are plucked by a quill). Both of these instruments have very feeble sound output compared with the modern piano. One of the drawbacks of the harpsichord is that the mechanism that plucks the string results in one level of loudness, no matter how softly or firmly the key is pressed. This makes it difficult for the performer to add expressive content to a performance since dynamic variation in the sound is not an option.

Early versions of the piano derived from the clavichord. As music composed for the piano evolved throughout the 18^{th} and early 19^{th} centuries, it placed greater demands on the instrument. In particular, performers (including the legendary composer and pianist Franz Liszt) played with greater vigor and energy, resulting in the physical failure of several instruments. This required a much sturdier and more robust design. In the year 1855 Henry Steinway developed the first modern piano, based on a cast-iron frame. The new stronger construction introduced by Steinway gave the piano a much louder, brighter tone and much greater expressive dynamic. Its overall power output completely surpassed its predecessors, and music composed for the piano moved into the modern era.

The full name of the piano is "piano-forte," which in Italian means "softly-loudly," celebrating the large dynamic range, and therefore large expressive range, achievable with the instrument. The frequency range covered by the piano is very large – over seven octaves in all, from A_0 at frequency 27.5 Hz to C_8 at frequency 4186 Hz. It has 88 black and white keys.

12.1.1 Sound Mechanism

When a piano key is pressed, it causes the mechanism called the *action* to sling a felt hammer toward a particular string, while simultaneously lifting the damper from that string. The piano string is therefore set into vibration by the percussive impact of the hammer. The felt on that hammer cannot be too hard (resulting in a harsh tone) or too soft (resulting in a weak muffled sound). Just before the hammer hits the string, the damper is lifted so that the string string can vibrate freely. While the key is held down, the note continues to sound. As soon as the key is released the damper is set back on the string and the sound stops.

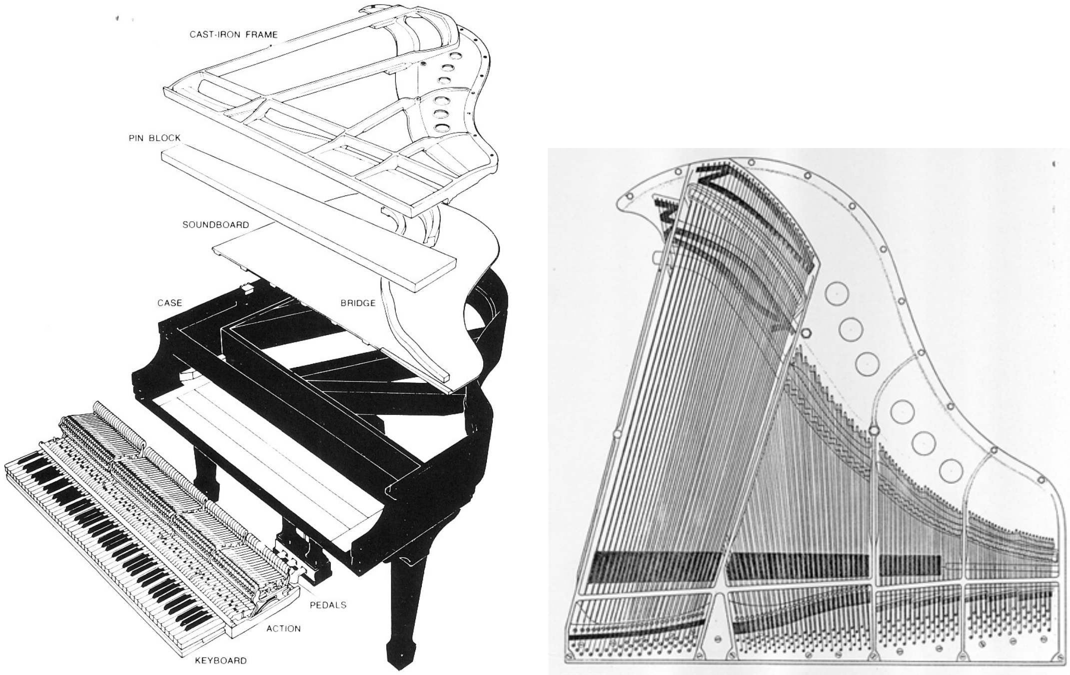


Figure 12.1: On the left, a view of the piano components, including the cast-iron frame, the soundboard and bridge, the case, and the keyboard and action. On the right, a view of the cast-iron frame with the strings and tuning pins. Lowest notes are on the left and highest on the right.[2]

12.1.2 The Strings

The lowest 10 notes on the piano have one steel string each that is heavily wound in copper to give it large mass per unit length. The next 18 notes have two strings each, also wound in copper. The remaining notes consist of three strings each, all the way to the top range of the instrument. The total tension of all the strings in a concert grand piano can be as high as 60,000 pounds (this amounts to the weight of approximately 15 automobiles!). The strings are stretched between a tuning peg on the front end of the piano, over a bridge, and anchored to the back end of the cast-iron frame (see figure 12.1). Since the strings are tightly pressed down against the bridge, they are able to transfer their vibrations directly to the sounding board. The tuning pegs are set into a block that consists of as many as 41 layers of cross-grained hardwood, and must hold the high tension of the strings constant so that intonation does not change as the instrument is played. The bridge, which is mounted directly to the large soundboard, is typically made from Norway Spruce, and serves to translate the vibrations of the strings to the soundboard.

12.1.3 The Action

The string is set into vibration by the percussive action of a felt hammer. The mechanism that causes the hammer to strike the string (and also to lift the damper off the string), called the action, is very complicated and involves a great number of parts. The action serves to "throw" the hammer at the string so that it is in "free fall" when it hits. After it bounces off the string, it is caught by the backcheck about halfway back to its resting spot. It mustn't bounce off the backcheck or else it might return to hit the string a second time. The backcheck must catch it without making noise, and make it available for another return to the string if the music calls for the same note to be played in rapid succession. There is much more detail to the function of the action that we will not cover, but suffice it to say that the piano action is the product of considerable evolution over the years. Figure 12.2 shows schematics of the action for both grand and upright pianos. In the grand piano the hammer is located below the horizontal strings, and is therefore flung upwards to strike the string. The advantage of this geometry is that the rebound of the hammer is naturally assisted by gravity, so that the action can more readily accommodate the playing of single notes in rapid succession. In upright pianos the strings are fixed vertically, and as a result, the hammer is flung horizontally to strike the strings. Without the assist of gravity, the action for upright pianos tends to be less responsive for this rapid repetition of single notes.

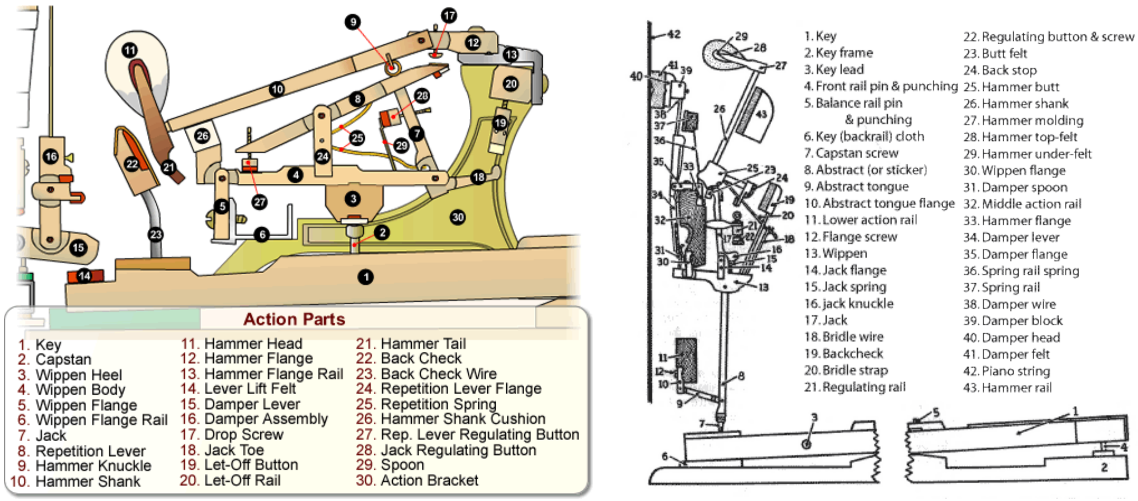


Figure 12.2: Detail of the action of a grand piano (left) and an upright piano (right). Numerous parts are involved in slinging the hammer toward the string, catching it on its return, and keeping it ready to strike again in rapid succession. In the grand piano the hammer moves *upward* to strike the string, so that gravity naturally assists with pulling the hammer away, enabling a faster more responsive action than for the upright piano, where the hammer is slung horizontally. [1]

12.2 String Inharmonicities

The origin of sound in a piano is a vibrating string. We've learned earlier that the frequencies on a vibrating string are given by

$$\boxed{f_n = \frac{n}{2L}\sqrt{\frac{T}{\mu}} \qquad n = 1, 2, 3, \ldots} . \tag{12.1}$$

When a string is set into vibration, the family of vibrational modes that result are integral multiples of the fundamental frequency. Figure 5.3 from chapter 5 shows the shape of these harmonic modes. Note that with increasing harmonic number there is increasing curvature in the string shape. If the string is flexible enough to bend with the curvature of the n^{th} harmonic, then the frequency with which it vibrates is nf_1. Strings used on the guitar or on any member of the stringed instrument family are quite flexible, and could be easily wound around two or three fingers. The harmonics they produce have frequencies in strong agreement with those predicted by equation 12.1.

In contrast, piano strings are quite stiff by comparison, and have much greater mass per unit length.

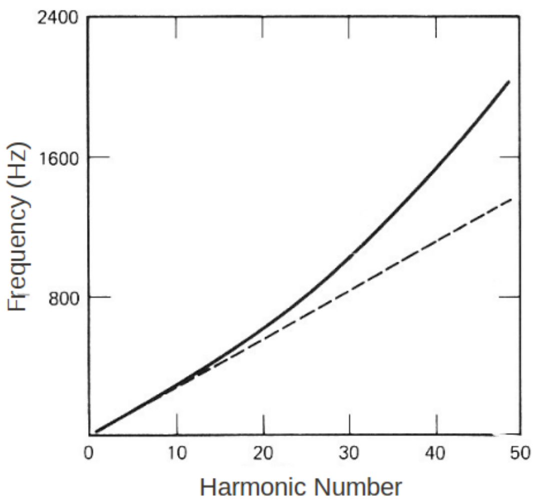


Figure 12.3: The harmonics of the lowest note on the piano, A_0. The dotted line represents the frequencies of the harmonics (y-axis) that would normally correspond to the harmonic number (x-axis). The solid curve represents the actual frequencies corresponding to the harmonic numbers resulting from the inharmonicity produced by the stiffness of the piano string. For harmonic numbers above 10, the departure from the standard harmonic values becomes noticeable.[3]

When a piano string is set into vibration by impact of the felt hammer, an entire family of harmonics is set into vibration. The fundamental has the longest wavelength of them all, at a value twice the string length. The higher harmonics have smaller wavelength, and with increasing n the wavelength becomes very small compared to the string length. Because the piano string is so stiff, it is not able to be bent with the tight curvature required for these higher harmonics, and as a result it "snaps back" to equilibrium *more quickly* than it would if it were able to bend appropriately. As a result, the frequencies for these higher harmonics are pushed to higher values than those calculated by equation 12.1. The difference between the calculated and the actual frequencies is small for the lower harmonics and becomes increasingly larger with higher harmonic number n (see figures 12.3 and 12.4). As a result of this non-linear effect, the harmonics become *overtones*. The shorter the string, the less able it is to bend with the curvature of the higher harmonics, so that the effect is most most pronounced for the highest notes on the piano.

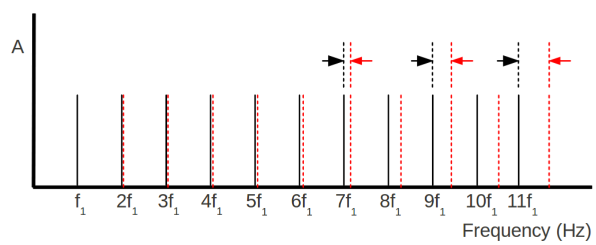


Figure 12.4: Sound spectrum of harmonic tone (black spectrum lines) and piano tone with inharmonicity. Note the red dotted lines, which are the overtones of the piano string, are shifted toward higher frequency from their "harmonic" values. The effect increases with increasing harmonic number. The effect is somewhat exaggerated in this schematic in order to illustrate the effect.

12.2.1 Stretched Octaves

Any octave on a piano is typically tuned in such a way as to eliminate beats resulting from the interaction of the second harmonic of the lower note and the fundamental of the higher note. Since inharmonicity causes the second harmonic of the lower note to be shifted slightly sharp, the higher note ends up being tuned slightly sharp as well, relative to the lower pitch's fundamental frequency. That it ends up sharp is not particularly noticeable to the human ear, but the effect magnifies over several successive octaves. Piano tuners will typically begin by tuning the octave at the center of the keyboard first, and then work their way up to the top and down to the bottom of the keyboard from that center point. The notes in the upper half of the piano become increasingly sharp, and notes in the lower half become increasingly flat, relative to the middle notes. The resulting difference in pitch between the very lowest notes and the very highest notes on the piano ends up being approximately 60 cents, or a little over $\frac{1}{2}$ of a semitone (one semitone of pitch difference is defined as 100 cents). While this may sound significant, the effect is not particularly noticeable when listening to music played on the piano. The red curve in figure 12.5 illustrates graphically the tuning shifts in cents over the range of the piano.

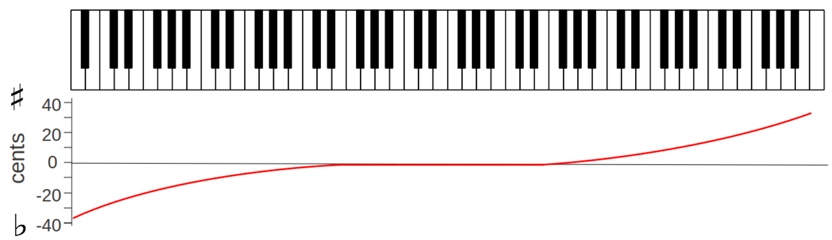


Figure 12.5: In order to account for the inharmonicity in the stiff piano strings, the low notes are tuned slightly flat (as much as approximately 30 cents) and the high notes are tuned slightly sharp (as much as approximately 30 cents). Note that 100 cents equals one semitone.

12.3 Sound from the Piano

The strings actually deliver very little acoustic sound to the air. Their vibrations are passed on to the soundboard via the bridge which is mounted directly to the soundboard. The soundboard, which vibrates resonantly with the string vibrations, then radiates the majority of the sound into the air. Opening the lid of the instrument helps project sound from the soundboard into the room. Some of the energy of the string vibrations reaches the soundboard through the space of air separating them.

12.3.1 Attack and Decay

The actual motion of the string is a superposition of both vertical and horizontal motion. The vertical component of string motion (*i.e.* vibrations perpendicular to the soundboard) transfers its vibrational energy to the bridge of the instrument, which is mounted directly to the soundboard. Absorption of these string vibrations by the soundboard causes this vertical mode of vibration to dampen quickly. The horizontal component of string motion (parallel to the soundboard) lingers longer, since it is much less efficient at transferring its vibrational energy to the bridge and therefore to the soundboard. The sound of the piano is characterized by a sudden, percussive burst of sound (called the "attack") produced by the action of the hammer on the string, followed by a relatively quick decay of sound from the vertical string vibrations, and a much slower decay of sound from the horizontal string vibrations. Figure 12.6 shows three diagrams illustrating the decay of sound from a piano. The first pane shows the overall decay, which consists of two components, a quick decay (arising from absorption of vertical string vibrations by the soundboard) and a slow decay (arising from the longer-lived horizontal string vibrations). Note that the y-axis is logarithmic, so that what looks like a linear drop-off in sound intensity is actual exponential.

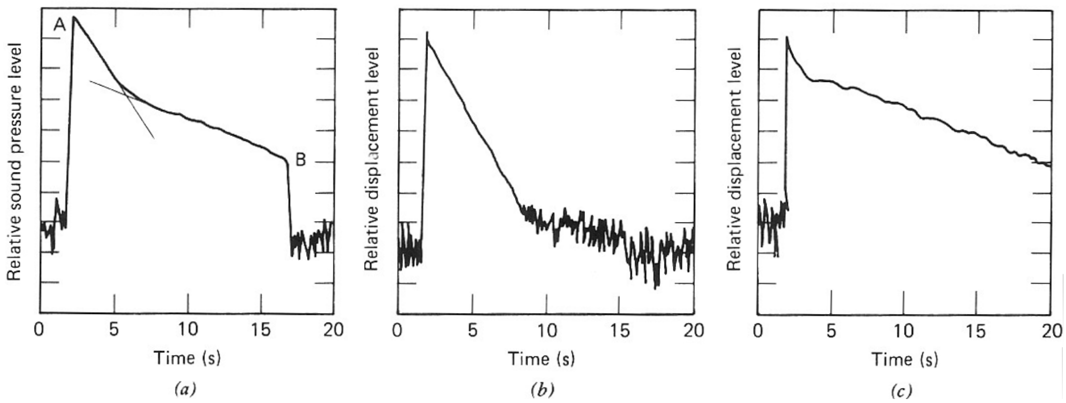


Figure 12.6: The sound from a struck piano string as it decays. The left pane shows all sound as it decays, and the middle and right panes depict each of the two modes of decay, from the string's vertical and horizontal vibrational motions, respectively. [4]

12.3.2 Beats and Warmth

One technique used in the tuning of the piano is to slightly mistune the three strings that make up a particular note so that a very slow beat results. It is thought that the presence of a long-lived beat can bring a sense of "warmth" to the sound. When a violin is played, the addition of slight vibrato brings warmth. When several violins play together, they are never quite in tune and the resulting sound has warmth owing probably to the long-lived beats they produce. Alternatively, if the three strings are brought into strict tune with one another, the sound lacks warmth. It is a common practice to slightly detune the trio of strings making up the piano strings to produce the sense of warmth pianos are known for.

12.4 Chapter Summary

Key Notes

- The modern piano, first developed by Henry Steinway in 1855, is based on a cast-iron frame, which gives it a brilliance and power unlike its predecessors. Forerunners of the piano include the clavichord (where the strings were hammered, as in the modern piano) and the harpsichord (where the strings were plucked). The full name of the piano is "piano-forte," which in Italian means "softly-loudly," celebrating the large dynamic range, and therefore large expressive range, achievable with the instrument.
- The piano embraces over seven octaves of pitch (from A_0 of frequency 27.5 Hz to C_8 of frequency 4186 Hz) with 88 black and white keys. The lowest strings are single and double strings that are each over-wound with another string.
- The middle and upper range notes have three strings each. The total tension of all the strings in a concert grand piano can be as high as 60,000 lb.
- The piano string is set into vibration by the percussive impact of a hammer. The action of the piano is the mechanism that sets the hammer (and the dampers) into motion. In some pianos the action consists of more than 7000 individual parts.

- The origin of sound in a piano is a vibrating string.
- The greater the tension T, the faster the string snaps back toward its equilibrium position, and therefore the higher the frequency of sound that results. What makes piano strings different than those in other string instruments is that they have significant stiffness, which adds an additional contribution to the speed with which the string returns toward equilibrium.
- Strings with such significant stiffness therefore vibrate with a higher frequency than predicted by the string harmonies equation, and is more pronounced for higher frequencies than for low.
- The effect of stiffness on frequency increase is related to the square of the frequency, so that $2f$ (i.e. the octave) will be shifted four times more than the fundamental frequency f. Therefore, the higher frequency components of a given note end up more like overtones than harmonics. The presence of these overtones (inharmonicities) help give the piano its unique sound.
- The strings actually deliver very little acoustic sound to the air. Their vibrations are passed on to the soundboard via the bridge. The soundboard then radiates the sound into the air.
- The piano sound is characterized by a sudden, percussive burst (its "attack"), followed by a slower decay of sound. There are actually two decay rates present, one longer lived than the other. The shorter is produced by vibrations perpendicular to the soundboard, whose energy is lost more quickly than vibrations that are parallel to the soundboard. In addition to this, the various overtones of a particular note have different decay rates, making the timbre of the piano change as the note dies away.

Exercises

Questions

1) What quality of piano strings causes them to vibrate at higher frequencies than are predicted by the standard frequency formula for string vibrations? What is the effect of these higher frequencies on the timbre of the instrument?

2) Why is it that the decay of sound from a piano has two components, one that dies out quickly and the other that dies out more slowly?

3) Why does a properly tuned piano need to be tuned with "stretched octaves"?

4) Are the higher modes of vibration for piano strings better characterized as harmonics or overtones? Explain briefly.

Problems

1. A piano string is 140 cm long with a linear density of 0.042 kg/m. The string is stretched with a tension of 2000 N. The hammer is designed to strike the string at a distance equal to $\frac{1}{7}$ of its length from one end. What is the maximum time that the hammer can be in contact with the string if we assume that it must break contact before the wave reflected from the far end returns to the point of contact?

2. A piano wire is very stiff, such that when first struck with large amplitude, the frequency of its oscillation is *higher* than

when its amplitude has nearly died out. Depict the time dependence of the string's amplitude. What might be the musical consequence of such behavior?

3. Explain why, for two strings of the same diameter, mass density, and tension, the shorter one shows more inharmonicity than the longer one in its vibrations.

4. Suppose you have a piano bass string of length $L = 2.0$ m and linear mass density $\mu = 0.060$ kg/m, and you want it to vibrate with pitch D_1. How much tension must you apply?

5. Explain, in terms of the vibrational modes affected, why the damper felts in a grand piano are located immediately above the hammers instead of elsewhere along the strings.

6. Suppose that all the strings of a piano were of the same material and also had the same diameter and tension. If the longest string (A_0) were 2 m in length, how long would the highest A-string (A_7) need to be? Is this practical? Explain briefly.

7. Show that two unison strings, tuned 2 cents (0.12%) different and initially in phase, will fall out of phase and therefore cancel one another after about 400 vibrations.

8. A piano tuner finds that two of the strings tuned to C_4 give about one beat per second when sounded together. What is the ratio of their frequencies? Show that their pitches differ by about 7 cents. (Reminder: one cent is $1/100^{th}$ of a semitone and corresponds to a frequency ratio of approximately 1.0006.)

References

[1] http://wengleemusic.com/blog/archives/120, http://www.teach-me-piano.com/piano-action.html

[2] The Physics of the Piano, E. Donnell Blackham, Scientific American, 1965; http://www.zainea.com/pint.gif

[3] John. S. Rigden, Physics and the Sound of Music, pg. 191, Wiley (1982)

[4] John. S. Rigden, Physics and the Sound of Music, pg. 192, Wiley (1982)

Chapter 13

The Human Voice

We use our voice so much, it might be easy to take it for granted. After all, more often than not we speak to one another without even considering the sheer marvel of human speech, or our ability to formulate audible variations and produce specific patterns of sound that communicate truth and meaning. The human voice is the oldest, as well as the most versatile, complex, and amazing of all musical instruments. While it shares some features with other instruments, its complexity and flexibility certainly sets it apart, and there are still aspects of it that are not well understood from a scientific standpoint. The instrument of the human voice most certainly predates any fabricated instruments.

13.1 Vocal System Components

All instruments have (in fact require) three components: an energy source, an oscillator, and a resonator, and the voice is no exception. The lungs constitute the first of the three, the energy source in the form of forced air that moves through the vocal folds, into the vocal tract, and out of the mouth. The vocal folds, a pair of ridges of soft tissue located at the top of the trachea, constitute the second component, the oscillator; they oscillate when set into motion from the forced air. Finally, the larynx, along with the pharynx, mouth, and nasal cavity, make up the vocal tract, which constitutes the third essential component, the resonator (see figure 13.1.

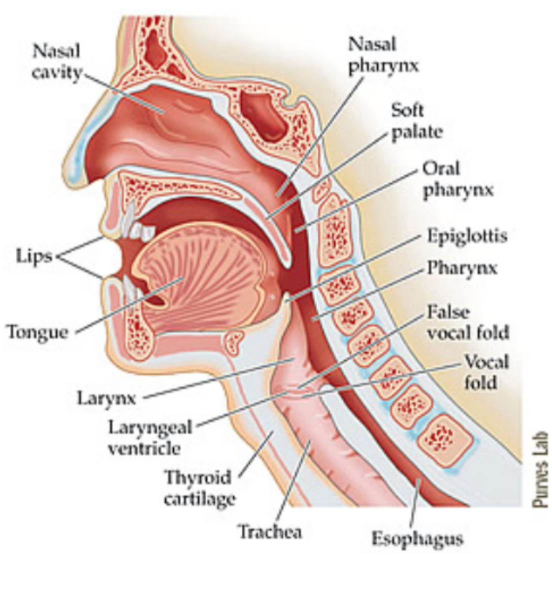


Figure 13.1: (***For color version see Appendix E***) The anatomy of the human vocal system. Air supplied by the lungs is forced up the trachea, and the vocal fold opens and closes at high frequency to modulate the moving air into sound waves. The resonant character of the vocal tract (consisting of the larynx, pharynx, mouth, and nasal cavity) determines the specific timbre of the spoken sound. Changes in the shape of the vocal tract cause changes in the resonant character and hence in the timbre of the spoken sound. [1]

13.1.1 Vocal Fold Operation

The vocal folds are located at the bottom of the larynx, a small boxlike cavity situated at the top of the trachea (the tube that leads to the lungs). The opening of the vocal folds is called the glottis, and the epiglottis, located above it, closes off the larynx when food is swallowed. The vocal folds are controlled by muscles and by the air flowing through them from the lungs. For sounds produced by the voice to have identifiable pitch, there needs to be *periodicity.* So the question naturally arises ... how do the vocal folds turn a continuous stream of air from the lungs into a periodic vibration that can be turned into sound? Surprisingly, the answer is connected to the reason that airplanes are able to fly!

There is a well-known principle of fluid dynamics called "Bernoulli's principle," which says that an increase in the speed with which a fluid (or air) flows corresponds to a decrease in pressure. Figure 13.2 shows the pattern of airflow past an airplane wing. Air flows over the top and the bottom of the wing. Because of the curved shape of the top, air flows farther over the top than underneath the wing. Since the air flowing above and below the wing makes the trip from the front to the back of the wing in the same amount of time, the air that goes above the wing travels *faster* than the air that goes beneath the wing. The result of this faster air on top means that the pressure above the wing is *lower* than below the wing. This pressure differential gives the plane the necessary *lift* to get off the ground, in a sense "sucking" the wing upward.

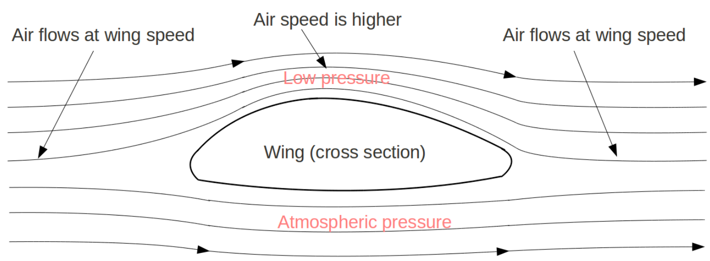


Figure 13.2: Airflow over the top and bottom surfaces of an airplane wing. Since air flowing over the top travels a farther distance than air flowing under the bottom of the wing, and does so in the same amount of time, it moves faster, resulting in a pressure drop at the top of the wing. The pressure differential between the top (lower pressure) and bottom (higher pressure) of the wing gives the plane the necessary lift to enable it to fly.

Figure 13.3 Shows a mechanical analog to the vocal fold operation. The mass M is held back by a spring from closing the gap at B. When air flows down the line, it first passes through the wide channel at A, and then through the narrower channel at B. The narrower constriction on the flow at B causes the air to speed up at that point. As a result of this increase in speed, the pressure drops at B. The drop in pressure "sucks" the mass M downward to close off the passage, at which point the airflow ceases. Pressure immediately builds up at A, eventually forcing the passage open again, at which point air flows through point B, causing the pressure to drop, pulling the mass back down again, *etc*. The periodic opening and closing of this passage results in a series of air pulses that pass to point C and beyond. Because of the sizable inertia of the mass M, the frequency of these pulses will be quite low – too low to produce pulses of sufficiently high frequency to constitute sound.

Now consider the operation of the vocal folds. The lungs provide pressure in the trachea below the vocal folds. When they are held closed by the vocal fold muscles, air pressure builds up beneath them. When the pressure is sufficient to force them open, air flows up through them at high speed, causing a drop in pressure, pulling them closed again. Once closed, pressure builds up again beneath them. This cycle of pressure buildup, vocal fold opening, air flow, pressure drop, and vocal fold closure, produces periodic air pulses that move into the vocal tract as sound waves. The frequency of the wave depends on the tension in the vocal fold muscles and the pressure of the air supplied by the lungs. And the frequency of the vocal folds directly determines the pitch of the tone being spoken or sung. A singer controls her pitch by adjusting the tension in the vocal folds, which adjusts the resonant frequency of their vibration.

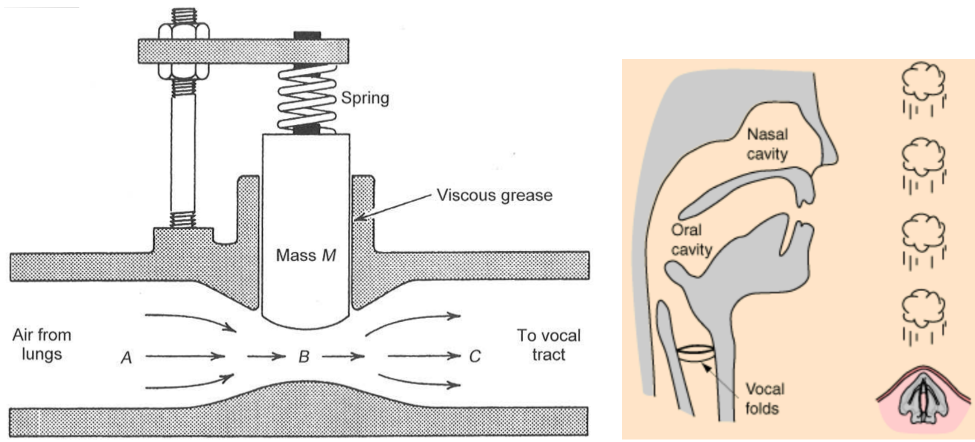


Figure 13.3: On the left, a mechanical analogy for the operation of the vocal fold. The mass is pulled down by the decrease in pressure at point B owing to the higher air flow speed. When the passage is closed off by the mass, pressure builds up at A and eventually forces the passage back open. On the right, depiction of pressure pulses that arise from the opening and closing of the vocal folds under pressure from the lungs and Bernoulli's principle. [2],[3]

13.1.2 Vocal Tract Resonances

The vocal tract consist of several chambers, including the larynx, the pharynx, the mouth, and the nasal cavity. It can be very roughly approximated as a cylindrical column of air closed at the vocal fold end (which corresponds to an antinode of vibration since this is the origin of the pulses moving into the vocal tract). We can calculate the resonant frequencies of this column, recognizing that it supports odd integral multiples of the fundamental:

$$\boxed{f_n = \frac{n}{4L} 20.1\sqrt{T_A} \qquad n = 1, 3, 5, \dots} \tag{13.1}$$

The effective length of this column for an adult male is approximately 17 cm, for which the first three resonant frequencies are at 500 Hz, 1500 Hz, and 2500 Hz. These correspond to *natural frequencies* of the vocal tract, or frequencies at which it is resonant. At or near these frequencies, which are called *formants*, the vocal tract will vibrate with large amplitude. Figure 13.4 shows how the spectrum of sound produced by the vocal folds can be shaped by the formants. Harmonics near the formant frequencies are supported and therefore emphasized in amplitude, and those far from the formant frequencies are not supported and therefore have smaller amplitude.

So, this is how the sounds of human speech are produced. Sound vibrations coming from the vocal folds and the larynx (which consist of a fundamental and harmonics) are shaped in amplitude by the formant frequencies of the vocal tract. Harmonic frequencies near the vocal tract resonances are strengthened in amplitude, while harmonics further from the resonances are weakened (as shown in figure 13.4). By changing the shape of the vocal tract, which involves opening and closing the jaw, moving the tongue and lips, *etc.*, the vocal tract resonances (formants) can be shifted in frequency,

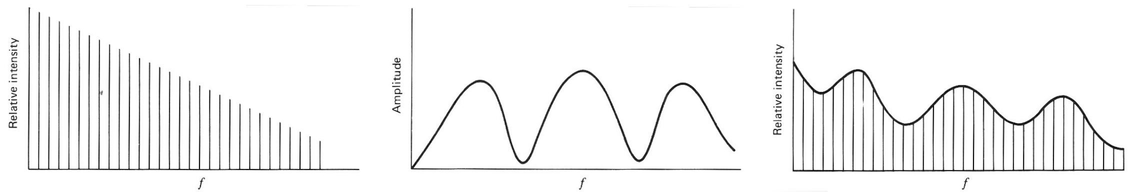


Figure 13.4: The sound produced by the vocal folds (left) contains many harmonics. The formants are resonant frequencies of the vocal tract (middle), representing regions where harmonics will be supported or not. The resultant shaping of the harmonic frequency amplitudes (right) produces the timbre of the resulting sound. Formant strengths and locations in frequency can be changed by action of the jaw, tongue, and lips.[4]

changing the regions in which harmonics are strengthened or weakened, resulting in very different sounds (timbres). Thus for a single pitch with its fundamental and family of harmonics, very different sounds (timbres) can result simply from changing the shape of the mouth and therefore the locations of the formants.

13.2 Speech: Vowels and Consonants

Human speech utilizes several different types of sounds. A distinct and individual element of speech sound is called a *phoneme*. Two very general categories of speech phonemes are consonants and vowels. Consonants are generally short, or transient in nature, while vowels are typically longer-lived, steady sounds. A more helpful categorization of human speech is one in which each of the sounds is placed into one of five basic categories: *plosives*, *fricatives*, *other consonants*, *vowels*, and *diphthongs*[5].

Plosives These are types of consonants consisting of a "burst" of air, formed by first completely blocking the vocal tract, and then suddenly opening it. Consider the following *unvoiced* plosives. When the lips do the initial blocking of air flow, the consonant "p" is produced. When the front of the tongue blocks the air, "t" is produced. When it is the soft palette, "k" results. The sound coming from a plosive is not periodic – it consists of a pulse of air that has no identifiable pitch. Also, they cannot be "held onto" the way some other consonants can. There are also plosives that involve the use of the voice. The vocal folds vibrate briefly as the pulse of air is released, for consonants including *b*, *d*, and *g*. Regardless of whether they are voiced or not, the sound coming from plosives is typically weak compared with other phonemes.

Fricatives These are different from the plosives in that they can be "held onto" for a period of time. They come in unvoiced and voiced pairs: *f* and *v*; *th* ("think") and *th* ("them"); *s* and *z*; *sh* and *zh*; and one more that has only an unvoiced version: *h*. Note that for both the voiced plosives and the voiced fricatives, there is periodic sound but no specific pitch. The sound associated with the fricatives (a "static hiss or sizzling sound") comes from the turbulent movement of air through a very small opening at high speed. This is a type of *white noise*, consisting of a broad range of frequencies with random amplitudes and no periodicity. As in the case of plosives, they can be formed at various locations in the vocal tract, including the lips and teeth (*f, v, th*), the tongue and palette (*s, z, sh, zh*), and the opening of the vocal folds (*h*). The frequency range of white noise associated with them is generally around 3000 Hz (for *sh*) and up to around 4000 to 6000 Hz (for *s*).

Other Consonants These include what are sometimes referred to as *semi-vowels*: *w*, *y*, *l*, *r*, and the "nasal" consonants *m*, *n*, and *ng*.

Vowels These consist of steady sounds with identifiable pitch. As such, each spoken vowel has some fundamental frequency and a family of harmonics. The distinct and recognizable sounds of the different vowels arise from their different *timbres*. The specific shape of the vocal tract used to speak a vowel will emphasize some of a vowel's harmonics (those at or near one of the vocal tract's resonant frequencies) and de-emphasize others (those far from the resonant frequencies). The sound spectrum that results from the specific vocal tract shape typically has three regions, or "envelopes," where harmonics have higher amplitude, separated by valleys where they have lower amplitude – see figure 13.4. These regions of emphasis are the three lowest resonances of the tract. By operation of the jaws, tongue, and lips, the shape can be changed to shift the location of these emphasis regions, thus changing the vowel.

We can visually see the difference in the vowels by considering the specific shape of their sound spectra. Figure 13.5 shows the approximate locations of the formants (y-axis) used to shape the sound spectrum of several different vowels (x-axis). The vertical range of grey boxes for each of the vowels locates the approximate position of the formants in frequency. As figure 13.4 demonstrates, the formants are fairly broad in frequency.

Diphthongs These correspond to quick transitions from one vowel sound to another, owing to the shape change in the tongue and mouth. Some vowels (for example the long "*a*" or long "*i*") are impossible to sustain for long, as in the word *late* – the *a* sound in the middle starts with the "*eh*" and ends with "*ee*." Other example words with diphthongs include boat (*oh-oo*), foil (*oh-ee*) and about (*a-ou*). Diphthongs can sometimes be challenging to sing, especially when the word has duration, since where in the musical phrasing to best place the transition may not be clear.

Example 13.1

Phonemes *What string of phonemes is used to say "Physics of Music?"*

Solution: The number of phonemes in this phrase is 13 (Ph-y-s-i-c-s o-f m-u-s-i-c), consisting of vowels, voiced and unvoiced fricatives, a nasal consonant, and plosives. The specific string of phonemes are: an unvoiced fricative (*ph*), vowel (*y*), voiced fricative (soft *s*), vowel (*i*), plosive (hard *c*), unvoiced fricative (hard *s*), vowel (*o*), voiced fricative (soft *f*), nasal consonant (*m*), vowel (*u*), voiced fricative (soft *s*), vowel (*i*), and last but not least, a plosive (hard *c*).

Example 13.2

Formants and Vowels *The location of the lowest three resonances of the vocal tract for a spoken vowel are located at 700, 1700, and 2400 Hz. Of the following words, which is likely being spoken: hood, road, glad, made, trees, and far? How about for the resonant locations at 200, 2200, and 3200 Hz?*

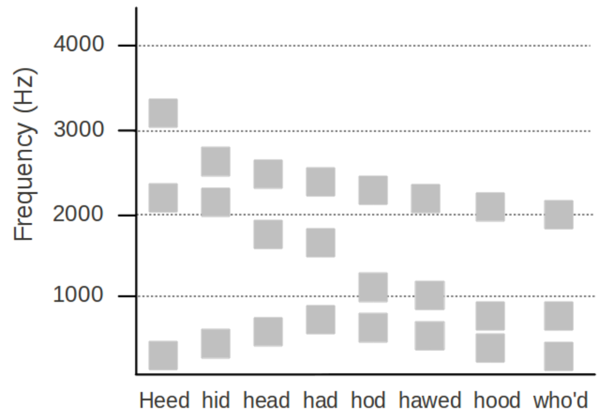


Figure 13.5: The three main formant locations in frequency (y-axis) as a function of the vowel sound (x-axis). Note that the formants are somewhat wide, indicated by the grey boxes.

Solution: According to figure 13.5, the first set of resonances best matches the frequency locations for the three boxes above the word "had," and therefore of the possibilities given, the word that comes closest to this is "glad." For the second set of resonant frequencies, the match most corresponds to the word "heed," and therefore "trees" is best represented by this set of resonances.

13.3 The Singing Voice

While a complete treatment of the singing voice is not possible in such a short chapter, a few features that distinguish it from the speaking voice are worth exploring briefly.

13.3.1 The Singer's Formant

Through proper voice training, a singer can develop an enhanced formant, most notable in the voices of operatic tenors, in the frequency range of approximately 2500 to 3000 Hz. The location of the larynx drops slightly during singing (as opposed to during normal speech), and the trained tenor singer has developed the ability to drop the larynx farther and expand the throat immediately above the larynx, creating an acoustic discontinuity in the vocal tract.

This allows for resonance vibration, *i.e.* standing waves, to develop within the larynx. The first of these standing waves has a frequency around 2500-3000 Hz. Figure 13.6 illustrates the enhancement of the vocal harmonics in the region of the Singer's Formant. The frequency location of this enhanced formant turns out to be in a region where the typical output of a full orchestra is not particularly strong. This is why the voice of a trained tenor can be heard in the concert hall over a full orchestra.

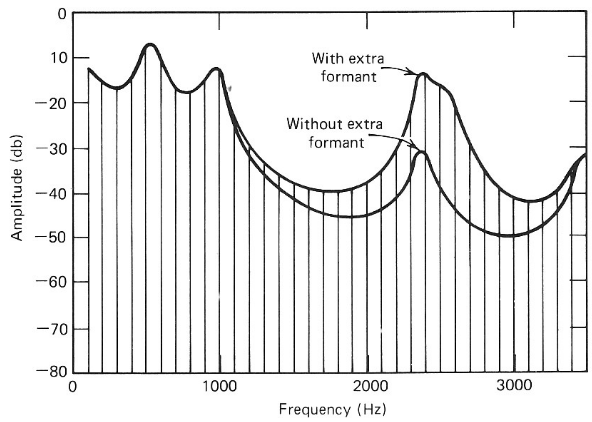


Figure 13.6: The Singer's Formant, which occurs in the frequency range 2500 - 3000 Hz. This formant serves to increase the amplitude of the harmonics in this range, enabling the voice to have increased projection in a concert hall.[4]

13.3.2 Throat Singing

A type of singing that has long been celebrated in central Asian cultures (especially in Tuva and Tibet), is called "throat singing." As in standard singing, throat singing involves vibration of the vocal folds to produce a low-pitched fundamental and a large number of harmonics. In addition to this, throat singing involves placing the jaw forward, narrowing the opening of the lips, and raising the tongue in the back, allowing for the raising and sharpening of the second formant. This results in an unusual vocal tract resonance that causes a single higher harmonic (typically one between the 6^{th} and the 12^{th}) to be heard as a separate tone from the voiced tone - see figure 13.7. Careful shifting of this formant changes the emphasis from one harmonic to another, enabling a high-pitched melody to be sung over the sustained bass note.

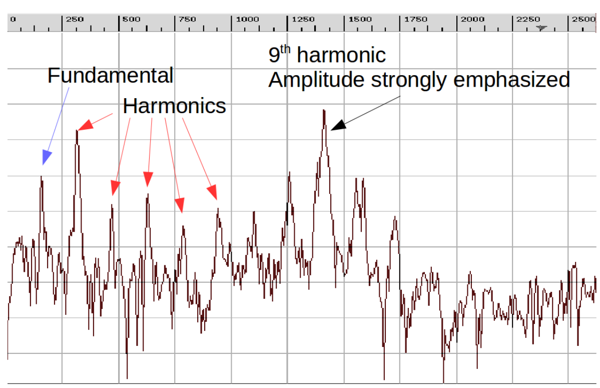


Figure 13.7: Sound spectrum from a throat singing performance. Note the 8^{th} harmonic is strongly emphasized, owing to the ability of the performer to strike an enhanced resonance in the larynx, thus giving a specific harmonic sufficient amplitude to be heard as a separate tone.[6]

On a related note . . .

Ethnomusicology

Ethnomusicology is the anthropology of music, a study of how music is created, structured, mentally represented, and used in different cultures. Music serves different purposes around the world, and ethnomusicology reveals the vast depth and beauty of it. A fascinating example is how songs in Aboriginal Australian cultures function as maps and aid in navigation. Songlines are full of detailed geography and describe the path for a traveler. They describe landmarks, water sources, and shelter, and can help someone navigate over hundreds of kilometers. While in other oral traditions stories tend to change through time, the songlines must be preserved accurately because they aid in survival. Small songs (verses of a songline) have been shown to be passed on verbatim. Ethnomusicologists have found recordings of someone singing a small song to be identical in rhythm, melody, and words to a recording of another person singing the same song over half a century before. How are songs preserved so accurately? The rhythmic text of the songs, the melodies, and accompaniment (like hands clapping or drumming on thighs) serve as mnemonics: syllables land on certain beats, and bodily movements and accompaniment serve as cues. Ethnomusicologist studies like this reveal not only the importance of music in a culture, but also how music affects cognition[7].

13.3.3 Vibrato

This is a type of frequency modulation (or adjustment) of pitch. A small muscle in the larynx, called the cricothyroid, is used to adjust slightly the pitch of the tone in a periodic fashion, typically with a modulation frequency of around 6 Hz. This is not to be confused with *tremolo,* which involves adjustment of the amplitude of the sound. Since vibrato is used more in some types of music than in others, a trained singer should be able to sing both with and without vibrato. Vibrato is also used in many instruments, and can be very effective in supplying "warmth" to the tone, especially for instruments played in ensemble. When two or more voices sing together, the addition of vibrato can help mask natural pitch variations that occur when they can't sustain exactly the same pitch.

13.4 Chapter Summary

Key Notes

- The human voice, the oldest of instruments, consists of three basic components: the lungs, the larynx (consisting of the vocal folds and the trachea, and the glottis), and the vocal tract (consisting of the larynx and the mouth).
- The lungs provide the source of air, the larynx provides the source of vibration, and the vocal tract provides the shaping of the harmonics which therefore determines the resulting timbre. For an adult, the vocal tract can be approximately modeled as a semi-closed 17 cm long cylindrical column. It therefore has resonances at approximately 500, 1500, 2500 etc. Hz, called formants.
- Human speech is divided into phonemes, of which there are several kinds, both voiced and unvoiced.

- Formants arise from resonances in the vocal cavity, and by changing the location of these formants, different timbres result, allowing the formation of different vowels.
- The act of singing (or talking) consists of activating the vocal folds (by air pressure from the lungs, permitting a sequence of air pulses into the vocal tract), and moving the jaw, the tongue, and the lips, which changes the resonant characteristics of the vocal tract, thus shifting the formant frequencies and producing the desired sound.
- Professional male singers have an extra formant, called the Singers Formant, which allows for the production of very strong sound output in the range 2000-3000 Hz, thus enabling the tenor to be heard well over the sound of the orchestra.

Exercises

Questions

1) What is a Singer's Formant, and what beneficial function does it provide?
2) Explain briefly how vowels are formed by the human vocal system.
3) Explain briefly how a singer adjusts the pitch of her voice.
4) Explain briefly how air flowing through the vocal fold ends up producing a periodic pattern of pulses, which end up producing a pitch.
5) Why are the vocal tract resonances odd harmonics of the fundamental?
6) How do fricatives differ from plosives?
7) How do voice plosives differ from unvoiced plosives?

Problems

1. What string of phonemes is used to say "Westmont College Physics"?
2. What happens to the distinction between voiced and unvoiced plosives in whispering?
3. Suppose that the representative vocal tract lengths for man, woman, and child are 17, 14, and 11 cm, respectively. For any given vowel, what approximate percentage difference would you expect to characterize all their formants when you compare them?
4. Identify the frequency, approximate musical pitch, and vowel identity of the sung note whose sound spectrum is shown in figure 13.8.
5. A 300 Hz vocal fold signal is sent through a) a cylindrical and b) a conical tract, each 17 cm long. Draw expected output sound spectra, one above the other for comparison. (Hint: Your procedure should be analogous to figure 13.4. It is important to locate the harmonics accurately on your frequency axis, even though they aren't on the book figure.)
6. From the frequency of the Singer's Formant, estimate the length of the larynx. Is your answer reasonable?
7. Suppose a vocal tract 17 cm long was filled with helium ($v = 970$ m/s). What formant frequencies would occur in a neutral tract?
8. Can you suggest why it is that an adult

may have difficulty learning to speak a new language without the accent of their home language?

9. The power P (in Watts) used to move air in or out of the lungs is equal to the pressure p (in N/m^2) multiplied by the flow rate $R_f$of air (in m^3/s). Find the power for:

   a) Quiet breathing (p = 100 N/m^2, flow rate = 100 cm^3/s)

   b) Soft singing (p = 1000 N/m^2, flow rate = 100 cm^3/s)

   c) Loud singing (p = 4000 N/m^2, flow rate = 400 cm^3/s)

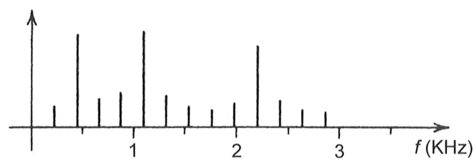


Figure 13.8:

References

[1] http://www.dukemagazine.duke.edu/issues/050608/music2.html

[2] Arthur H. Benade, Fundamentals of Musical Acoustics, Oxford Press, Inc. (1976)

[3] http://hyperphysics.phy-astr.gsu.edu/hbase/music/voice.html

[4] J. Sundberg, The Acoustics of the Singing Voice, Scientific American 236 (1977)

[5] Donald E. Hall, Musical Acoustics, 3^{rd} edition, Brooks/Cole p.296 ff. (2002) - much of the discussion of speech derives from this reference.

[6] This sound spectrum sample is from a throat singing performance of Huun Huur Tu at the Philadelphia Folk Festival, August 2006.

[7] Will, U. (2004). Oral memory in Australian Aboriginal Song Performance and the Parry-Kirk debate: A cognitive ethnomusicological perspective. In: Hickmann,E. and Eichmann,R.(eds): Proceedings of the International Study Group on Music Archaeology, vol.X. Thanks to Meg Seymour for this contribution.

Chapter 14

Musical Temperament

Do, a deer, a female deer,
Re, a drop of golden sun...[1]

14.1 Development of Musical Scales

Have you ever wondered how the traditional musical scale used in the Western world came to be? The most common musical scale used in the West is based on a 12-tone scale of equally spaced "semitones," out of which a major and a minor 8-tone "diatonic" scale is constructed. Many of us first learned about this scale from Julie Andrews in "The Sound of Music"[1] when she and the von Trapp children sing "Do Re Mi." Is this 12-tone scale, from which we get this major diatonic scale, "inevitable"? Did it develop accidentally? Is there a compelling logic supporting its structure? And if so, what is it? What is a musical scale, and where does it come from? And why do different cultures adopt different musical scales?

What are some other possibilities for defining the musical scale? Perhaps a more interesting question we can ask is how we might go about forming our own musical scale from scratch. On what basis would we decide on the frequency values corresponding to the various notes of a new musical scale? Would we appeal to observations we make in nature? to mathematics? to other forms of art?

The topic of musical temperament is complex and fascinating. It concerns musical interval frequency ratios and the individual frequency values assigned to each note of a musical scale. Many people are so familiar with listening to Western music that they may take their musical scale as a natural given. But there is nothing "sacred" about the 12-tone scale we use in Western music, although as we shall see, there is a strong logical and mathematical rationale for its structure.

In ancient days, before musical instruments became highly developed and played in ensemble, people likely sang in pitches and intervals that simply pleased them, without thought of standardizing a particular set of pitches based on some logical rationale. In similar fashion, as the first musical instruments were created, they were likely fashioned to make sounds that were simply satisfying. Over time, different musical scales in many diverse and varied cultural settings developed, both independently and influenced by others. Also over time, the structure of musical scales used within individual cultures evolved. People created ever-changing and ever-new music, instruments became more elaborate and complex, performers performed the music, and as a result musical scales evolved and became more sophisticated.

14.2 Pitch Sensation and Circularity

... and that brings us back to do, oh, oh, oh ...[1]

As we've seen, the frequency range of human hearing is quite large. From the lowest to the highest frequencies perceived by the human auditory system ($\sim$ 20 Hz – $\sim$ 20,000 Hz), there is no *a priori* reason why we might expect that any frequency should "sound" the same as any other. If we were unfamiliar with the notion of musical pitch, we might imagine that over the entire frequency range there would be a continuously varying sensation of pitch with no similarity between different frequencies. And yet our sense of pitch as a function of frequency has a very definite "periodicity," or "circularity," associated with it.

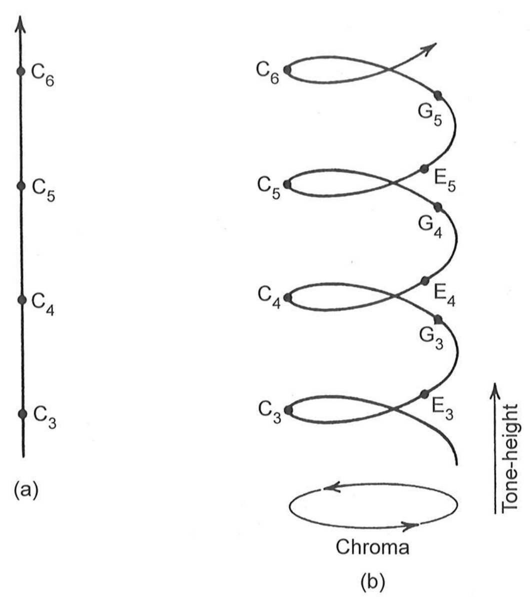


Figure 14.1: In a) a "one dimensional" way of considering the perception of pitch, where as frequency increases, pitch increases through successive octaves marked by the "C" notes of increasing tone height. In b), a two-dimensional depiction of pitch where similar points on the successive spirals correspond to similar chroma. As pitch rises, tone height increases, and the chroma values repeat periodically through successive octaves. [2]

Consider what happens when we start with two tones of identical frequency. When one of the two is slowly raised in frequency, our sensation of the two together is first to hear beats, then roughness, then two separate pitches. As they continue to move apart, we hear them continue to separate further in pitch, but only up to a point. As the second tone approaches twice the frequency of the first, its pitch *returns to its starting point*, only "higher" in value. We then hear the two tones share the same note, separated by one octave. All of the notes in between unison and the octave are said to have different *chroma*, or color, and notes with the same chroma but separated by one or more octaves are said to have different *tone height*.

As the tones are separated further in frequency past the octave, the second tone moves through the same chroma values as it did over the first octave, but at increased tone height. This pattern continues as the second tone moves through successive octaves above the first. While chroma applies to a continuous variation in frequency, pitch refers to tones with specifically defined values of frequency. In Western music the octave is divided into 12 separate pitches called semitones, each of which has a different chroma value from the others. The specific frequencies that define the pitches depend on the temperament we consider.

Given that the circularity in pitch perception is common to all humanity, the question of temperament really boils down to the question of how the gap between octaves is to be filled – how many notes will make up the octave, and to what frequencies the individual notes are to be tuned.

14.3 Intonation and Temperament

Musical intervals played by single or multiple instruments can often be produced according to the best musical sense and taste of the performers. Some instruments allow great flexibility in determining the *intonation* of musical intervals, such as the stringed instruments or the human voice. Since stringed instruments have no frets, performers are able to vary the length of the string, and therefore the frequency, over a continuous range. There is great freedom, therefore, to adjust the intonation of musical intervals to produce those sounds that are most pleasing to the ear. The same is true for vocal ensembles, especially a Capella (*i.e.* with no instrumental accompaniment).

Instruments that are "locked in" to their pitches do not have the same amount of freedom. For example, while woodwind and brass players have some ability to adjust the intonation by adjustment of their blowing, pianists have no freedom whatsoever. The piano strings are fixed in length and are activated by the strike of hammers in such a way as to allow no variation in pitch. For instruments with little or no flexibility of this sort, a system of pitches must be established in order for them to be tuned. And this, of course, raises the question of how, exactly, they should be tuned.

Even though keyboard instruments are the least flexible of the instruments, there is, of course, flexibility when the instrument is being tuned in the first place, after which the note values are set. A keyboard instrument is tuned to a particular *temperament*, which is a system that determines to which frequency each note is set. Ultimately the goal is to make the instrument and the music it produces sound "best" for its purpose. Historically several temperaments have developed over the years, each of which has both advantages and drawbacks. We will discuss three of these in some detail.

14.3.1 Western Music and the Diatonic Scale

The diatonic scale forms the basis of the Western European music. In Greek "diatonic" essentially means "progressing through tones." It consists of a 7 tone scale that repeats every octave, consisting of 5 whole tones and two semitones. Since whole tones consist of two semitones, the entire scale spans 12 semitones. Two common versions of the diatonic scale are the "major" and "minor" scales.

14.4 Musical Scale Considerations

A *temperament* is a specific recipe for how many notes make up the musical scale and how each is to be tuned. Western music has traditionally adopted the 12-tone scale spanning the octave. Since cultures through time have developed musical scales and adopted intervals based on what sounded most satisfying to them, we can wonder whether there is anything special or particular that brought about the development of the 12-tone scale in Western cultures.

Let's start by reminding ourselves of the structure of a complex tone arising from a tonal instrument. We've learned that instruments based on vibrating strings or air columns are largely tonal, meaning that the sounds they produce consist of a fundamental pitch and a family of integral multiple harmonics. Figure 14.2 shows the individual harmonics and their note values (in musical notation) for a tone tuned to "C." You'll notice that even within the first 8 harmonics, we can identify several musical intervals common to the diatonic scale. The ratios of adjacent harmonic frequencies reveal the intervals of the fifth, the fourth, and the major and minor third. We also see in the ratios of more distantly spaced harmonics the additional intervals of the major and minor sixth, and the minor seventh. We can identify notes, C, E, G (the musical triad which is the foundation of Western music) and B♭ (see figure 14.3), and would identify more notes when moving further up the ladder of harmonics.

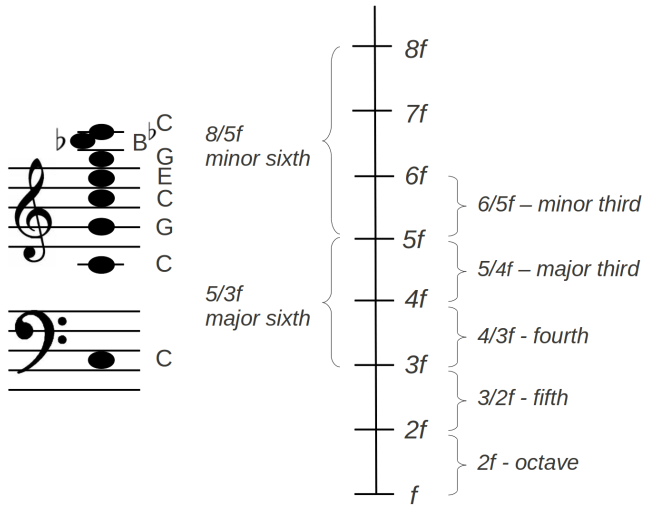


Figure 14.2: The harmonics of a "C" note, and the musical notes they correspond to. Note that the relation between neighboring and more distant harmonics correspond to several very common musical intervals used in Western music. By the 8^{th} harmonic we can already identify four notes in the diatonic scale including the very common triad. The interval of the minor seventh (between the 4^{th} and 7^{th} harmonics) is also present, though not noted in the figure.

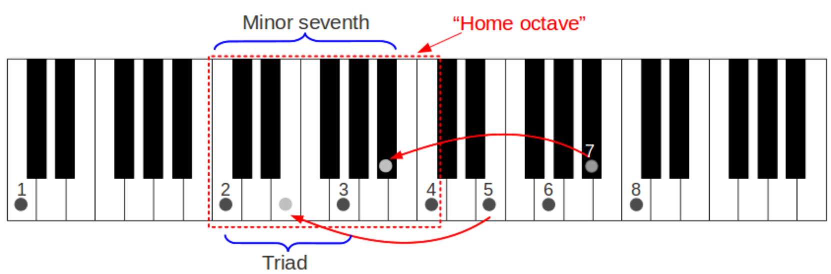


Figure 14.3: (***For color version see Appendix E***) Within the harmonic structure of the sound from a tonal instrument are already several notes that are part of the Western diatonic scale. Some higher notes (the E and the B$^\flat$) are moved down by an octave to place them in a "home octave" and identify them as members of the diatonic scale. Also present within the first 8 harmonics are the notes C, E, G, and B$^\flat$.

This figure indicates that the sound output of a tonal instrument already has within it a set of harmonics that correspond to several notes in the diatonic scale, and frequency ratios that correspond to intervals commonly used in Western music. We could imagine, then, a strong case can be made for constructing a musical scale built on the content already present in the sound spectra of tonal instruments.

14.4.1 Pythagoras and Musical Intervals

The Greek philosopher and mathematician Pythagoras (ca. 500 BC) is well known for his identification of the relationship between string length and the note produced when plucked. For two strings vibrating simultaneously, he noted that particular length ratios give rise to pleasing musical intervals – see figure 14.4. He and his followers, lovers of numbers as they were, pointed out the specific relationship that exists between small-integer fractions and pleasing musical intervals. This can be expressed mathematically:

$$\boxed{\frac{f_2}{f_1} = \frac{n}{m} \qquad n, m = \text{ integers}}. \tag{14.1}$$

Two strings vibrating simultaneously with length ratio 2:1 produce the octave, 3:2 produces the fifth, 4:3 the fourth, *etc.* They claimed that these integer ratio intervals are the ones on which music should be based. Also, they claimed that the smaller the integers that make up the fraction, the more "consonant" the musical interval produced, and therefore the more important it would be to music. This observation is consistent with the prominent role that the musical fifth and fourth (the two intervals with the smallest integer ratios beyond the octave) have played in the history and development of Western music.

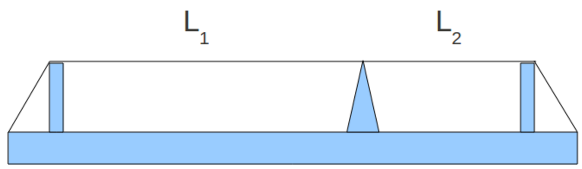


Figure 14.4: An arrangement used to divide a vibrating string into two portions. When the length ratio is the fraction of low integers, the strings form musical intervals common to the Western scale.

14.4.2 Two Pure Tones and Interval Intonation

Imagine that we combine two pure tones with frequencies f_1 and f_2 such that $f_2 = \frac{3}{2}f_1$. With the frequency ratio of 3:2 they form a musical fifth. The musical fifth is one of the most commonly used intervals in Western Music. The combination of these two tones with this exact frequency ratio sounds very pleasing to the ear. What happens if we detune this musical fifth? If we were to change the frequency of the tones so that their ratio becomes say 3:1.9, we would predict that the sound would quickly become much less pleasant. After all, why else do musicians take care to tune their instruments carefully before a performance? A nicely tuned ensemble can create very pleasant music, but a poorly tuned one can create music that is very challenging to to listen to. And yet the truth is that for *pure tones*, there is no decrease in the level of smoothness or pleasantness of their sound when we detune the interval. A trained musician would be able to tell that the interval is no longer a perfect musical fifth, but the unpleasantness that she might normally anticipate is not present. Why is this the case?

Recall that a pure tone consists only of a fundamental with no harmonics. Each is a pure sine wave. The only time that two pure tones sound unpleasant when played together is when their frequencies are so close to each other that they produce beats, or roughness of tone. Otherwise their frequencies can be set to *any* relative values with any ratio at all, and they would still sound pleasant. Does this then challenge the Greek claim that exact integer ratio intervals are preferred over others? As long as we use only pure tones, it does not appear to be the case.

However, and this is a crucial point, music is built on sounds produced by *tonal* instruments, not pure tones. Instruments produce tones that are much more rich and complex than pure tones. It is the *presence of harmonics* in the tones, as we will now see, that make the Pythagorean intervals the preferred ones to use for making music. As long as our desire is to produce sound using tonal instruments that is pleasant and free from harshness, we *can* make a case for using the exact Pythagorean intervals.

14.4.3 Two Instrument Tones and Interval Intonation

When the sounds from two tonal instruments are combined to form musical intervals, the situation is quite different from the pure tone case because of the additional presence of harmonics. Consider the musical octave, the most harmonious of musical intervals (next to that of the unison). The ratio of frequencies is 2:1 so that every harmonic of the higher tone overlaps with every other harmonic of the lower tone. Now mistune the octave slightly (see figure 14.5), and every harmonic of the upper tone then produces beats with every other harmonic of the lower tone, creating harshness in the sound.

Note also that each pair of clashing harmonics has a different beat frequency from the others, adding to the sense of harshness.

So what is the take-away from this observation? Basically, we can see that a slight mistuning of the octave in either direction produces beats in the harmonics, giving the mistuned octave a harshness of tone, perceived by the ear as dissonance. The pure octave interval represents the most pleasant sound compared with a mistuning in either direction. Figure 14.6 depicts this notion graphically. When the ratio is exactly 2:1 the most harmonious sound (i.e. zero harshness) results. When the ratio deviates from 2:1 in either direction, the sound becomes harsh owing to the clashing harmonics, as indicated by the increase in harshness along the y-axis (in arbitrary units since this is a subjective judgment).

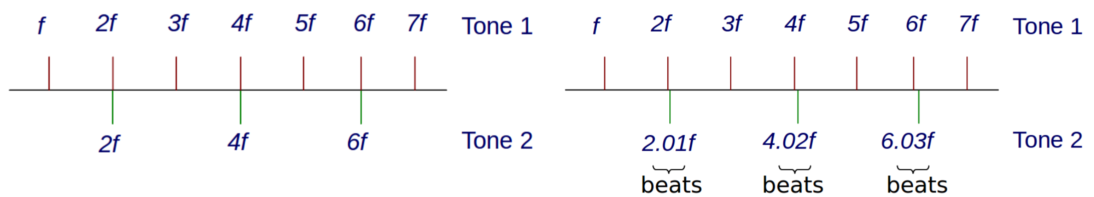


Figure 14.5: A pure (left) and a slightly mistuned (right) octave. Note that for the pure octave, all harmonics of the higher tone overlap completely with the odd harmonics of the lower tone, and there is no opportunity for any two harmonics to produce harshness through beats or roughness. For the mistuned octave, note each of the harmonics of the higher tone are located very close in frequency to the odd harmonics of the lower tone, and beats result.

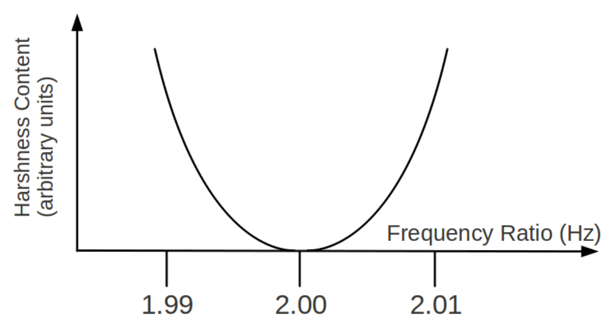


Figure 14.6: A plot depicting the sense of harshness that results from slightly mistuning the octave in either direction. The x-axis is the frequency ratio for the interval, and the y-axis represents a measure of harshness (in arbitrary units), starting from 0.

The same is true for other Pythagorean intervals when played by tonal instruments. Consider the musical fifth (see figure 14.7). The two tones share many harmonics in common: every even harmonic of tone 2 has frequency in common with one of the harmonics of tone 1. When the interval is tuned to exactly 3:2, all harmonics either overlap completely or are separated by considerable distance, and no beats are present. If one of the tones is slightly mistuned, every even harmonic of tone 2 will produce beats with harmonics of tone 1, yielding an unpleasant sound. You'll recall that for the case of *pure tones*, a slight mistuning does *not* produce harshness, since they have no harmonics that can clash and produce beats. But for *tonal* instruments the potential for harshness of tone occurs whenever two of their harmonics are spaced closely enough in frequency. A plot similar to that in figure 14.6 could be made for each of the other integer ratio intervals identified in table 14.1, demonstrating that each represents a sort of "minimum of harshness" relative to slight mistunings from those ratios in either direction. This, then, can be seen as part of the rationale for adopting the low-integer ratio intervals identified by the Greeks.

The lowest integer ratio intervals (those near the top of table 14.1) are the ones with the most overlapping and/or widely spaced harmonics, and therefore those with the least opportunity for producing

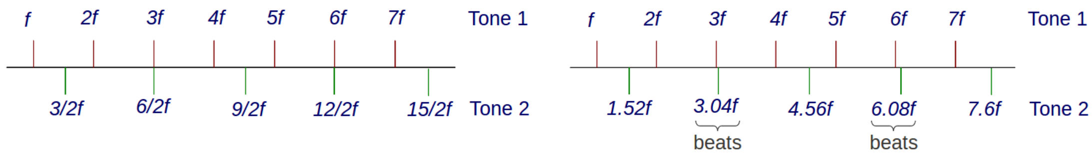


Figure 14.7: Harmonics of the two tones making up the interval of the fifth. Note that they share several harmonics. Each of the even harmonics of tone 1 overlaps with the harmonics of tone 2. Also note that nowhere do their harmonics come close enough to one another to produce beats, and therefore tension, or dissonance.

beats or roughness of tone. These should therefore be the most harmonious sounding intervals. As we move down the table to those intervals with higher integer ratios, the number of overlapping harmonics between the two decreases, and some of the harmonics of the two move closer in frequency, increasing the potential for beats. Also for these intervals the first pair of overlapping harmonics moves higher and higher in harmonic number. As an example, consider the whole tone interval – see figure 14.8. The lowest two overlapping harmonics for this interval are the 8^{th} harmonic of tone 2 and the 9^{th} harmonic of tone 1. Also, the 7^{th} harmonic of tone 2 is quite close in frequency to the 8^{th} harmonic of tone 1, allowing for the presence of beats and therefore harshness for this interval. The whole-tone interval is widely judged to be *dissonant* compared with the more consonant lower integer ratio intervals, and this analysis of the spacing of harmonics can help us understand what might contribute to this outcome, at least in part.

14.4.4 Consonance and Dissonance

The beats produced by closely spaced harmonics of two separate tones may not be immediately perceptible or obvious to the ear (since the higher harmonics are typically lower in amplitude relative to the fundamental), but their presence may nevertheless add a subliminal sense of harshness to the sound. The notion of consonance and dissonance in music is a complex and subjective topic. Consonance is associated with a sense of rest, of having "arrived" at resolution and peace in a musical

Table 14.1: Musical intervals according to the Pythagorean prescription deriving from resonances on a string, and the harmonics the two tones share in common. Note that the last interval corresponds to a *tempered* semitone.

Interval	Frequencies	First Pair of Matching Harmonics
Unison	f_1 and $f_2 = f_1$	f_2 with f_1
Octave	f_1 and $f_2 = 2f_1$	f_2 with $2f_1$
Fifth	f_1 and $f_2 = \frac{3}{2}f_1$	$2f_2$ with $3f_1$
Fourth	f_1 and $f_2 = \frac{4}{3}f_1$	$3f_2$ with $4f_1$
Major Third	f_1 and $f_2 = \frac{5}{4}f_1$	$4f_2$ with $5f_1$
Minor Third	f_1 and $f_2 = \frac{6}{5}f_1$	$5f_2$ with $6f_1$
Major Sixth	f_1 and $f_2 = \frac{5}{3}f_1$	$3f_2$ with $5f_1$
Minor Sixth	f_1 and $f_2 = \frac{8}{5}f_1$	$5f_2$ with $8f_1$
Whole Tone	f_1 and $f_2 = \frac{9}{8}f_1$	$8f_2$ with $9f_1$
Semitone	f_1 and $f_2 = 1.0595f_1$	No matching harmonics

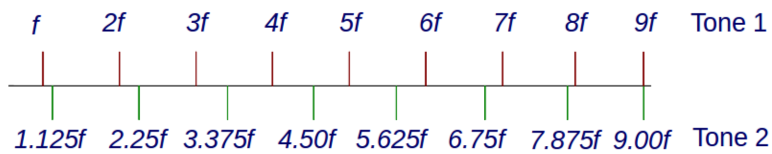


Figure 14.8: Harmonics of the two tones making up the interval of the whole tone, which has frequency ratio $\frac{f_2}{f_1} = \frac{9}{8}$. Note that the lowest harmonics that overlap between the two tones are the 8^{th} harmonic of tone 2 and the 9^{th} harmonic of tone 1. Note that for this "large integer" ratio interval, the 7^{th} harmonic of tone 2 can produce beats with the 8^{th} harmonic of tone 1, making this interval potentially dissonant.

phrase. Dissonance is associated with tension, unrest, a desire to move toward a resolution with greater consonance.

We are limited in our ability to account for musical dissonance, partly because the "definitions" of consonance and dissonance change over time as the nature of musical composition and the tastes of listeners changes and evolves. What was once considered dissonant may later be accepted as consonant. As well, a musical interval cannot be isolated from its musical context: the same interval might be alternately judged consonant or dissonant depending on the notes and intervals surrounding it in a musical phrase.

That said, we can nevertheless imagine that the presence of closely spaced harmonics between two tones could communicate to a sense of tension or dissonance to the listener. Historically, the unison and the octave (for which all harmonics of one of the tones overlaps with the other) have been judged as the most consonant intervals, followed by the fifth and the fourth (for which several harmonics overlap and none clash). The less consonant intervals are the major and minor thirds, and the major and minor sixths. The least consonant (*i.e.* dissonant) intervals are the whole tone and the semitone, which even when "perfectly" tuned still contain closely spaced pairs of harmonics that produce harshness of tone. These historic notions of consonant and dissonant intervals are consistent with the Pythagorean claim of the lowest integer ratios having the most consonance and those with higher integers ratios having the least.

In conclusion, if we desire to construct a musical scale based on intervals that yield pleasant, smooth, beat-free sound, the musical intervals of choice do appear to be those identified by Pythagoras, *i.e.* those intervals that can be expressed as integer ratios (see equation 14.1). When the two tones are tuned to the exact ratios prescribed by the Pythagorean scheme (especially to the lower integer ratios), harmonics of two tones either overlap exactly or are spaced far apart. When the intervals are slightly detuned from the exact integer frequency ratio values, the harmonics that once exactly overlapped begin to beat against one another. Even if the beats are somewhat subtle or not immediately noticeable, they do nevertheless contribute harshness to the interval.

So that brings us to the question of how to build a musical scale. How many notes should fill the octave? What recipe should we use for assigning frequencies to each note? And can we construct a musical scale where every interval corresponds to one of the integer ratios in the Pythagorean prescription? We would want all fifths within the octave to have frequency ratio exactly $\frac{3}{2}$, all fourths to have ratio $\frac{4}{3}$, all major thirds to have ratio $\frac{5}{4}$, *etc.* Is it possible to construct such a scale so that no matter which note of the octave we choose, all intervals relative to that note are the exact ratios we seek?

14.5 Common Musical Temperaments

As it turns out, no scheme of assigning frequencies to all the notes of the musical scale is free of difficulties. Let's consider three different popular temperaments used in Western music.

14.5.1 Pythagorean Temperament

...sol, a needle pulling thread... [1]

Pythagorean tuning is well suited to music that emphasizes musical fifths and octaves, and does not modulate significantly from the home key. This generally applies to music prior to the 15th century. One of the drawbacks of this temperament is that while most intervals of the fifth in the octave are perfect and therefore very smooth, other intervals are not perfect, and in particular the major thirds are quite sharp. Another drawback is that there is very little freedom to modulate from the home key to more distant keys, as the intervals become quite harsh and dissonant. As musical thirds became more important to musical expression, and as the desire to modulate from the home key grew, other alternative temperaments were developed.

The basic building block of the Pythagorean temperament is the interval of the perfect fifth. We can produce all of the notes of the diatonic scale (the solfège) by starting at some arbitrary note and moving away from that note in perfect fifths, accumulating the notes of the diatonic scale as we move. Let's see how this works. Beginning at C_3, we can first move up by one fifth to reach G_3. The frequency ratio of these two tones is

$$\frac{f_{G_3}}{f_{C_3}} = \frac{3}{2}. \tag{14.2}$$

Moving up a fifth from G_3 we reach D_4, whose frequency relative to G_3 is $\frac{f_{D_4}}{f_{G_3}} = \frac{3}{2}$ and whose frequency relative to our starting point of C_3 is therefore

$$\frac{f(D_4)}{f(C_3)} = \left(\frac{3}{2}\right) G_3 = \left(\frac{3}{2}\right)\left(\frac{3}{2}\right) C_3 = \left(\frac{3}{2}\right)^2 C_3. \tag{14.3}$$

As we continue this movement upward to n fifths above C_3 the ratio of this note's frequency to f_{C_3} is $\left(\frac{3}{2}\right)^n$. Figure 14.9 depicts the successive fifths as we move up the keyboard from C_3.

The next step in the Pythagorean construction of the diatonic scale is to move the accumulated notes downward so that they all fit into one octave. This means moving some notes down one octave and some down two. Whenever a note is moved down an octave, its frequency is reduced by a factor of 2. The resulting values of the frequencies relative to f_{C_3} are shown in figure 14.10. After moving up 5 fifths above C_3 and then moving them down into a common octave, we've accumulated all but one of the tones of the major diatonic scale.

The final step in producing the entire diatonic scale is to move *down* from the initial C_3 note by a fifth (which yields an F_2), and then move the resulting note up one octave - see figure 14.11.

At this stage, we can rearrange the 8 notes we've accumulated so far to create the diatonic octave. Figure 14.12 shows each of their frequency ratios relative to our starting note, C_3.

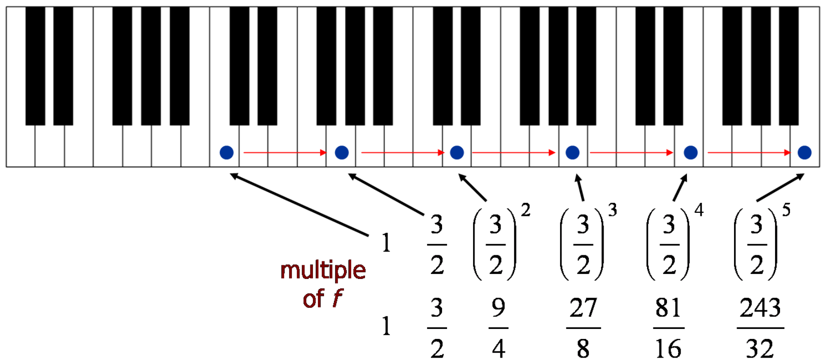


Figure 14.9: Moving up in fifths starting from the note C_3, whose frequency is defined as f. Note that as we move up the keyboard, the sequence of fifths above C_3 have frequencies that are powers of $\frac{3}{2}$ relative to f.

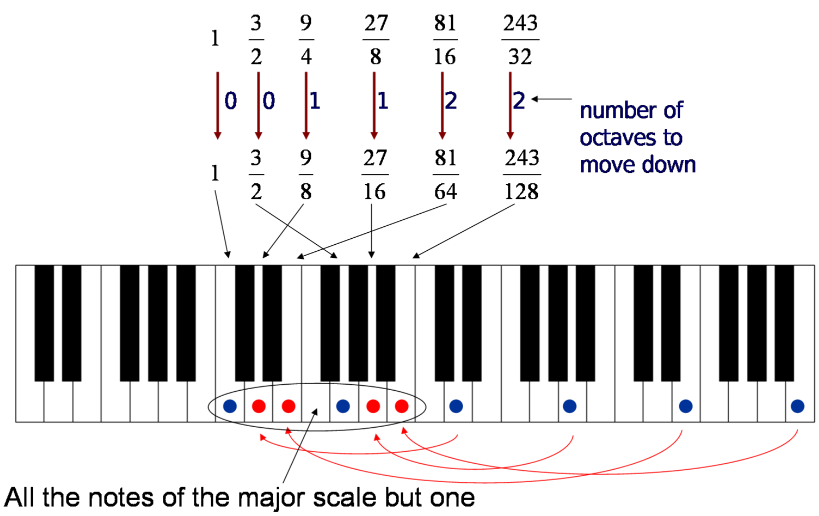


Figure 14.10: (***For color version see Appendix E***) Moving the acquired notes into a common octave, we have 6 out of the 7 required notes: C, D, E, G, A, and B.

Before we move any further, I'd like to pause and anticipate how many notes will eventually fill out our scale. The semitone is a natural product of our Pythagorean construction. It corresponds to the interval between mi and fa, and to the interval between ti and do'. The semitone represents the smallest interval, and a whole tone corresponds to two semitones. Over the span of our 7 notes there are two whole tone intervals (do-re and re-mi) followed, by a semitone (mi-fa), followed by 3 whole tones (fa-sol, sol-la, and la-ti), followed by a semitone (ti-do'). Therefore if we fill the octave with the rest of the notes that sit in between each whole tone interval, we would end up with 12 tones total spanning the octave. This is the rationale behind our 12-tone scale.

We can find the ratios that correspond to the whole and semitone intervals in the Pythagorean construction by taking the ratios of frequencies for adjacent notes. The result is that the whole tone and semitone intervals are found to be $\frac{9}{8}$ and $\frac{256}{243}$, respectively – see figure 14.13.

The question now arises – how adequate is this scale? Will it serve our musical needs? It is still missing several semitones, but let's consider what we have so far. Note first that the fifths and fourths of the scale are fine - they have frequency ratios of $\frac{f_{sol}}{f_{do}} = \frac{3}{2}$ and $\frac{f_{fa}}{f_{do}} = \frac{4}{3}$, respectively. The major third, however, is sharp relative to its preferred value. The ideal ratio of the major third, derived from

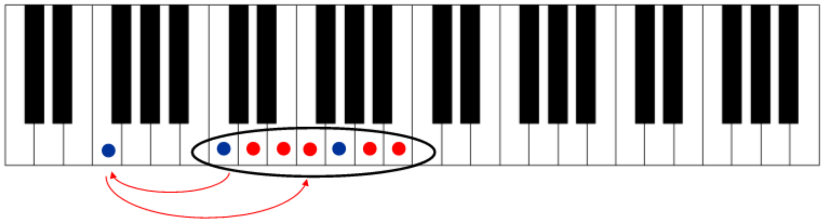


Figure 14.11: (***For color version see Appendix E***) Moving down one fifth and then up one octave identifies the last note of the major diatonic scale in the Pythagorean temperament.

multiple of f	1	$\frac{9}{8}$	$\frac{81}{64}$	$\frac{4}{3}$	$\frac{3}{2}$	$\frac{27}{16}$	$\frac{243}{128}$	2
	do	re	mi	fa	sol	la	ti	do′

Figure 14.12: Re-arranging the notes in order to create the octave. Frequency values are relative to the starting note.

the vibrational resonances of a string (see table 14.1) should be $\frac{5}{4} = 1.25$, whereas according to the Pythagorean temperament it comes out to $\frac{f_{mi}}{f_{do}} = \frac{81}{64} = 1.265$, which is sharp. Additionally, the minor third $\frac{f_{fa}}{f_{re}} = \frac{4}{3}\frac{9}{8} = \frac{32}{27} = 1.185$ which is flat relative to the desired value of 1.200. It is clear, then, that this scale system has drawbacks.

How do we then fill out the rest of the octave with the remaining semitones? The procedure is identical to what we've done so far. We continue ascending and descending by musical fifths starting from C_3 and then move the resulting notes down or up the appropriate number of octaves to locate them in the home octave. Without tracing the detail of this procedure, I've summarized the results in figure 14.14.

There are a few shortcomings in this musical scale worth mentioning. There are two different semitone intervals. Notice that the do-do$^\sharp$, re-re$^\sharp$, mi-fa, sol-sol$^\sharp$, la-la$^\sharp$, and ti-do' intervals are equal to $\frac{256}{243} = 1.053$, whereas the do$^\sharp$-re, re$^\sharp$-mi, sol$^\sharp$-la, and la$^\sharp$-ti intervals are $\frac{2187}{2048} = 1.068$. Note also that there is an inherent ambiguity in the assignment of note fa$^\sharp$ (or sol$^\flat$) – there are *two* values for this note depending on whether we arrived at it by ascending or descending by intervals of a fifth from our starting point. These inconsistencies represent significant problems for this temperament. We'll come back to it shortly. Let's first consider a second temperament.

14.5.2 Just Temperament

This scale attempts to maximize the number of consonant intervals having exact frequency ratios within the octave. Remember that the major third for the Pythagorean temperament ended up sharp. The just scale starts from a perfect triad, one that has a perfect fifth *and* a perfect third. The triad is central to Western music, consisting of simple trio of notes do-mi-sol, or C-E-G with frequencies in the ratio $1 : \frac{5}{4} : \frac{3}{2} = \frac{4}{4} : \frac{5}{4} : \frac{6}{4} = 4 : 5 : 6$. In a similar fashion to the construction of the Pythagorean scale, we'll start with a triad at C_3 and move up and down by successive triads. The starting point, along with the frequency ratios for the notes relative to the lowest note in the triad, is

do	mi	sol
$\frac{4}{4}$	$\frac{5}{4}$	$\frac{6}{4}$

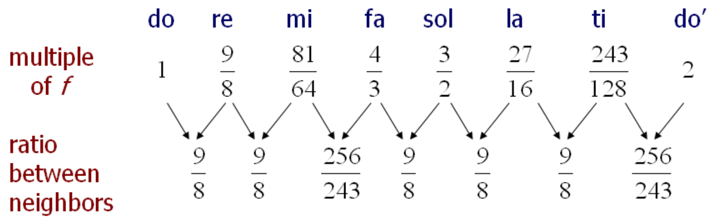


Figure 14.13: Frequency ratios corresponding to the whole tone and semitone in the Pythagorean temperament.

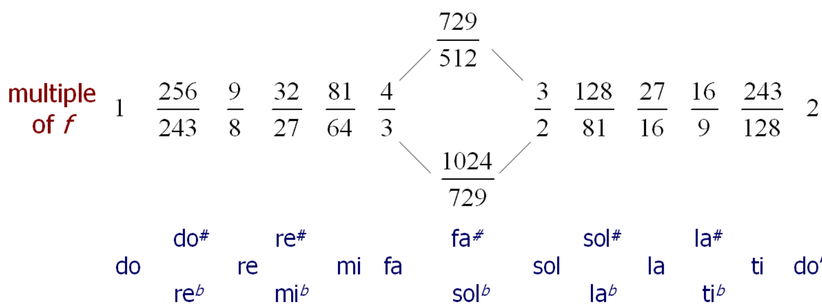


Figure 14.14: The complete chromatic scale in the Pythagorean temperament.

from which we will move up one triad and down one triad:

		do $\frac{4}{4}$	mi $\frac{5}{4}$	sol $\frac{6}{4}$		
fa' $\frac{4}{6}f$	la' $\frac{5}{6}f$	do $\frac{6}{6}$		sol $\frac{6}{4}f$	ti $\frac{15}{8}f$	re' $\frac{9}{4}f$

Then we bring notes to within the home octave by descending or ascending one octave, and then arrange the notes in order to end up with the notes depicted in figure 14.15, with the intervals between neighboring notes indicated as well.

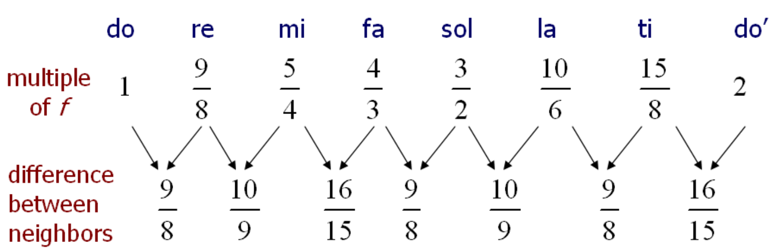


Figure 14.15: The diatonic scale in just temperament, and the interval ratios between them.

Note a major shortcoming with this scale is that there are two different whole tone intervals. The whole tone interval corresponding to do-re, fa-sol, and la-ti is $\frac{9}{8} = 1.125$, whereas the whole tone interval corresponding to re-mi and sol-la is $\frac{10}{9} = 1.111$. The semitone interval in just temperament has ratio $\frac{16}{15} = 1.0667$. Some other interval ratios are problematic as well. The minor third fa-re is $\left(\frac{4}{3}\right)\left(\frac{8}{9}\right) = \frac{32}{27} = 1.185$, which is slightly flat compared to its ideal value of $\frac{6}{5} = 1.200$.

The chromatic scale in just tuning is acquired in the same fashion as it was for the Pythagorean tuning, *i.e.* by continuing to move up and down from our starting point in triads, pulling the notes back into the home octave, and rearranging the notes in order. The end result is that there are more note assignment ambiguities in the just scale than there were in the Pythagorean.

The just temperament has, overall, a larger number of pleasing intervals than the Pythagorean temperament, but it shares the lack of freedom to modulate from the home key in which the instrument is tuned, to different keys. Our scale construction above is referenced to the key of C, for which we assign the solfège notes do, re, mi, ... *etc.* Imagine we wanted to tune the instrument to the key of D instead of C, starting the new solfège sequence from D, as shown in table 14.2. Compare the values for each of the keys that both scales have in common – compare the D key in the D- and C-tuned scales, the E key in both, *etc.* Note that the E and A keys end up being *tuned to different frequencies* between the two tunings. This means that if the keyboard is tuned to the key of C and the music is played in the key of D, sour intervals will result since the instrument is not tuned in D. Once tuned to a particular key, the instrument is basically restricted to play music in that key, or in very closely related, if the objective is to have sound that is pleasing.

Both the Pythagorean and just tuning scales share these harmonically difficult features. While each contains intervals that are "perfect" in their home key (the just scale contains more of these than the Pythagorean), they also contain intervals that are off from their ideal values (for example the sharp major third in the Pythagorean tuning). In both there is very little freedom to modulate away from the home key into other keys without inheriting sour intonation.

Table 14.2: Notes in the just temperament tuned in the keys of C and of D. Frequency ratio values are relative to C.

	do	re	mi	fa	sol	la	ti	do
	C	D	E	F	G	A	B	C
C-major	1	$\frac{9}{8}$	$\frac{5}{4}$	$\frac{4}{3}$	$\frac{3}{2}$	$\frac{5}{3}$	$\frac{15}{8}$	2
D-major	$\frac{9}{8}$	$\frac{81}{64}$	$\frac{45}{32}$	$\frac{3}{2}$	$\frac{27}{16}$	$\frac{15}{8}$	$\frac{135}{64}$	$\frac{9}{4}$
	D	E	F$^\sharp$	G	A	B	C$^\sharp$	D

Unfortunate as it may be, it is impossible to build a scale to fit within the octave that contains exact fifths and/or thirds without causing other intervals to have poor intonation. This may seem counter-intuitive, but considering the following exercise. If we start from the note corresponding to the low A_0 on a keyboard, and move up exactly 12 perfect fifths (each having frequency ratio 3:2), we will end up on the note A_7. We could also make that same journey by moving from A_0 to A_7 through 7 perfect *octaves* (each having frequency ratio 2:1). When we compute the final frequency for A_7 through these two separate procedures (*i.e.* one through 12 perfect fifths and the other through 7 perfect octaves), believe it or not, we would end up with *different* final frequency values for A_7! The inability to end up at the same frequency for our final note underscores the inability to construct a perfect temperament that contains both perfect fifths *and* perfect octaves (which are arguably the two most commonly used musical intervals) without inheriting severe inconsistencies in other intervals. These difficulties were recognized by the 17^{th} century, and shortly thereafter a new temperament was developed, to which we move next.

14.5.3 Equal Temperament

The question seems very basic ... is there an intonation system that will have both of these highly desirable features:

1. True intonation – can *all musical intervals* within the octave be set to their perfect values, according to the natural resonances of a string (those listed in table 14.1)?
2. Freedom to modulate – can we modulate from the home key to any other key (there are 24 major and minor keys within the 12-tone system!) without encountering sour intonation?

We've seen that it is impossible to create a tuning system that has perfect intonation for all intervals within the octave. As soon as we begin to set some of the intervals to perfect intonation, other intervals automatically become imperfect – we can't avoid that. Furthermore, if we succeed in producing an intonation in which at least *some* of the intervals are perfect, and all other intervals, while not perfect, are sufficiently pleasant, we still lack the ability to modulate away from the home key without encountering considerably worse intonation. Apparently we will need to strike some compromise. So let's modify slightly the question we posed above. Can we create a system that has the following two desirable features:

1. "Near" true intonation – can musical intervals within the octave be set close enough to their perfect values so that, while not perfect, yet all of them are still relatively pleasant?
2. Freedom to modulate – can we modulate from the home key to other keys and still have the same level of pleasantness?

There is an intonation system that has these desired features, called equal temperament.

This intonation system is based on setting all semitones to *exactly* the same frequency ratio, and all octaves to be perfect, *i.e.* with ratio 2:1. There will be no distinction from one semitone to another – all will be exactly the same as all others. This will grant us the freedom to modulate from one key to another since there is no distinction between keys – they all contain the same semitone intervals, and therefore they will share the same values for all the other (slightly compromised) intervals as well. So the question then becomes what is this ratio to which we need to set our semitones? We want to construct a 12-tone scale, so that our octave is spanned by exactly 12 semitones. Let's call the frequency ratio for the semitone "a" and begin our construction at "do." The first semitone above "do" (which will be "do$^\sharp$") will have frequency "a" times "do." All other notes in the scale will have the same frequency ratio relative to the semitone above or below them. In other words,

$$\frac{f_{do^\sharp}}{f_{do}} = \frac{f_{re}}{f_{do^\sharp}} = \frac{f_{re^\sharp}}{f_{re}} = ... = \frac{f_{do'}}{f_{ti}} = a. \tag{14.4}$$

These 12 semitones will span the entire octave, and our requirement when we get to the top is that

$$f_{do'} = 2f_{do}. \tag{14.5}$$

If we assign "do" to have frequency f, then do$^\sharp$ will have frequency $f_{do^\sharp} = af$. After that, "re" will have frequency $af_{do^\sharp} = a(af) = a^2 f$, then the next semitone "re$^\sharp$" will have frequency $af_{re} = a(af_{do^\sharp}) = a(a^2 f_{do}) = a^3 f_{do}$ and so on until we reach the top, at which point $f_{do'} = a^{12} f$. Since this ratio of the octave must necessarily be equal to 2, then we can compute the value of a:

$$a^{12} = 2 \quad \rightarrow \quad \boxed{a = \sqrt[12]{2} = 1.059546}. \tag{14.6}$$

The notes belonging to the equal temperament scale are listed in appendix C. A peculiar feature of this temperament is that the *only* interval that is set to its perfect value is the octave. Every other interval is set at a frequency ratio that is close to, but not exactly equal to, its ideal value. Figure 14.16 shows a comparison between the ideal values of the interval frequency ratios (top) and the equal temperament ratios (bottom).

Table 14.3: Frequency ratios for notes of the solfège in equal temperament

Note	Frequency
do	1.0000
do$^\sharp$	1.0595
re	1.1225
re$^\sharp$	1.1892
mi	1.2599
fa	1.3348
fa$^\sharp$	1.4142
sol	1.4983
sol$^\sharp$	1.5874
la	1.6818
la$^\sharp$	1.7818
ti	1.8877
do'	2.0000

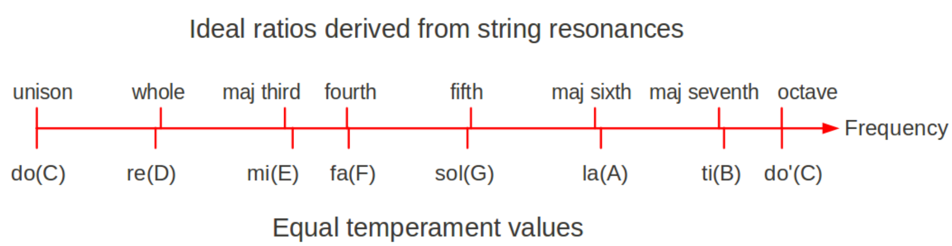


Figure 14.16: A comparison of the frequency ratios for common intervals in the equal temperament system (bottom) compared with their ideal values derived from the resonances of a vibrating string (top).

A quick glance at some of the more important intervals shows how much of a compromise we've managed to inherit. The interval of a fifth, which should ideally be $3:2 = 1.5$ is now equal to 1.4987, slightly flat by 0.087%. This means that, according to our earlier discussion of tonal instruments, we should hear some slight beating in the harmonics from this slight mistuning. A trained ear can hear them, but fortunately the beats are sufficiently weak that most people will hear this interval with little or no harshness. Other consonant interval ratios are listed in table 14.3. The largest mismatch is in the thirds and sixths. Therefore this system too has its drawbacks for some intervals. But the important point to note is that even though *all* intervals are off from their ideal values (except for the octave), they are close enough to their ideals that the ear still hears them as relatively pleasant (especially compared with the difficult intonations we encountered for some intervals in the Pythagorean and just scale systems).

The major benefit in adopting this temperament is the freedom for keyboard instruments to be played together with all other instruments in any key, and for the freedom of modulation within a piece of music to any other key without a change in the interval ratios that will make one key sound more or less pleasant than any other. All keys will have the same exact interval ratios for fifths, fourths, third, *etc.* All will be off in exactly the same way, but by little enough that the ear is happy with the sound.

The piano has become a very popular instrument today, and for this reason people have largely become accustomed to listening to equal temperament intervals without noticing its imperfect features. Compare this with someone who was hearing it for the first time after having gotten used to Pythagorean or just temperaments, for whom the new sound was likely a little more displeasing than it is for us today. However, the tempered scale represents the best compromise available for keyboard instruments with fixed intonation, and is therefore widely used in keyboards around the world. Johann Sebastian Bach was so thrilled at the advent of this new temperament system (since freedom of modulation is such a large part of his musical expression) that he celebrated it by writing two separate sets of 24 preludes and fugues, each written in one of the 24 major and minor keys available in the octave. The idea was in part that you could sit down and play through all of the pieces in one sitting without needing to retune the instrument.

Before the equal temperament system was adopted, a public performance would often need to be interrupted between pieces written in different keys in order to retune the instrument to render the sound appropriately pleasing. With the equal temperament scale, many pieces in different keys could be played in one sitting, which suited Bach just fine!

14.6 Chapter Summary

Key Notes

- Musical scales from around the world differ from one another, and one can ask whether one system is more “inevitable” than another.
- The human auditory system perceives a sense of circularity to pitch. As we increase tone height, the chroma value of tones changes in a repeating pattern (which we associate with octaves).
- Intonation refers to the tuning of musical intervals on an instrument, and temperament corresponds to a particular system of notes assigned to the scale. Instruments that are “locked in” to tuning (*e.g.* the piano) require a temperament system on which to base its tuning. Other instruments that have flexible intonation (*e.g.* the violin) can plan intervals of different intonation and are not locked into any particular temperament.
- Western music is based on the diatonic scale consisting of 5 whole tones and 2 semitones. Two common versions of this scale are the “major” and “minor” scales.
- Pythagoras and his followers determined what they considered to be harmonic intervals based on integer ratio string lengths playing together. Ratios involving small integers were more harmonic, and ratios involving larger integers less so.
- Intonation is important for the formation of pleasant sounding intervals, especially for sound from tonal instruments.
- Consonance and dissonance are subjective categories which change through the ages, but it is thought that consonance generally brings a sense of peace and resolution, while dissonance brings a sense of tension and lack of resolution. Intervals that are judged to be dissonant are thought to involve the presence of beats between various harmonics of the two tones.

- Common musical temperaments in the Western world are the Pythagorean, based on perfect fifths, the just scale, based on perfect triads, and the tempered scale, based on the semitone having frequency ratio $a = \sqrt[12]{2} = 1.059546$.
- Each temperament system has benefits and drawbacks. The tempered scale is the most commonly used system today, since it affords the musician the ability to play in any musical key without having to retune the instrument, and great flexibility in modulating from key to key without inheriting sour interval intonation.

Exercises

Questions

1) Briefly describe what the advantages and disadvantages the tempered scale has over the Pythagorean and just scales.
2) Why is the Tempered scale the most useful scale (compared with the Pythagorean and just) to use in contemporary music? What are their benefits and drawbacks?
3) Describe briefly the idea behind the construction of the tempered scale.
4) Describe briefly the ideas behind the construction of the Pythagorean scale and the just scale.
5) Describe "chroma" and "tone height." How are they different?
6) What is the pitch relationship between two tones with the same chroma and different tone height?
7) What is meant by the circularity of pitch sensation?
8) Regarding the integer ratios corresponding to musical intervals, did the Greeks claim that there exists a general correlation between the actual values of the integers and the sense of interval consonance? If so, what was their claim?
9) Two instruments are playing a perfect interval. When one is then slightly mistuned, why does the interval become dissonant? Is this also true of pure tones? Why or why not?

Problems

1. At what point would you divide a 65 cm long guitar string so that the two segments sound pitches one octave apart?
2. Start at the lowest note on the piano, A_0, and move upward from that note by seven successive octaves.
   a) What note do you end up on?
   b) How many musical fifths are between these starting and ending points?
3. In the previous question, you established the number of octaves and fifths between the notes A_0 and A_7. If we assume that the frequency of A_0 is f,
   a) What is the frequency of A_7 regarded as *exactly* 7 octaves above A_0?
   b) What is the frequency of A_7 regarded as *exactly* 12 fifths above A_0?
4. Examine the Pythagorean diatonic scale shown in the chapter. How many perfect fourths are possible from the eight notes of this scale?
5. How many perfect major thirds are possi-

ble from the eight notes of the just diatonic scale?

6. Suppose we want to produce a whole tone interval in two different ways: a) first by moving up a perfect fifth and then moving down a perfect fourth; and then b) by moving up a perfect fourth and then moving down a perfect minor third. What ratio, relative to the starting note, is the frequency of the whole tone when created using these two different methods? (Note: you will get two different results, neither of which has the unique claim of being the right size!)

7. Imagine that we tune the notes F_3 and A_3 to frequencies 176 and 220 Hz, respectively.

    a) Show that this would produce a justly-tuned third, *i.e.* a perfect third.

    b) Draw the harmonics of F_3 and A_3 and indicate which harmonics are common to both.

    c) If we call frequencies of the two tones f_1 and f_2, and the notes are slightly mistuned, show that the beat rate would be equal to $f_b = 4f_2 - 5f_1$.

    d) If the mistuned frequencies are 174 and 220 Hz, what will be the perceived beat frequency? What is the approximate pitch of this beat tone?

8. Imagine that we would like to generate a musical scale that spans the entire octave using the following scheme: Start with a frequency f, move upward a perfect fifth, then downward a perfect fourth. This process will generate a whole tone. Then staring from that new note, move up a fifth and down a fourth to generate the next whole tone. Continue in this fashion until we reach the octave. Identify difficulties inherent in the musical scale resulting from this procedure.

9. In problem 3, you should have found that the frequency of A_7 is $128f$ when it is computed by moving up 7 octaves from A_0 with frequency f. On the other hand, when you compute the frequency of A_7 by moving up 12 fifths from A_0, you should have found that it ends up at frequency $129.75f$. Clearly this is a problem since one note can't have two different frequencies. In terms of the temperament of the piano, what compromise would need to be made? Is this consistent with what is done in equal temperament tuning? Comment briefly.

10. A new equally tempered scale is invented with 16 steps spanning the octave from f to $2f$. On this new scale, what is the frequency ratio of the 4^{th} note of the scale relative to the 1^{st} note f? How about for the 12^{th} note?

11. The sounds in a touch tone phone correlate to the following frequencies: 697, 770, 850, 941, 1209, 1337, and 1477 Hz. What are the closest notes to these on the musical scale?

12. Verify by multiplication that a fifth plus a fourth equals an octave in any tuning, as does a major sixth plus a minor third.

13. Using the frequency ratios in figure 14.15, verify that the intervals C:G, E:B, F:C', G:D', and A:E' are perfect fifths in the just diatonic scale. Determine the frequency ratio for the imperfect fifth, D:A.

14. Using table 14.1:

    a) What is the frequency ratio for a perfect major sixth?

    b) Do C_4 and A_4 in equal temperament have this ratio?

    c) When these two notes are played together, what beat frequency will be heard, and what is the approximate pitch of this beat tone?

References

[1] The Sound of Music, Rodgers and Hammerstein II (1959), film directed and produced by Robert Wise (1965)

[2] Donald E. Hall, Musical Acoustics, Third Ed., BrookesCole pg. 408 (2002)

REFERENCES

Chapter 15

Musical Acoustics

The term acoustics applies to the broad physical science of sound, pertaining to the production, propagation, and perception of sound waves. Its meaning in Greek is "of (or for) hearing," and therefore applies to all of sound. A course in the physics of music could easily be titled the "Acoustics of Music." On the other hand, the term is often more narrowly used to refer to the study of how sound interacts with the interior of an enclosed volume such as a room, studio, or concert hall. This is the sense in which we will consider the term, and this chapter will explore the ways in which the environment in which music is played shapes and affects the sound we ultimately hear.

15.1 The importance of good acoustics

Let's start by considering the *absence* of a surrounding acoustic environment. Does such a possibility exist? Imagine, for example, sitting at the top of a flag pole and singing. The sound moves out in all directions, not reflecting off of anything to return to the ears. In this special case, the sound intensity falls off as $1/r^2$, just as we've seen in earlier chapters, and we are reminded that this intensity principle only applies if there is nothing nearby to reflect and therefore redirect the sound, *i.e.* when there is *no* surrounding acoustic environment present.

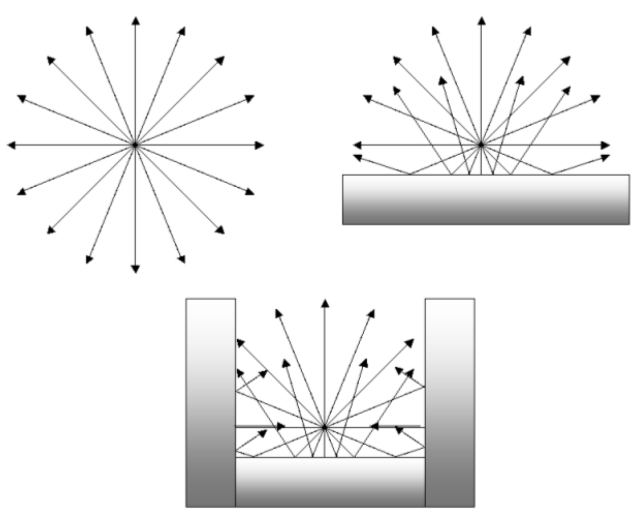


Figure 15.1: Sound intensity falls off as $1/r^2$ from the top of the flagpole, but near the ground reflection of sound back upwards causes the intensity of the perceived sound to fall off more slowly as a function of distance from the source. In a room where there are even more reflecting surfaces the sounds falls off even less as a function of distance. The latter two cases are examples for which sound does *not* fall off as $1/\mathrm{r}^2$.

As you descend the flagpole toward the concrete ground (continuing to sing, of course), your voice sounds louder to you as you descend because of the sound reflected from the ground underneath you. You are now entering an acoustic environment where sound is affected and shaped by local surfaces. Under more "normal" circumstances (other than sitting on a flagpole, that is), there are almost always surfaces in the local environment that reflect and absorb some of the sound emitted by a source. When you listen to my lecture, you are not only hearing the direct sound from my voice, but the reflected sound from the ceiling panels, walls, desktop surface, *etc.* Any listening experience involves receiving *direct* sound from the source *and* indirect sound reflected from local surfaces.

The acoustical environment in which music is played is an important component to the overall listening experience. Not only is it important during a musical performance for the audience to receive a well-rounded listening experience, and for that sound quality to be distributed uniformly throughout the concert hall, but it is also essential that the musicians on-stage hear one another well to coordinate their artistry as an *ensemble*, and therefore for sound to be able to travel from one side of the stage to the other. Concert halls must be designed to minimize audience noise while at the same time supporting the sound projected by the orchestra. Echoes must be avoided, and sounds reflecting from various surfaces in the auditorium need to reach the listener soon after the direct sound arrives. We will explore many aspects of what makes a good acoustic environment.

15.2 Reflection, Dispersion, and Absorption

A good room acoustic will translate into high quality of sound reaching the listener. Again, this arriving sound consists of both direct sound (from the sound source) and indirect sound (through reflections off surfaces). We need therefore to understand the nature of sound reflection and absorption if we are to assess the quality of an acoustic environment.

The study of acoustics really boils down to the the consideration of two basic interactions between sound and surfaces. Whenever a sound wave impinges upon a surface, such as a wall, a couch, a curtain, *etc.*, it can *reflect* (either "regularly" or by being "dispersed") and it can become *absorbed*. There is always *some* reflection and *some* absorption for sound interacting with all surfaces – some are very good reflectors, others are very good absorbers.

Furthermore, the nature of reflection depends on the relative size of the sound wavelength and the object from which it is reflecting. If the size of the object is larger than about 4λ, *regular* reflection occurs (see below). If the size of the object is less than this, *dispersion* of the sound can occur, which tends to spread out and distribute it more uniformly throughout the room. When the size of the object is far smaller than the sound wavelength, it doesn't really affect the sound at all.

15.2.1 Reflection

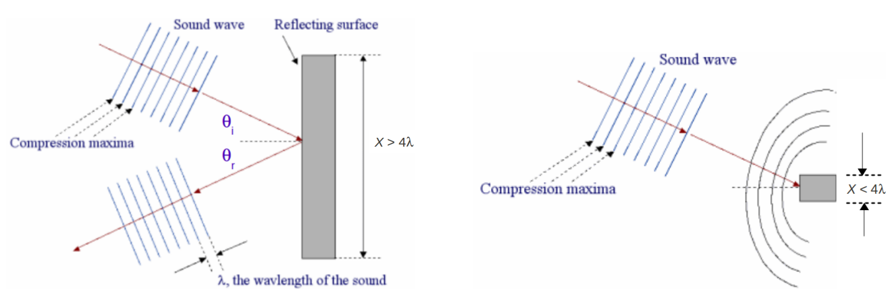


Figure 15.2: Reflection of a sound wave from a large surface. The surface dimensions are much larger than the wavelength, so that "regular" reflection occurs, meaning that the angle of incidence is equal to the angle of reflection.

"Regular" reflection occurs when the size of the reflecting object has dimension $d > 4\lambda$. Sound reflects in such a way that its incoming angle θ_i (measured with respect to a line perpendicular to the surface) equals the angle of reflection θ_r at which it leaves the surface - see figure 15.2. Musical sound ranges in wavelength from about 2 cm (15,000 Hz) to about 17 m (20 Hz). When the surface is flat and straight, incoming parallel rays of sound are reflected such that the outgoing rays are also parallel to one another. When the surface is convex, incoming parallel rays spread out from one another into the room, whereas for concave surfaces, incoming parallel rays converge toward one another. In the former case, sound is nicely distributed throughout the room, which is acoustically desirable, whereas in the latter case, sound tends to be "focused" into parts of the room, which is acoustically undesirable. This latter case tends to create "hotspots" in the room, which detracts from the quality of the acoustic environment. Therefore convex surfaces serve very well in concert halls, whereas concave surfaces should be avoided.

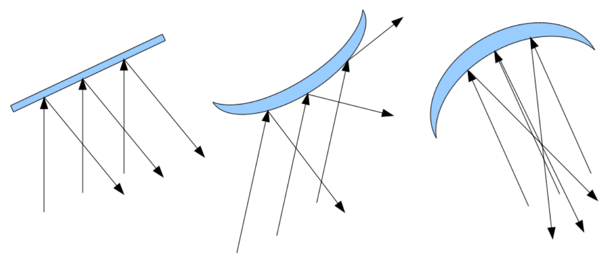


Figure 15.3: Regular reflection from straight and curved surfaces. Notice that convex surfaces disperse the sound throughout the room, whereas concave surfaces tend to focus sound. The former is desired in order to fill the room uniformly, the latter is not desired as it can create "hot spots" in the room.

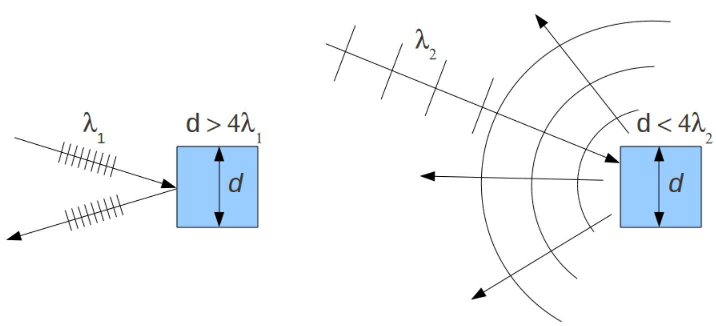


Figure 15.4: Reflection of two different wavelengths from the same box. The box has dimensions such that $4d > \lambda_1$ (the smaller wavelength wave on the left) for which regular reflection occurs, and $4d < \lambda_2$ (the larger wavelength on the right), for which reflection is diffused and the wave spreads throughout the room.

15.2.2 Dispersion

When sound reflects from small objects (defined roughly as those that have dimensions *smaller* than about 4λ), they do not reflect in a "regular" way, but rather become dispersed into the surrounding environment - see figure 15.2b. Thus, depending on the wavelength, an object in a room may disperse the sound of some frequencies and not others.

Example 15.1

Regular and Dispersive Reflection *A box sitting on a coffee table has dimensions 30 cm on a side. What is the frequency below which sound is dispersively reflected from the box, and above which sound is regularly reflected? See figure 15.4.*

Solution: Regular reflection occurs when the dimension of the reflecting object is greater than about 4λ. Therefore, since the box has dimension 30 cm = 0.30 m on a side, wavelengths greater than about $\frac{0.30}{4}$ = 0.075 m will be dispersively reflected, and wavelengths less than about 0.075 m will be regularly reflected. The frequency corresponding to wavelength 0.075 m can be calculated from

$$f\lambda = v_s \quad \rightarrow \quad f = \frac{v_s}{\lambda} = \frac{345 \text{ m/s}}{0.075 \text{ m}} = 4600 \text{ Hz}$$

Therefore, frequencies less than about 4600 Hz will be dispersed by the box, and frequencies greater than about 4600 Hz will be regularly reflected from the box side.

15.2.3 Absorption

All material objects reflect some of the sound incident on them and absorb the rest. The most highly reflective surfaces reflect most of the sound energy falling on them and absorb very little. Highly absorptive materials do the opposite – they absorb most of the sound energy that falls on them and reflect very little. An example of the former is hardwood floors, and an example of the latter is heavy drapery.

15.2.4 Absorption of Sound by Air

As a wave propagates through the air, some of its energy is absorbed by the air. This absorption is generally quite small, and in many cases we can ignore the effect of the air on sound, compared to the effect from solid surfaces such as walls, curtains, carpet, and furniture. The contribution of air becomes important when sound propagates long distances, such as in large auditoriums or outdoors. High frequencies are more readily absorbed by the air than low.

15.2.5 Absorption By Materials

The acoustic environment of a room or hall can be greatly altered by changing the sizes and shapes of surfaces making up the interior volume and the types of materials from which they are formed. Hard, smooth surfaces tend to reflect more than they absorb, while softer, more porous surfaces tend to absorb more than they reflect.

Each material can be characterized by an *absorption coefficient*, denoted by α, which ranges between 0 and 1, and whose value indicates how much of the sound incident on it is absorbed. Therefore a value $\alpha = 0$ corresponds to a perfectly reflecting material, for which no sound is absorbed (0% absorption), and a value of $\alpha = 1$ corresponds to a perfectly absorbing material, for which all sound is absorbed (100% absorption). Both of these values represent extremes that are rare in materials – all materials reflect a little *and* absorb a little. Note that the absorption coefficient is a unit-less quantity.

Absorption coefficients are also wavelength/frequency dependent – the size of the features that cause absorption for one frequency might not absorb other frequencies as well. A surface that may reflect fairly well at low frequencies (for example a low-pile carpet), with a correspondingly low coefficient

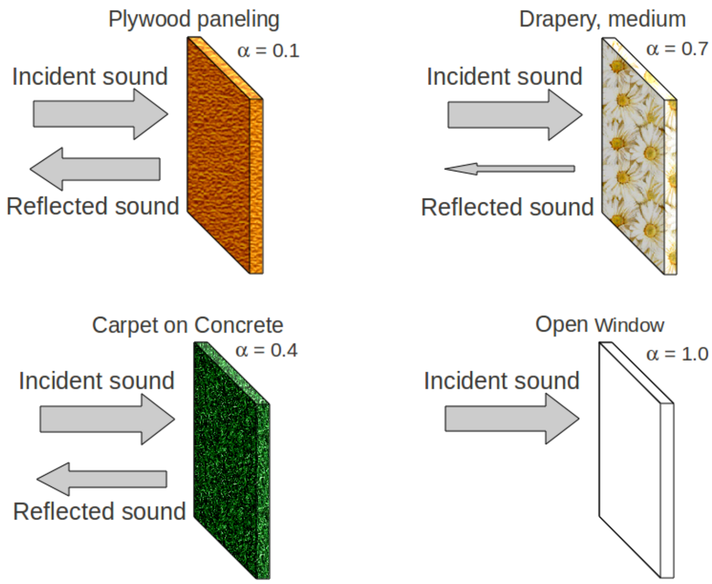


Figure 15.5: Four surfaces with different absorption coefficients. Note that the amount of reflected sound energy is inversely related to the absorption coefficient – the more sound is absorbed, the less is reflected.

(α closer to 0) may absorb better at high frequencies (α closer to 1). Table 15.1 lists values for many common materials found in rooms and concert halls.

15.2.6 Absorptivity

The actual amount of sound energy absorbed by a surface is dependent not only on its material makeup but is also directly proportional to its surface area. The total absorptivity (otherwise referred to as the *absorption*, which is how we will refer to it) for a surface is defined as

$$\boxed{\text{Absorption:} \qquad A = \alpha S} \tag{15.1}$$

where A is the absorption, S is the surface area in square meters (m^2), and α is the absorption coefficient of the material. The unit for absorptivity is the "sabin," in honor of Walter Sabine, a pioneer in the science of acoustics. One sabin is equivalent to one m^2.

Example 15.2

Absorption for Different Surfaces *What is the total absorption of the following surfaces for sound of frequency 1000 Hz? a) 2.5 m^2 of heavy drapery, b) an unpainted concrete block wall measuring 3×5 square meters, and c) a ceiling of suspended acoustic tile, measuring 5×4 square meters?*

Solution: a) The absorption coefficient for heavy drapery (at 1000 Hz; see table 15.1) is 0.72. The absorption is therefore

$$a_d = \alpha_d s_d = 0.72\,(2.5)\,\text{m}^2 = 1.8 \text{ sabins}$$

Table 15.1: Absorption coefficients for various surface materials. These numbers are approximate; actual materials may vary somewhat from these values.[1]

	Absorption Coefficient at Frequency (Hz)					
Surface Material	125	250	500	1000	2000	4000
Acoustic tile, rigidly mounted	0.2	0.4	0.7	0.8	0.6	0.4
Acoustic tile, suspended in frames	0.5	0.7	0.6	0.7	0.7	0.5
Acoustical plaster	0.1	0.2	0.5	0.6	0.7	0.7
Ordinary plaster, on lath	0.2	0.15	0.1	0.05	0.04	0.05
Gypsum wallboard, 1/2"on studs	0.3	0.1	0.05	0.04	0.07	0.1
Plywood paneling, 1/4"on studs	0.6	0.3	0.1	0.1	0.1	0.1
Concrete block, unpainted	0.4	0.4	0.3	0.3	0.4	0.3
Concrete block, painted	0.1	0.05	0.06	0.07	0.1	0.1
Concrete, poured	0.01	0.01	0.02	0.02	0.02	0.03
Brick	0.03	0.03	0.03	0.04	0.05	0.07
Vinyl tile on concrete	0.02	0.03	0.03	0.03	0.03	0.02
Heavy carpet on concrete	0.02	0.06	0.15	0.4	0.6	0.6
Heavy carpet on felt backing	0.1	0.3	0.4	0.5	0.6	0.7
Platform floor, wooden	0.4	0.3	0.2	0.2	0.15	0.1
Ordinary window glass	0.3	0.2	0.2	0.1	0.07	0.04
Heavy plate glass	0.2	0.06	0.04	0.03	0.02	0.02
Draperies, medium velour	0.07	0.3	0.5	0.7	0.7	0.6
Upholstered seating, unoccupied	0.2	0.4	0.6	0.7	0.6	0.6
Upholstered seating, occupied	0.4	0.6	0.8	0.9	0.9	0.9
Wood/metal seating, unoccupied	0.02	0.03	0.03	0.06	0.06	0.05
Wooden pews, occupied	0.4	0.4	0.7	0.7	0.8	0.7
Open window, archway, *etc.*	1.0	1.0	1.0	1.0	1.0	1.0
Air, per 100 m^3, relative humidity 50%	0.0	0.0	0.0	0.30	0.90	2.40

b) The absorption coefficient for an unpainted concrete block at 1000 Hz is 0.29, so that

$$a_w = \alpha_w s_w = 0.29\,(3 \times 5)\,\mathrm{m}^2 = 4.35 \text{ sabins}$$

c) The absorption coefficient for a suspended acoustic tile at 1000 Hz is 0.99, so that

$$a_c = \alpha_c s_c = 0.99\,(4 \times 5)\,\mathrm{m}^2 = 19.8 \text{ sabins}$$

Example 15.3

Adding Drapery *The walls of a particular room measuring 5 m × 4 m × 3 m consist of plywood paneling.*

a) Calculate the total absorption of just the walls, at a frequency of 1000 Hz.

b) Suppose we want to triple the overall absorption of these walls by adding a certain amount of drapery. What area of drapery should we add?

Solution: The wall material has absorption coefficient 0.1 at 1000 Hz. The absorption of the walls without the drapery is

$$A_{old} = A_w = \alpha_w S_w = 0.1\,(2 \times 5 \times 3 + 2 \times 4 \times 3)\,\mathrm{m} = 0.1 \cdot 54 \text{ m} = 5.4 \text{ sabins}$$

We now want to cover the walls with drapery, but we don't yet know how much area we need in order to double the total absorption – let's call the area of the draperies S_d. The coefficient of drapery at 1000 Hz is 0.7. In order to calculate the new absorption, we need to start from the absorption of the walls alone, then "subtract" the portion of the walls that will be covered by drapery, and then add the drapery:

$$A_{new} = 3A_w = \alpha_w S_w - \alpha_w S_d + \alpha_d S_d = \alpha_w S_w + (\alpha_d - \alpha_w)\,S_d$$

$$\alpha_w S_w + (\alpha_d - \alpha_w)\,S_d = 3\alpha_w S_w \quad \rightarrow \quad S_d = \frac{2\alpha_w S_w}{(\alpha_d - \alpha_w)} = \frac{A_{new} - A_{old}}{(\alpha_{new} - \alpha_{old})}$$

15.2.7 Absorption and Room Features

Taking into account *all* of the surface features in a room and their individual sizes and absorption coefficients, we can calculate the *total* absorption A_{tot} for the room. The total absorption, measured in units of sabins, or m^2, corresponds to the "equivalent" number of square meters of 100% absorbing surface in the room. Think of two extreme examples:

1. Imagine a room in which all surfaces are made of a special material that absorbs *all sound falling on it* – a perfectly absorbing room. Since all surfaces have absorption coefficient of 1, then the total absorption A_{tot} for the room will be *equal to its total surface area.*
2. The other extreme example would be a room with *perfectly reflecting surfaces*, for which the absorption coefficient for all surfaces is 0. In this case, the total absorption A_{tot} for the room would be 0, meaning that sound is 100% reflected.

Both of these extreme examples do not represent physically realizable situations, although the first (a totally absorbing room) is more feasible than the second. The total absorption for physical rooms would fall in between these two extremes, meaning that the total absorption A_{tot} would have a value in between 0 and A sabins, where A is the total square area of all the walls, ceiling, floor, and objects in the room. A room might have, for example, a total absorption of $\frac{A}{2}$ sabins, implying that the average absorption coefficient for all surfaces is somewhere around 0.5.

Example 15.4

Total Absorption for a Room *A room has dimensions 6 m (width) × 8 m (length) × 3 m (height). The walls and ceiling are 1/2" wallboard, the floor is a wooden platform floor. The room has an archway measuring 2 m × 2.5 m and two glass windows, each measuring 1 m × 2 m. What is the total absorption of this room at 1000 Hz?*

Solution: The absorption coefficients for the materials are the following: ceiling and walls, $\alpha_c = \alpha_w = 0.04$; wooden floor, $\alpha_f = 0.2$; archway, $\alpha_a = 1.0$; and glass windows, $\alpha_g = 0.1$.

The total absorption of the ceiling is

$$A_c = \alpha_c S_c = 0.04\,(6 \cdot 8)\,\mathrm{m}^2 = 1.92 \text{ sabins.}$$

The total absorption for the floor is

$$A_f = \alpha_f S_f = 0.2\,(6 \cdot 8)\,\mathrm{m}^2 = 9.6 \text{ sabins.}$$

The total absorption for the walls (minus the area for the archway and two windows) is

$$\begin{aligned} A_w = \alpha_w S_w &= 0.04\,(6 \cdot 3 + 6 \cdot 3 + 8 \cdot 3 + 8 \cdot 3 - 2 \cdot 2.5 - 2 - 2)\,\mathrm{m}^2 \\ &= 0.04\left(75\ \mathrm{m}^2\right) = 3.0 \text{ sabins.} \end{aligned}$$

The total absorption for the two glass windows is

$$A_g = \alpha_g S_g = 0.1\,(2 + 2)\,\mathrm{m}^2 = 0.4 \text{ sabins.}$$

And finally, the archway represents an opening to the outside. Since sound does not "reflect" from an opening but rather leaves the room, the "absorption" coefficient is equal to 1.0. The total absorption of the archway is then

$$A_a = \alpha_a S_a = 1.0\,(2 \cdot 2.5)\,\mathrm{m}^2 = 5.0 \text{ sabins.}$$

The total absorption for the entire room is then

$$A_{tot} = A_c + A_f + A_w + A_g + A_a = (1.92 + 9.6 + 3.0 + 0.4 + 5.0)\,\text{sabins} = 19.9 \text{ sabins.}$$

Given that the total surface area of the room is $A_{room} = (2 \cdot 6 \cdot 3 + 2 \cdot 8 \cdot 3 + 2 \cdot 6 \cdot 8)\,\mathrm{m}^2 = 180\ \mathrm{m}^2$, a total absorption of 19.9 sabins means that the average absorption coefficient is around $\frac{19.9}{180} = 0.11$, corresponding to a pretty "live" room with loud acoustic.

15.2.8 Noise Reduction

The room considered in the previous example ended up being fairly "live," owing to the many hard reflective surfaces present. We can "soften" the acoustic (*i.e.* increase the overall absorption) by replacing some of the surfaces with more highly absorptive materials. Increasing the absorption reduces the ambient noise in the room caused by environmental noises. The more "live" the room, the higher overall background noise level, owing to the high reflectivity of the room. Increasing the overall absorption therefore serves to diminish the background noise. When the total absorption of a room changes as a result of modification, the resulting *noise reduction* is defined as

$$\boxed{\text{Noise reduction:} \quad NR = 10 \log \frac{A_2}{A_1}} \tag{15.2}$$

where A_1 and A_2 are the initial and final total absorption values for the room. The units of NR are db, and a positive value corresponds to a net *reduction* of the noise level in the room.

Example 15.5

Noise Reduction *Imagine that we now wish to "soften" the acoustic for the room in the previous example. We do so by installing suspended acoustic tile over the entire ceiling, and covering the floor with heavy carpet on felt backing. What net noise reduction do we achieve by these modifications?*

Solution: The new absorption coefficient for the ceiling, consisting now of suspended acoustic tile, is $\alpha_c = 0.7$, and the new absorption coefficient for the carpet-covered floor is $\alpha_f = 0.5$. The total absorption for the ceiling now becomes

$$A_c = \alpha_c S_c = 0.7\,(6 \cdot 8)\,\text{m}^2 = 33.6 \text{ sabins}$$

and the new total absorption for the floor becomes

$$A_f = \alpha_f S_f = 0.5\,(6 \cdot 8)\,\text{m}^2 = 24 \text{ sabins}$$

so that the total absorption for the room now becomes

$$A_{tot} = A_c + A_f + A_w + A_g + A_a = (33.6 + 24 + 3.0 + 0.4 + 5.0)\ \text{ sabins} = 66 \text{ sabins}.$$

The new absorption is $A_2 = 66$ sabins and the old absorption was $A_1 = 19.9$ sabins so that the net noise reduction resulting from our modifications on the room is

$$NR = 10 \log \frac{A_2}{A_1} = 10 \log \frac{66}{19.9} = 5.2 \text{ db}.$$

We have thus reduced the ambient noise in the room by approximately 5.2 db. We could do more by installing soft furniture, placing pictures on the wall, putting drapery over or around the windows, *etc.*

15.3 Sound Resonance in a Small Enclosure

We tend to think of sound waves as being basically "one-dimensional" as they propagate through space (the air molecules move back and forth along one direction, that of wave motion). There are, however, more complicated arrangements that can exist when sound is confined to small enclosures. Recall our discussion of the one-dimensional string, and the straightforward way in which the harmonic frequencies could be expressed mathematically. We derived the various wavelengths that could be supported by the string given the boundary condition of being tied down at the ends. We then derived the sequence of supported higher frequencies using these wavelengths and the speed of wave propagation on the string. In this case, the higher modes of vibration worked out to be *integral multiples* of the fundamental (i.e. harmonics). Then, recall the more complicated patterns of vibration on a *two*-dimensional surface, the drum head. The boundary condition for this system corresponds to the circular perimeter, where the skin is pulled tight over the rim. The higher modes of vibration for this system, though utilizing mathematics too advanced for this course, nevertheless have a well-defined mathematical form. These frequencies are not, however, integral multiples of the fundamental, but are rather *overtones*, and the sound from this system is not particularly "tonal," since overtones don't support the vibration of the fundamental.

So now let us move to three dimensions. When we consider an enclosure in which standing waves of sound vibrate, we'll find that it, too, has a complicated sequence of resonant frequencies. For a *rectangular* enclosure, the mathematical expression for these higher modes of vibration is not particularly complicated. The formula resembles that for the one-dimensional tube,

$$\boxed{f_{n_x,n_y,n_z} = \frac{v}{2}\sqrt{\left(\frac{n_x}{L_x}\right)^2 + \left(\frac{n_y}{L_y}\right)^2 + \left(\frac{n_z}{L_z}\right)^2}} \tag{15.3}$$

where n_x, n_y, *and* n_z are integers, and L_x, L_y, *and* L_z are the dimensions of the enclosure along the x, y, and z axes. Let's see how this plays out with an example. Let's assume that you are in an enclosure, say a shower stall (where your singing voice always seems better!) measuring 0.8 m along x, 1.3 m along y, and 2.5 m along z. The lowest mode vibration will be "one-dimensional," *i.e.* as if the shower stall is simply a "tube" with a sound resonance along *one* axis. The longest wavelength supported by the enclosure is along the vertical, since this is the largest of the three dimensions. This mode of vibration is denoted f_{001}, meaning that the resonance has $n = 1$ along z, and 0 (*i.e.* no resonance) along x and y. Let's compute this resonant frequency. Note that since in this example n_x and n_y are 0, the formula above simply reduces to a very familiar formula, that of an open-open tube:

$$f_{001} = \frac{n_z}{2L_z}v = \frac{1}{2L_z}20.1\sqrt{T_A}.$$

Assuming a room temperature of 22°C (295 K), we see that the resonant frequency for the lowest supported mode of vibration in the shower stall is

$$f_{001} = \frac{1}{2\,(2.5)}20.1\sqrt{295} = 69 \text{ Hz},$$

which corresponds to a fairly low tone. The next highest one-dimensional mode of vibration will be along the y axis (the next longest dimension), which is

$$f_{010} = \frac{1}{2\,(1.3)} 20.1\sqrt{295} = 132 \text{ Hz}.$$

Table 15.2: Table of resonances in a small rectangular enclosure

n_x	n_y	n_z	$f_{n_x n_y n_z}$ (Hz)
0	0	1	69.0
0	1	0	132.8
0	0	2	138.1
0	1	1	149.7
0	1	2	191.6
1	0	0	215.8
1	0	1	226.5
1	1	0	253.4
1	0	2	256.2
1	1	1	262.6
0	2	0	265.6
0	2	1	274.4
1	1	2	288.5
1	2	0	342.2
1	2	1	349.0
1	2	2	369.0
2	0	0	431.5
2	0	1	437.0
2	1	0	451.5
2	1	1	456.8
2	1	2	472.1
2	2	1	511.4
2	2	2	525.2

What are the next highest frequencies? It may not be particularly intuitive; it does not correspond to the $n_x = 1$ mode with frequency $f_{100} = 216Hz$, but rather (in order) $f_{002} = 138Hz$, $f_{011} = 150Hz$, *and* $f_{012} = 192Hz$ (you should calculate these yourself for practice). The values for the higher frequencies are listed in table 15.2, and plotted in figure 15.6.

The following plot shows the large number of resonances which are fairly closely spaced. Note especially that the higher frequencies are not, in general, integral multiples of the fundamental.

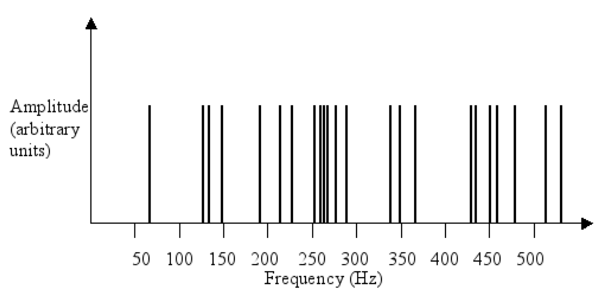


Figure 15.6: Sound spectrum showing the large number of sound wave resonances for rectangular enclosure with dimensions $0.8 \times 1.3 \times 2.5$ m^2.

From the large number of resonances, it is easier to understand why singing in the shower makes the voice sound better. The fundamental of the sung pitch, plus the accompanying harmonics are largely supported by the enclosure, and if you happen to hit the pitch of one of the resonant frequencies, your entire voice is well supported by the enclosure. So singing in pitch tends to be easier with the rich feedback of sound you receive from the enclosure.

The modes for which only one of the three indices n_x, n_y or n_z is non-zero can be viewed in much the same way as the one-dimensional tube vibrations we studied for the open woodwind instruments. However, when two of the three indices are non-zero, it's better to think of the vibration patterns along the same lines as those we studied for the two-dimensional drumhead. When all three indices are non-zero, the volume is filled with a complicated 3-dimensional vibrational pattern that's not particularly easy to visualize.

15.4 The Acoustical Environment

Our singing on the flagpole example from earlier in the chapter underscored the fact that when there is no acoustic environment present (or more accurately, no local reflecting surfaces), the sound intensity emitted by the source drops with the square of the distance, and any sound that is perceived from a distance corresponds only to "direct" sound, since there is no reflected sound present. The falloff of sound intensity in a room is quite different, owing to the presence of many reflecting surfaces (see figure 15.7). In a medium-sized room, when we are very close to the source, intensity falls off as the distance squared, as it did in the outdoor case; reflected sound is relatively weak since it originates from the far walls, ceiling, *etc.* As we move farther from the source, reflections from local surfaces contribute more substantially to the intensity reaching the listener. After a certain distance, the intensity remains constant (*i.e.* it levels off) as we move farther away, owing to the support from the reflective surfaces. For small rooms, the sound intensity levels off to a constant value not far from the source.

The presence of reflecting surfaces in a room not only increases the amount of sound received by a listener (reflecting surfaces re-direct some of the sound to the listener that would otherwise have been lost), but it also adds an acoustic enhancement through its time of arrival relative to the direct sound, contributing *reverberation*.

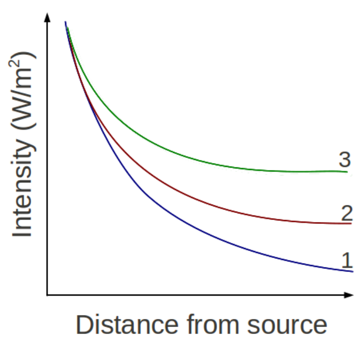


Figure 15.7: The intensity of sound as a function of distance. Curve 1: no surrounding acoustic environment, sound intensity falls off as the distance squared from the source. Curves 2 and 3: In large-medium and medium-small rooms, intensity deviates from inverse distance squared dependence sooner, and levels off at some distance as reflected sound supports the intensity.

15.4.1 Buildup and Decay of Sound

Imagine that a key on a piano is suddenly struck, after which the key is held down briefly. The sound is one of an abrupt "attack" of sound, followed by a smooth steady tone. A listener some distance away first hears the sharp sound directly from the instrument, *i.e.* the sound that travels directly from the source to the observer (see figure 15.8), which is then joined by several reflected versions of the sound from the walls, ceiling, floor, and furniture. Each reflected sound comes with a slight delay depending on the distance it traveled. Sound reflecting once before reaching the listener tends to arrive sooner, followed by sound reflected two and three times. The time delays between the arrival of the direct and first reflected sound, and the arrival of other versions, are very small, not able to be detected as separate sounds. Rather, the initial attack of sound is "rounded out" or "blurred" by the presence of this fast sequence of slightly delayed reflections – sound which normally would have been abrupt and sharp (the "attack" of the piano sound) has been augmented in volume and "spread out" in time by the presence of reflected sound.

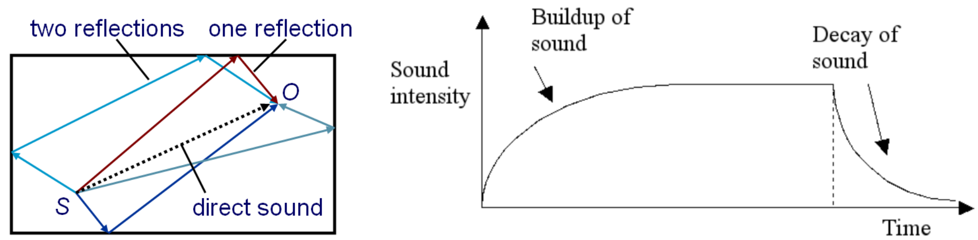


Figure 15.8: (***For color version see Appendix E***) Sound reaching an observer consists of direct plus reflected sound. The buildup of sound results from the slight delay between the direct and reflected sound.

The reverse of this process occurs when the note is ended abruptly a second or two later. When the direct sound from the piano first ceases, the direct sound ends first for the listener, during which the various reflected sounds are still "in flight." As the listener receives each successive cessation of reflected sound, the volume dies out until at last (typically on the order of a second later) the sound ceases entirely. The shape of the sound intensity during cessation is essentially the reverse of the buildup – see figure 15.8.

15.4.2 Reverberation

The persistence of sound after the direct sound has ceased is referred to as reverberation. The time it takes for a specific sound to die away is related to a room's reverberant character. By definition, the *reverberation time* is defined as the time it takes for a sound to drop by 60 db in sound level, equivalent to a drop in intensity of 1 million. At 1000 Hz, this corresponds to a "perceived" drop of a factor of about 70 (see figure 7.7 of chapter 7 to convince yourself of this).

The reverberation time is characteristic to a room, dependent on its size, and the degree to which its surfaces absorb sound. How might we go about predicting the reverberation time for a given room? We can at least reason along the following lines. The persistence of sound in a room is enhanced by the presence of reflecting surfaces, not absorbing surfaces. Hence we would predict that the reverberation time should be approximately inversely related to the total absorption: $T_R \propto \frac{1}{A_{tot}}$. Basically, the more absorbing surfaces, the faster the decay of sound – energy is removed from the room by the surfaces. Also, the length of time it takes for the sound to die out increases when the reflected sound must travel farther. Therefore we would predict that the reverberation time will be roughly proportional to the size (volume) of the room: $T_R \propto V_{room}$. The reverberation time is given by

$$\boxed{\text{Reverberation time (in seconds) for a large enclosure:} \qquad T_R = \frac{0.16V}{A + \alpha_{air} V}} \tag{15.4}$$

where V is the volume of the room in m^3, A is the total absorption of all surfaces in the room in sabins, and α_{air} is the absorption coefficient for air. When the volume of an enclosure becomes large, the absorption by the enclosed air becomes an important contributor to the reverberant character of the room as it contributes more substantially to the overall absorption of sound. However, for small and medium environments it is often fine to ignore the air term in the denominator:

$$\boxed{\text{Reverberation time, small to medium enclosure:} \qquad T_R = \frac{0.16V}{A}}. \tag{15.5}$$

Remember that *both* the volume of the room V and the total absorption A contribute to the reverberation time. While large halls tend to have long reverberation, and small halls and rooms tend to have short reverberation, it is entirely possible to have a large hall with low reverberation time (large absorbing surfaces), as it is possible to have a small room with large reverberation time (lots of reflecting surfaces). The best acoustic environment is one in which the room size and the type and location of absorbing surfaces are combined in such a way as to maximize the acoustic needs of the room, whether they be for music, speech, rehearsal, or others.

15.4.3 Benefits and Drawbacks of Reverberation

Do we desire long reverberation times, or short, in order to maximize the listening experience? The answer, in part, depends on what kind of music we are considering. Large cathedrals with large reflecting surfaces, where organ and vocal ensemble performances typically occur, have rather long reverberation times, lending a spacious, “heavenly” element to the sound. Music that fares well in such environments is “broad” and “expansive,” with less need for precise, clean articulation. Large reverberation tends to “mix” the sounds of former notes with succeeding ones, giving more of an overlap to the sound. Baroque ensembles, on the other hand, don’t fare well as well in such large reverberant environments, since the music is characterized by a precise and quick articulation of sound that would blur with the larger reverberation time. So let’s consider for a moment what the effects of reverberation are.

Generally speaking, longer reverberation time brings greater “warmth” to the sound, and shorter reverberation time brings greater clarity and precision to the sound. Speech fares best in an environment with relatively low reverberation, owing to the need for crisp articulation of sound with appropriate separation of the consonants, *etc.* Public lectures generally fare best, therefore, in smaller, more intimate environments with low reverberation times where words don’t blend together. When large reverberant halls are necessary, especially where electronic amplification is present, slower speaking delivers better clarity and understanding.

Short reverberation times allow for crisp articulation, good for speech and music with clean precision. This environment lacks some of the “warmth” and expansive atmosphere of larger halls. Would we fare well, then, with no reverberation time? The complete absence of reverberation gives an eerie sense, an “empty” environment where sound is not supported well and therefore is quiet. Speaking in such an environment requires extra effort to be heard. The walls of the enclosure, which give rise to reverberation, also support the sound by reflecting it back into the room to support the direct sound. Since the path it takes is longer than that for direct, it follows the direct sound in time.

Reverberation allows for the *build-up* of sound in an enclosure. A listener receives the direct sound from the instrument, followed quickly by several reflections from various parts of the surrounding room. The direct sound takes the least time to arrive, since it travels along a straight line at the speed of sound. The next sound to arrive will be a reflected one, along the shortest of all possible reflection paths. This will be followed by other reflections. However, it’s important to note that the time gap between the direct and various reflected sounds is typically too short to discern, and so the experience is as "one sound" whose volume is enhanced by the presence of reflections which add to the direct sound. Indeed, the quality that makes an outdoor Greek amphitheater function well is the presence of reflected sound that enhances the direct. When a "pin" is dropped on the stage, the direct sound alone is probably not enough to be heard on its own from a seat in the middle of the audience, but the bowl-like shape of the seating area serves to allow many reflections to enhance the sound.

The time it takes for sound to build is similar to the reverberation time for the decay of sound, but the two are not necessarily equal, since the origination of sound is from a *point* in the room, whereas decay begins from sound spread throughout the room. Imagine a lone oboist on stage who begins to play a long, melodious note, and you listen from a seat in the middle of the hall. The sound originates from a point source on stage, and your listening spot is quickly filled with the rich addition of direct and reflected sound. In order for the sound to be deemed acoustically pleasant, the first reflected sound shouldn't lag the direct sound by more than about 0.030 seconds, a time called the "initial delay gap." This first reflected sound might be, for example, reflected from the ceiling at a point somewhere between you and the instrumentalist. The second reflected sound, following close on the heels of the first, might come from one of the side walls, *etc.* The time it takes the sound to climb to its maximum volume will be on the order of a second or so for a medium-sized room. As the oboist continues to play the note long, an equilibrium in sound level is established, and all energy being generated by the oboe exactly balances the energy removed from the room by absorption. When the oboe releases the note, the sound proceeds to die out over the course of a second or so.

15.4.4 Initial Time Delay Gap

Another important feature of a room acoustic has to do with how "smoothly" the sound from a source builds in volume from reflected sound support. If you are in a very large room with your eyes closed, your ears can sense the expansive size of the environment when the sound has the faint character of an echo. The ear discerns a noticeable delay between the arrival of sound traveling directly from the source and sound reflecting from one of the room surfaces, signaling a large acoustic enclosure.

The distance in time between the direct sound and the first reflected sound is defined as the "initial time delay gap." Figure 15.9 illustrates the principle.

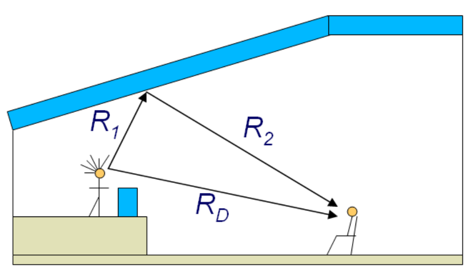


Figure 15.9: Illustration of the initial time delay gap, corresponding to the time between the arrival of direct sound traveling path R_D, and reflected sound, traveling paths R_1 and R_2.

Two paths of sound from the lecturer are shown – the direct and the first reflected sounds. There are, of course, other versions of reflected sound not shown, but the first is the most important since this is the one that immediately follows the direct sound. The reflected sound travels paths R_1 and R_2 for a total distance of $R_1 + R_2$. The direct sound travels distance R_D. What is the time difference between the two? The time it takes the direct and reflected sounds to travel from the lecturer (me) to the student (you) are, respectively,

$$t_d = \frac{R_D}{v_s} \qquad t_R = \frac{R_1 + R_2}{v_s}$$

so that the difference between the two, defined as the initial time delay gap, is

$$\boxed{\text{Initial time delay gap:} \quad t_I = t_R - t_D = \frac{R_1 + R_2 - R_D}{v_s}} \tag{15.6}$$

where v_s is the speed of sound, or 345 m/s at room temperature. In a good acoustic environment for musical performance, this delay should not exceed approximately 0.030 s, preferably below 0.020 s. Table 15.3 lists some of the initial time delay gaps for various famous concert halls.

Table 15.3: Initial time delay gaps for 4 famous concert halls, along with their average reverberation times.

Concert Hall	t_I **Midmain Floor**	t_I **Midbalcony**	Reverberation time
Boston Symphony Hall	0.015 s	0.007 s	1.8 s
Cleveland Severance Hall	0.020 s	0.013 s	1.7 s
New York Carnegie Hall	0.023 s	0.016 s	1.7 s
Philadelphia Academy	0.019 s	0.010 s	1.4 s

Only since 1900 has acoustical knowledge been applied to the design of concert halls. Old concert halls, built much earlier than this, likely have good acoustics, but that is precisely why they are old. It doesn't necessarily mean that the people that constructed it had some secret knowledge, but that if the acoustics were poor (and I'm sure many were built that were poor), the hall would most likely have been torn down and replaced by something else at some point. Avery Fisher Hall is a good example of a hall that had many renovations in hopes of improving its acoustics. Finally, an acoustician was in charge, and had the place literally gutted and re-built according to good acoustical principles, after which it became famous for its good sound.

The science of acoustics is largely understood, but it is still a developing science with questions that still require answers.

15.5 Chapter Summary

Key Notes

- When sound waves interact with a surface, they can undergo reflection, dispersion, and absorption. The nature of reflection depends on the shape of the surface. Convex shapes are helpful for acoustic environments intended for speech or performance since they serve to disperse sound widely, whereas concave surfaces detract from the acoustic quality of a room since they can focus sound and create hot spots.
- Different materials absorb sound to differing degrees as characterized by their absorption coefficients. The absorptivity of a room can be derived by calculating the contribution of all the surfaces to the absorption of sound.

- Noise reduction can be calculated from a change in the absorption of a room after changing some of the surfaces.
- Small enclosures can produce complicated resonance frequency structures owing to the three-dimensional nature of the resonating sound waves.
- The buildup and decay of sound depends on the degree to which sound reflects efficiently from the surfaces and can enhance or detract from the acoustical environment depending on use of the space.
- Reverberation is a measure of how long it takes for sound to die away in a room.
- The initial time delay gap is a measure of the time between the direct and first reflected sound.

Exercises

Questions

1) Define briefly the initial time delay gap. Is it desirable to have a long or a short initial time-delay gap, and why?
2) What is reverberation time? What main feature of a room contributes to its ultimate value? Is it desirable to have a long or short reverberation time?
3) Is the nature of regular reflection of sound from objects in a room the same for all frequencies? Explain briefly.
4) Why are there often so many more resonant frequencies for a 3-dimensional enclosure compared to a 1-dimensional system (such as vibrations on a string)?
5) What is the difference between regular reflection and dispersive reflection?
6) Why is it that a material might have a different absorption coefficient for different frequencies?
7) Is it always important to take into account absorption of sound by air? When is it most important?

Problems

1. At a frequency of 500 Hz, what is the absorption of a piece of plywood paneling that measures 2 m × 3 m? How does this differ from its absorption at a frequency of 125 Hz?

2. A room has dimensions 5.5 m (width) × 7 m (length) × 3.5 m (height). The walls consist of gypsum wallboard, the ceiling has suspended acoustic tile, and the floor is covered with vinyl tile on concrete. There are two large glass windows on the wall, each measuring 2 m × 1.5 m. Calculate the total absorption of this room at 1000 Hz. We then remove the tile and install heavy carpet, and hang three paintings on the wall, each with absorption coefficient 0.35 (at 1000 Hz), one measuring 0.8 m × 0.4 m, and the other two measuring 1 m × 1.4 m. what is the noise reduction we have achieved with this modification?

3. A closet has dimensions 2 m × 1.5 m × 3 m. What are the three lowest resonant frequencies of this enclosure?

4. For the example of the shower stall in section 15.3, verify the resonant mode frequencies f_{010}, f_{200} and f_{221}.

5. Calculate the frequencies of the first three resonances of a room with dimensions 5 m × 10 m × 2.5 m. Do they have any significance acoustically?

6. What is the reverberation time for the room in question 2 before the modifications were made? How about after the modifications?

7. Explain why ceilings in newer auditoriums commonly consist of several large panels slanted different ways to make a wrinkled or corrugated pattern instead of being entirely flat.

8. Visit Winter Hall Room 210 and observe the large wooden panels sitting above the stage area. Comment on their shape, and whether this is acoustically desirable or not, and why.

9. It turns out that 1/4"and a 1/2"plywood panels have practically the same absorption coefficients for high frequencies, but quite different ones at low frequencies. Explain why this is the case, and explain which of the two will have higher absorption coefficient for low frequencies, and why.

10. Imagine that our gymnasium dimensions are 35.5 m in width, 52 m in length, and 16.5 m in height. All absorption coefficients are given for sound at 1000 Hz. The walls are painted concrete block with absorption coefficient 0.07, the ceiling is made of material that is basically equivalent to plywood paneling with absorption coefficient 0.1, and the floor is hardwood with absorption coefficient 0.05. (Of course we are ignoring the bleacher stands and other items in the room, which would serve to absorb more than the surfaces so far mentioned.)

    a) What is the reverberation time of this gymnasium for a frequency of 1000 Hz?

    b) Estimate the largest initial time delay gap for the gym. (Hint: Imagine standing at the very center of the floor and speaking to someone a few feet away from you.)

    c) Comment on the acoustics of this enclosure for the performance of music.

    d) Of course our gymnasium is much better acoustically than the enclosure depicted above. Why is it that music for Chapel sounds much better than it would if played in the gym described above? Comment briefly.

11. Why is the initial time delay gap larger for a seat position in the midmain floor of a concert hall than for a seat in the midbalcony region?

12. A room has dimensions 30×20×8 m and has a reverberation time $T_R = 1.2$ s. Is this reasonable when compared to the total actual surface area for the room? Comment briefly.

13. Read a short passage aloud to estimate a typical speaking rate in terms of (a) syllables per second, and (b) phonemes per second. Suppose intelligibility requires that each syllable or phoneme, on average, drops approximately 10 db in sound level before the next one comes along. This is equivalent to saying that the time intervals between them should be $1/6^{th}$ of the reverberation time T_R. Use this criterion with your estimated syllable and phoneme rates to make two estimates for the maximum acceptable T_R for speech. Discuss whether you think it more appropriate to apply this criterion to syllables or to phonemes.

14. In the room shown in figure 15.10, how long does it take for (a) the direct sound, (b) the ceiling reflection, and (c) the back wall reflection to get from the sound S to the listener L? Discuss the acceptability of these delays.

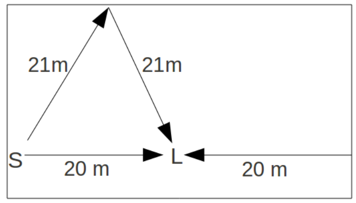


Figure 15.10:

References

[1] Backus and Doelle, Environmental Acoustics, pg. 227, McGraw Hill (1972)

References

Chapter 16

Recording and Playback

Over the last 100 years several technologies have developed that enable us to record, store, and play back music. Some of the earlier technologies are more straightforward, so we will start with them. The more modern, digitally-based technologies for music storage and playback are more sophisticated, and we will consider them briefly in chapter 18. At present, we will investigate three basic media used for the production of music: vinyl records and the turntable, magnetic tape, and AM and FM transmission and reception. While these seem like very different technologies, they are each based on the same fundamental phenomenon of *electromagnetic induction*, having to do with the interaction of *electric currents* and *magnetic fields*. Our plan is not to cover this topic in great detail, but only in sufficient depth to understand the basic function of these three technologies. We will also briefly consider some basic sound system components – amplifiers and speakers.

16.1 The Transducer

The *transducer* is at the heart of recording and playback technologies. In the broadest sense, a transducer is anything that changes one type of signal into another. In our case, we will be interested in devices that change mechanical motion into electrical signals or electrical signals into mechanical motion. Transducers will be involved at several stages of both the recording and the playback of music. For example, the movement of the needle, or stylus, during the playback of music from a vinyl record on a turntable is transformed into an electrical signal by the cartridge (a transducer). This electrical signal is then routed to an amplifier (which increases the size of the electrical signal), after which it is sent to the speaker, where the electrical signal is turned into the mechanical motion of the speaker cone (by a transducer on the back of the cone) that then fills the room with sound.

16.1.1 Electricity and Magnetism

Electric and magnetic phenomena were recognized by the early Greeks. Up until around 1820 it was thought that the two were distinct and separate, but in the early- to mid-19th century, Michael Faraday, James Clerk Maxwell, and others discovered that the two were manifestations of the same *electromagnetic* phenomenon.

Electric currents have magnetic phenomena associated with them. What is an electric current? Inside wires, which are made of conducting metal, are many small electric charges that are free to move about inside the conductor. These charges are associated with the atoms that make up the metal. When the charges are driven to move along the wire by an electric field provided by a power supply, they constitute *electric current*. The larger the flow of charges through the wire, the higher the current. Electric current is measured in units of Amperes, or amps.

16.1.2 Electric Current and Magnetic Field

In 1820 it was discovered that electric current in a wire can deflect a nearby compass needle. Moving charges produce magnetic fields, so when they move together down the wire's length as electric current, the magnetic field they produce circulates around the wire in concentric circles (see figure 16.1).

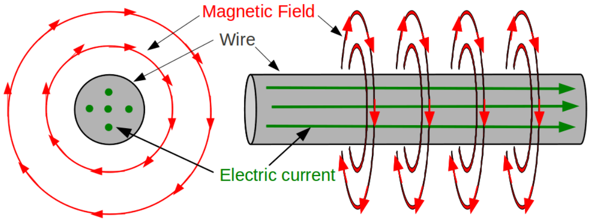


Figure 16.1: (***For color version see Appendix E***) The magnetic field associated with moving charges in a wire. Note that the field is "circumferential," meaning that it encircles the wire with a direction consistent with the "right hand rule."

As long as current flows, the magnetic field surrounds the wire. As soon as the current stops, the magnetic field disappears. The direction of the field can be determined by the "right hand rule." When you grasp the wire with your right hand and point your thumb along the direction of positive current flow, the magnetic field encircles the wire with a direction given by your curving fingers. The field appears as concentric circles around the wire, as shown in the figure. This phenomenon of magnetic field generation by electric currents forms the basis of electromagnets.

16.1.3 Force Between Electric Current and a Magnet

Now we are ready to consider the electromagnetic forces that are at the heart of transduction. You have probably noticed that when two magnets are brought close together, they each exert a force on the other, either attractive or repulsive in nature, depending on the orientation of their magnetic "north" and "south" poles. In figure 16.2, a wire carrying a current I experiences a force from the magnetic field B of the permanent magnet. The direction of the force is not entirely intuitive. The force results from the size and direction of the magnetic field B and of the current I, both of which point in different directions from one another. The resulting force on the current points neither along the direction of the magnetic field nor along the direction of the current, but in a direction perpendicular to them both! The "right-hand-rule" enables us to determine the direction of the force. If you take the fingers of your right hand, point them along the direction of the current I, and then curl them toward the direction of the magnetic field B, then your thumb will point in the direction of the force F. Likewise, the reverse also occurs: the electric current exerts a force on the magnet, and the direction is opposite. If one of them is mounted firmly and the other not, the latter will move as a result of the force. This basic interaction forms the basis of the electric motor, where electric current can be made to cause motion. As we shall see, this principle is also at the heart of the operation of the loudspeaker.

16.1.4 Current Generation by Moving Magnets

Given that an electric current has been found to produce a magnetic field, the question naturally arises as to whether the reverse is possible. Can a magnetic field generate electric current? Several scientists in the early 19^{th} century attempted to find an answer to this question, but it was Michael Faraday discovered the connection in 1893. It turns out that a *static* magnetic field cannot generate electric current, but a *moving* or *changing* magnetic field can! The relation is captured in an equation

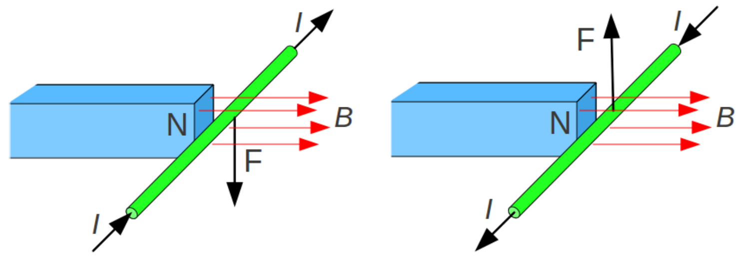


Figure 16.2: (***For color version see Appendix E***) The force exerted by an external magnetic field on a current-carrying wire. The direction of the force depends on the direction of the magnetic field and the current, and can be found by the right-hand-rule. If the fingers of the right hand are pointed along the current I, and then curl toward the direction of the magnet field B, the thumb points in the direction of the induced force.

bearing the name "Faraday's Law" in honor of its discoverer. When a magnet is moved in the vicinity of a conductor, a current is induced in the conductor's interior. If that conductor happens to be a wire-loop, the current circulates in the loop. When the movement of the magnetic is reversed, the current in the loop reverses as well - see figure 16.3. Note that it is not important to identify which is moving, the magnet or the coil. As long as there is *relative* motion between the two, the effect occurs. Also, the faster the magnet or the coil moves, the higher the current generated. This phenomenon of generating electric current by moving a magnet near a wire or by moving a wire near a magnet forms the basis of electric generators. For example, a large wheel with blades at the bottom of a waterfall can rotate under the action of the falling water. Magnets mounted to the axle of that wheel rotate near large wire coils, and electric current is generated as a result of these moving magnets.

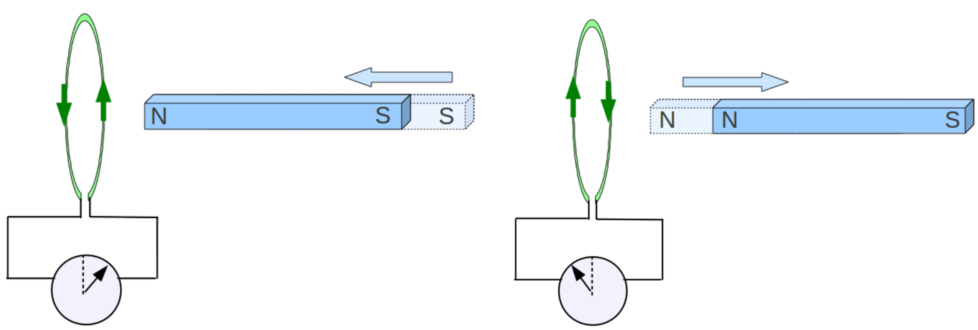


Figure 16.3: (***For color version see Appendix E***) A magnet moving in the vicinity of a wire loop generates current in the loop as long as the magnet is moving. Once it stops, the current ceases. The current changes direction when the magnet is moved in the opposite direction. The faster the magnet or the coil moves, the higher the current generated.

16.1.5 Summary of Important Electromagnetic Interactions

In summary, three basic interactions between magnetic fields and electric current, important to the operation of transducers, are the following:

1. Electric current moving through a wire has a circulating magnetic field associated with it. The direction of the field is found using the right-hand-rule. See figure 16.1. This phenomenon of magnetic field creation by electric current forms the basis of electromagnets.
2. A magnetic field exerts a force on a current-carrying wire. Likewise, the current exerts a force

on a magnet in the opposite direction with equal magnitude. If one of them is mounted, the other will move in the direction of the force exerted on it. See figure 16.2. This phenomenon of producing motion out of electric current forms the basis of motors.

3. A moving magnet or changing magnetic field induces current in a wire. It doesn't matter whether the magnet moves or changes and the wire stays still, or the wire moves near a magnetic field, since the motion between the two is relative. See figure 16.3. A very important point is that the faster the motion of the wire or magnet, the stronger the induced current. This phenomenon of creation of electric current from motion forms the basis of electromagnetic generators.

16.2 The Phonograph

From our example earlier on, you may be asking why we would want to understand the phonograph cartridge and turntable, when vinyl record technology is by now quite old and not nearly as common as the CD or MP3 player. The phonograph was the first of the high-fidelity recording and playback systems, and still remains today the medium of choice by some, although its popularity for most has been surpassed by the CD and other digital technologies. The operating principle behind the phonograph provides a great and intuitive example of transduction.

The phonograph system derives the sound from a large vinyl record with one long, spiral microscopic groove cut into its surface, starting from the outside edge and spiraling in toward the center. The record sits on a turntable, which rotates at $33\frac{1}{3}$ rotations per minute, and the stylus rides in the groove as the record rotates – see figure 16.4. The physical back-and-forth and up-and-down motion of the stylus, as it follows the complicated shape of the record groove, is translated into electrical signals by a transducer located in the cartridge, in which the stylus shaft is mounted. These electrical signals are then sent to the amplifier unit to drive the speakers and create the sound.



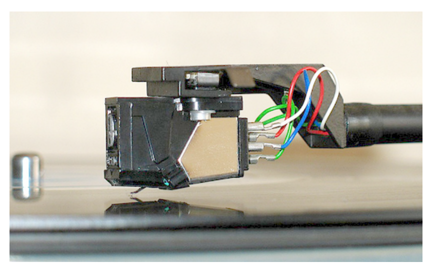


Figure 16.4: (***For color version see Appendix E***) A vinyl LP record, a turntable, and a closeup view of the cartridge and stylus sitting on the surface of the rotating record.[1][2][3]

16.2.1 The Phonograph Cartridge

The cartridge is a critical element in the phonograph system. It contains a transducer that turns the physical motion of the stylus into electrical signals. If the quality of the cartridge is not able to faithfully translate the sound recording on the record into accurate electronic signals, the quality of the rest of the sound system will suffer. The cartridge is therefore crucial to the high fidelity of the phonograph system. Figure 16.5 shows a schematic of a moving magnet (MM) cartridge. As the stylus rides in the groove of the moving record, it wiggles left and right, up and down. The magnet located at the end of the shaft connected to the stylus moves in similar fashion next to two small wire coils. Current induced in these two coils are then sent to the amplifying unit. Note the important role of transduction here. The moving stylus (physical movement) is translated into two electrical signals by the special relationship between moving magnets and the creation of electric current. There are also moving coil (MC) versions of cartridges, where the electric coils are mounted on the end of the stylus shaft (which move relative to fixed magnets).

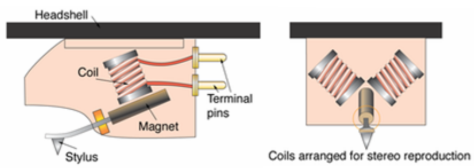


Figure 16.5: Schematic of a moving magnet cartridge. As the stylus rides in the record groove, it moves left and right, up and down. The motion of the stylus causes the small magnet in the cartridge to move in similar fashion next to two wire coils. The moving magnet induces current in each of these coils, which is then sent to the amplifier to be turned into sound.[4]

16.2.2 The Record Grooves

Let's understand how the recording of a sound waveform might appear on the surface of a vinyl disk or record. You'll recall that the waveform for a sound can be very complicated. The simplest kind of sound we first studied was that which results from simple harmonic motion, or a pure tone, which has the shape of a sine wave. The connection between waveforms and record grooves is that the grooves are cut in very much the same shape as the actual waveform – a conveniently intuitive result. When the grooves of a master record are first cut (from which copies of the records are stamped), the cutting element is driven by the electrical signals of the soundwave being recorded. The walls of the record groove are cut at an angle of 45^o with the vertical (see figure 16.7), and the detailed shape of the groove walls is governed by the sound waveform being recorded. Then, as a copy of the record is played, the stylus rides inside this track as the record rotates.

Consider the shape of the record groove that would correspond to a recording of a pure tone. If we were to magnify it, how would it appear? Figure 16.6 shows a very simple picture of how such a track would be shaped. As the record rotates, the stylus rides in the groove, and it moves back and forth in simple harmonic motion according to the track shape. The magnet at the end of the shaft on which the stylus is mounted also moves in simple harmonic motion and generates electrical currents in the two coils mounted inside the cartridge (see figure 16.5), and the induced currents also thus appear sinusoidal in shape. These currents are sent to an amplifier, which sends larger versions of these currents to the speakers. The speaker cones are driven by these sinusoidal currents so that they vibrate in simple harmonic motion and the sound of a pure tone fills the room.

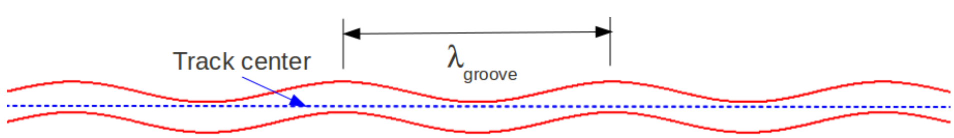


Figure 16.6: The basic shape of a record groove corresponding to a recording of a pure tone, which has the shape of a sine wave.

Note in the figure that the wiggles in a record groove have wavelength λ, but that this is clearly not the wavelength of the recorded sound, but is simply the physical distance between wiggles on the surface of the record. However, when the record rotates at $33\frac{1}{3}$ rotations per minute, the periodicity of this wiggle does translate into a particular frequency of sound. Waveforms of more complicated shape than that of a pure tone would yield more complicated groove tracks. However, no matter how complicated the waveform of music encoded on the surface of the disk, the shape of the groove walls does reflect (to a large extent) the actual waveform of the sound.

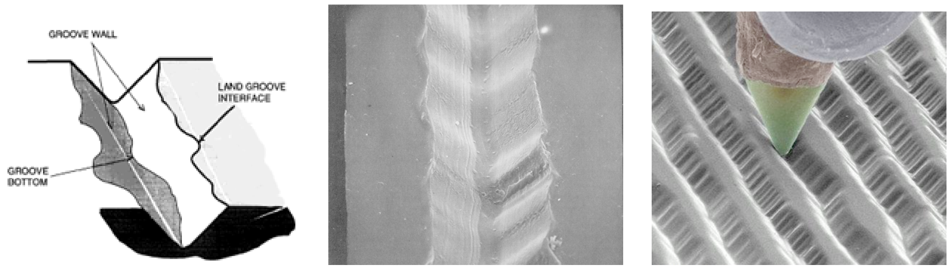


Figure 16.7: A schematic of a record groove, and magnified views of actual grooves (about 500× for image in middle, and 200× in right).[5]

The modern LP record is 12 inches in diameter. The outer track, where the music begins, has diameter 11.5 inches, and the innermost track where the music ends has diameter 4.75 inches. The average LP has about 1,500 feet of groove on each side, which amounts to about $\frac{1}{3}$ mile. While one side is typical divided up into several separate tracks of music, the groove forms one continuous spiral from the outer to the inner edge.

Example 16.1

Record grooves – distance variations

a) *How long does it take for an LP record to make one rotation?*

b) *How fast, in cm/sec, is the needle moving relative to the groove at the outer edge of the record? At the inner edge?*

c) *How far apart are successive bumps corresponding to a tone with frequency 440 Hz in the groove at the outer edge?*

d) *How far apart are the same successive bumps for the innermost track?*

Solution:

a) The rotation rate is $33\frac{1}{3}$ rotations per minute (RPM), and the time it takes for one full rotation is

$$T = \frac{1}{f} = \frac{1}{33.3333 \text{ rot}/\cancel{\text{min}}}\left(\frac{60 \text{ sec}}{\cancel{\text{min}}}\right) = 1.8 \text{ sec}$$

b) The speed of the needle relative to the groove can be calculated from the circumference $C = \pi D$ where D is the groove diameter, and the rotation period T. For the outer edge,

$$C_o = \pi D_o = \pi\left(11.5 \cancel{\text{in}}\right)\left(\frac{2.54 \text{ cm}}{\cancel{\text{in}}}\right) = 91.7\text{cm} \quad \rightarrow \quad v_o = \frac{C_o}{T} = \frac{91.7 \text{ cm}}{1.8 \text{ sec}} = 51.0 \text{ cm/s}$$

and for the inner edge,

$$v_i = \frac{C_i}{T} = \frac{\pi\left(4.75 \cancel{\text{in}}\right)\left(\frac{2.54 \text{ cm}}{\cancel{\text{in}}}\right)}{1.8 \text{ s}} = 21.1 \text{ cm/s}$$

c) In order to calculate the distance between successive bumps, we can calculate how many bumps will appear in one full circular groove, corresponding to one rotation of the disk. When the record is rotating, the frequency with which the needle encounters the bumps is 440/s. Since it takes 1.8 seconds to make one complete rotation, there will therefore be 440 /s (1.8 s) = 792 bumps/circular track. The distance d between bumps for the outer track will be the circumference divided by this number,

$$D_o = \frac{C_o}{792 \text{ bumps}} = \frac{91.7 \text{ cm}}{792 \text{ bumps}} = 0.116 \text{ cm/bump}$$

d) and for the inner track,

$$D_i = \frac{C_i}{792 \text{ bumps}} = \frac{21.1 \text{ cm}}{792 \text{ bumps}} = 0.027 \text{ cm/bump}$$

16.2.3 Monophonic and Stereophonic Encoding

In a monophonic recording, the left and right sides of the groove are shaped similarly. In a stereo (2 channel) recording, each side of the groove is impressed with a different pattern, one for each channel of sound. The movement of the stylus in this case is quite complicated. Where the two sides of the groove have bumps facing inward toward the center of the groove, it narrows, pushing the stylus upward as it moves past that point. Wherever the channel widens, the stylus moves downward (see figure 16.8).

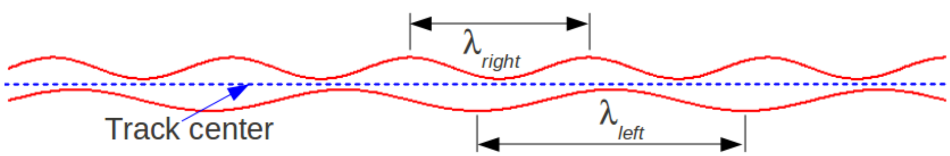


Figure 16.8: A record groove with two different track wavelengths on the two walls. In places where the channel narrows the stylus moves upward and where the channel widens it moves downward, in addition to being moved left and right by the two walls. The two coils in the cartridge can interpret this 4-way movement and decode the music into its two stereo channels.

Hence in addition to moving left and right as directed by the side walls of the groove, the stylus also moves up and down, depending on the variation in channel width. These two separate motions of the stylus are decoded by the cartridge coils. See figure 16.9 for how the left-right and up-down movement of the stylus can be detected. When the magnet moves upward, the induced current I (depicted in the plots below the coils) is positive in both coils, since the magnet moves *toward* each coil; when the magnet moves downward, negative induced current appears in both coils; when the magnet moves left, the induced currents are positive in the left coil (since the magnet approaches it) and negative in the right coil (since the magnet moves away from it); and when the magnet moves right, induced currents are negative in the left coil and positive in the right. In this way the specific motion of the magnet, and therefore of the stylus, can be decoded from the complicated motion of the magnet. The relative size and sign of the currents induced in the two coils can then be decoded into the two separate stereo channels.

16.2.4 Bass Equalization and Pre-Emphasis

Recall that earlier in the chapter we learned that when a magnet is moved near a wire, the strength of an induced current signal is highly dependent on the relative speed with which the magnet and wire

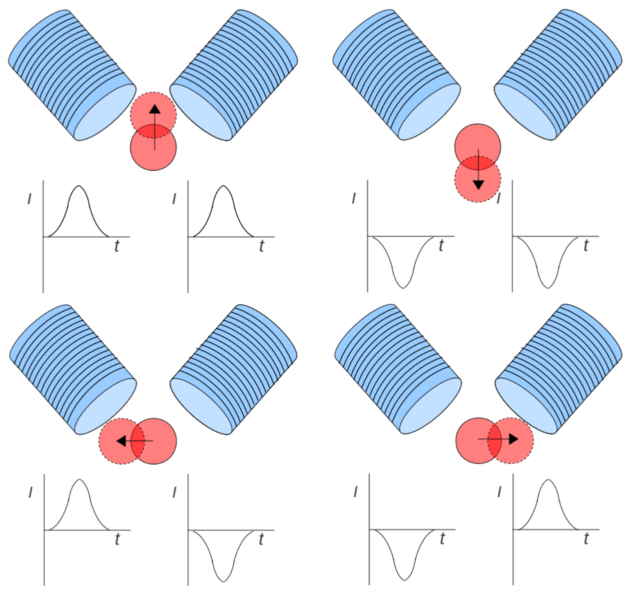


Figure 16.9: Magnet and dual coils in a stereo phonograph cartridge. Two coils (upper cylinders) are positioned near the magnet (circle - shaft is seen end-on).

are moved. The faster the relative motion, the larger the induced current. Therefore a record groove corresponding to a low volume tone should have very little left-right swing, and a groove corresponding to a loud tone should have a very large left-right swing. Imagine two grooves on a record with very different frequency but the same amount of left-right swing amplitude (see figure 16.10). When the stylus moves through the portion of the record with the high frequency groove, the rapid movement of the needle back and forth will cause the magnet also to move rapidly near the induction coils, and a large current will be induced because of the high speed of the magnet. When the stylus moves through the low frequency tone on the groove, even though the groove has the same left-right swing amplitude, the stylus moves much more slowly since its wiggle frequency is much lower. Then the magnet also will move more slowly near the coils, resulting in a much weaker induced current. The net result is that for tracks with the same left-right swing amplitude but different tone frequencies, the induced current will be much weaker for the low frequency tone.

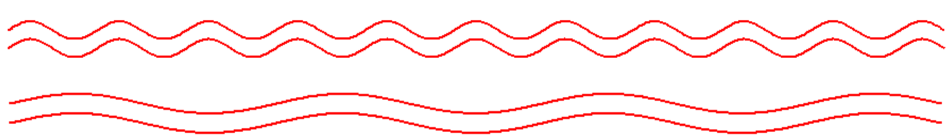


Figure 16.10: Two record grooves with different track wavelengths, corresponding to a high frequency tone (upper) and a low frequency tone (lower). Both tracks have the same stylus swing amplitude, but the induced current for the lower frequency tone will be much lower than that for the high frequency tone, owing to the lower relative speed of the magnet and the induction coils.

Is this a serious problem? Ostensibly no, since in the recording process we would simply need to make sure that lower frequency tone tracks are given a much larger swing amplitude so that they produce the right volume when played back. But there is, in fact, a practical problem here. The large swing amplitude for the bass tone tracks would either cut into their neighboring tracks, which is unacceptable, or the tracks would have to be separated by such space as to limit how much music could be placed on one disk. There are two competing concerns here – we want to fit enough music on one LP (approximately 30 minutes per side), and we want bass tones to play with the appropriate volume. Luckily, a trick is employed that satisfies both of these concerns.

Below a certain frequency, called the "turnover frequency," the bass tones are artificially reduced in volume before being recorded onto the record surface, meaning that the groove ends up with lower swing amplitude. The lower the frequency, the greater the adjustment to the volume before recording (see figure 16.11, solid curve). This process of adjustment is called *bass equalization*. If the record were played back with no recognition of the bass equalization made during the recording process, the resulting bass tones would sound very weak and the music would lack fidelity to the original performance. When the record is played back on a system that recognizes and compensates for this adjustment, the bass tones will be appropriately amplified in volume according to the dashed curve in the figure. Some amplifiers have a special "phono" input for connecting phonograph units. The circuitry inside the amplifier knows to apply bass equalization correction to signals connected to this port.

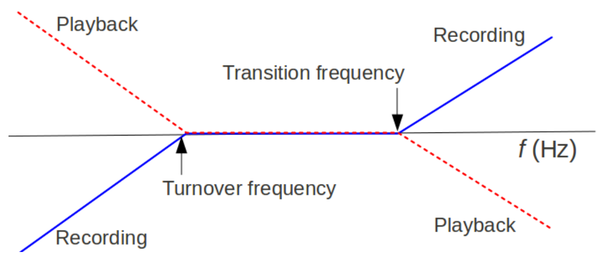


Figure 16.11: The amplitude modification of low and high frequency tones prior to the recording process for LP records. Low tones are reduced in volume before recording, in order to keep track width appropriate, and high frequencies are emphasized in order to assist with hiss and crackle noise reduction.

In summary, then, in the process of recording music onto a record disk, the lower frequencies are progressively lowered in volume before they are cut into the groove, and correspondingly raised in volume in the playback stage. This allows grooves to be narrow, and hence allows for the appropriate audio time to be stored on one record side. The frequency below which this electronic compensation is made during the recording and playback stages is called the turnover frequency.

The story with high frequencies is different. Small amplitude signals of high frequency have no problem being heard. Very small imperfections in the grooves are also heard easily, such as dirt and scratches. These generate unwanted high frequency noise such as hissing and crackling. In order to reduce the effects of this noise, a different trick is employed. In the recording process, before musical frequencies above a certain frequency called the *transition frequency* are recorded, they are artificially *amplified* in signal. The higher the frequency, the higher the applied amplification (see figure 16.11 again). The result of this *pre-emphasis* is that the wiggles in the grooves corresponding to these frequencies end up having larger swing amplitude than the music requires. If played back on a system that does not apply the appropriate compensation for these highly amplified tones, the sound is too loud in the upper frequency range. So, during playback, the required compensation is to appropriately lower the amplitude of high frequency signals (according to the dotted curve in the figure) to reverse the pre-recording emphasis. The effect of this is to reduce the sound from the imperfections in the grooves, resulting in better signal-to-noise ratio. The primary purpose of this pre-emphasis is to reduce surface noise on the LP.

16.3 The Turntable

The turntable provides a means for rotating the record at a very steady RPM (rotations per minute), and holding the cartridge at the proper position in the record groove. There are three types of drive systems for turntables, the rim drive, the belt drive, and the direct drive. The speed of rotation determines the pitch of the music, and must therefore be kept very even. Low quality turntables can introduce slight variations in rotation speed, producing unwanted effects called "wow" and "flutter," which correspond to slow and rapid unevenness in pitch. "Rumble" is unwanted low-frequency mechanical noise caused by the turntable mechanism.

The tonearm is precisely engineered to place the stylus in the record groove. The most common type is the pivoted tone arm. Optimally, the center line of the cartridge should be exactly parallel to the tangent line of the record groove, and the tone arm is responsible for this alignment. The vertical tracking force (VTF) is the net downward weight of the stylus in the groove, and is usually set to about the weight of 1.75 to 2 grams of mass. This may not sound like a lot, but because of the very small footprint of the stylus in the groove, the pressure that this small weight applies, over such a small area, amounts to about 340 pounds per square inch! With that in mind it is not difficult to understand why the groove walls wear out with use.

Because of the way the tone-arm geometry is arranged, there is a slight force pulling the tonearm toward the record's center, causing unequal wear on the two groove walls. If the stylus gets bumped up out of the groove, this force will also cause the stylus to "skate" across the record toward the center. In order to compensate for these effects, an "antiskating" force is applied to help balance the force so that the wear on the two record walls is even.

The mechanical assembly of the stylus, cartridge, and tonearm has its own resonant "bouncing" frequency. Slight warps in the vinyl record could potentially cause bouncing resonances in stylus, so tonearms are typically to be designed so to have little or no resonant response in the frequency range of 0 to 7 Hz, which is typically where record warps cause variations in height.

16.4 The Tape Deck

The process of recording and playback using tape decks is based on the same basic electromagnetic principles. Tape recording and playback technology is based on the ability to "freeze" a pattern of magnetized particles, or domains, in the surface layer of a long thin plastic tape. The tape typically consists of a thin cellulose acetate base, coated with a much thinner layer of iron oxide, which contains countless magnetized particles able to be oriented in any direction by exposure to an external magnetic field.

The magnetic particles in a blank, unrecorded tape point in random directions in the iron oxide layer (see figure 16.12). The particles in a tape with recorded music point in directions that reflect the shape of the musical waveform. In much the same way as we saw for phonograph recording, where the detailed shape of the record groove was directly related to the shape of the musical waveform, if you had eyes to "see" the magnetization pattern in a recorded tape, you would recognize the musical waveform as well. In places where the waveform amplitude is high (*i.e.* loud passages), a large number of the randomly orientate particles would be lined up, whereas in passages where the amplitude is low (quiet passages) only a few of the randomly orientated magnetic particles would be lined up.

16.4.1 The Recording Process

In the recording process, the goal is to store the shape of the musical waveform into a pattern of magnetism in the iron oxide layer. Iron is one of the few elements whose atoms have a specific electron configuration that produces a magnetic field of its own. The atoms of iron are similar to very tiny bar magnets. Small numbers of iron oxide molecules in the thin tape layer cluster and align together to

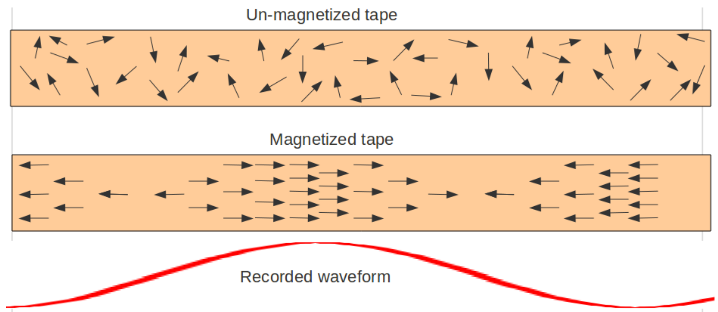


Figure 16.12: Sections of recording tape. The blank, unrecorded tape (upper panel) contains randomly oriented magnetic particles, where as the recorded tape (lower panel) the magnetic particles are aligned in the pattern of the musical waveform (in this case a portion of a pure tone).

form magnetic "domains," which have a well-defined magnetic field. These domains can be rotated by exposure to a strong magnetic field. The basic idea of recording, then, is to store musical information in the iron oxide layer, in the form of a magnetization pattern along its length, that conforms to the waveform of music recorded on it. Here's basically how the process of recording works. The audio signal is transformed into an electrical current by the process of transduction from a microphone. The current is fed to the wire coil at the top of the recording head, creating a magnetic field that fills the ring around which it is wrapped (see figure 16.13). The ring is made of iron, a magnetizable material, so the magnetic field lines conform to its shape, appearing as large closed circles.

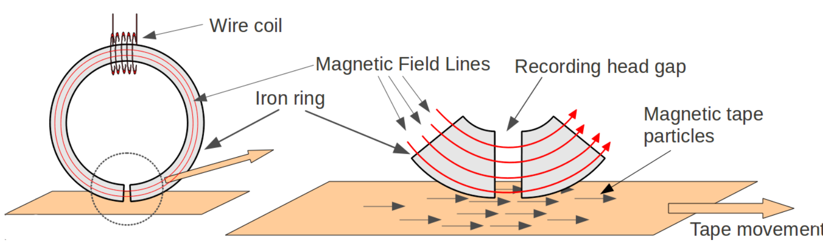


Figure 16.13: (***For color version see Appendix E***) The tape head, capable of both recording and playback. When recording, variable currents in the wire coil produce variable magnetic fields in the ring. As the tape moves under the head, the fluctuating magnetic field aligns the magnetic particles in the tape. During playback, the magnetic particles in the tape create a magnetic field in the ring. This fluctuating field induces currents in the coil that can be amplified and sent to the speakers.

At the bottom of the ring is a very small gap, which is typically 1 to 4 μm, or millionths of a meter, wide. The magnetic field lines cross this gap right above where the tape is located, and the presence of the field causes the magnetic particles in the tape layer to align as the tape moves underneath it. As the musical waveform changes in a complex way in time, the current in the coil varies in the same fashion, creating a fluctuating magnetic field in the ring that alternately points clockwise and counterclockwise, and therefore left and right in the gap at the bottom, with varying strength. When the field points left, the particles are aligned left; when the field points right, the particles are aligned right; when the field is weak, few of the randomly oriented particles are lined up; and when the field is strong, most of the randomly oriented particles are lined up. As the tape moves underneath the head (at a rate of $1\frac{7}{8}$ inches per second for a cassette tape), the fluctuating magnetic field, following the shape of the changing waveform of music, is turned into a pattern of magnetization that varies with direction and strength. The waveform is thus "frozen" into the length of the tape in the iron oxide layer.

16.4.2 The Playback Process

The pattern of magnetization in a recorded tape creates a weak magnetic field just outside the surface of the tape. The field along the length of the tape points alternately left and right, according to the pattern of magnetism frozen into the iron oxide layer. When this tape is passed underneath and very close to the gap of the playback head, the iron ring picks up on this surface field and conveys it to the pickup coil at the top (look again at figure 16.13). The fluctuating magnetic field in the ring induces alternating currents in the coil which are then passed on to the system for amplification before being sent to the speakers. The playback process is essentially the reverse of the recording process. Some tape players use the same head for recording and playback; others have separate heads for each process.

In the earliest monophonic tape machines, the entire width of the tape was used for one track of music. Later, the tape was divided into two separate tracks in order to record and playback stereo music - see figure 16.14. Two track tapes must be rewound in order to be played back again. Four track tapes were introduced so that music could be recorded and played back in two separate directions, requiring simply that the tape reel be turned over to continue. Cassette deck tapes employ a narrow 4-track tape in a plastic case that can be flipped over to play the other side.

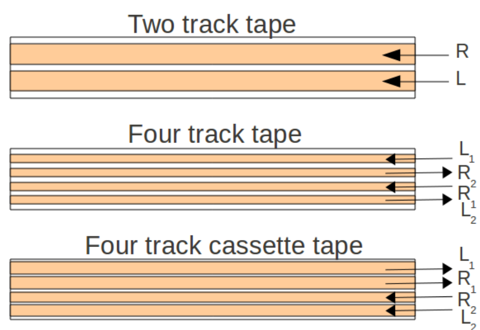


Figure 16.14: Three different tape types: High fidelity 2-track stereo, high fidelity bidirectional 4-track stereo, and 4-track lower fidelity bidirectional cassette tape.

16.4.3 Considerations for Good Fidelity

As stated before, tape that has not yet been recorded has randomized magnetization. When played, this produces a hissing sound, consisting of all frequencies (so-called "white" noise). When quiet music is being played, this hiss can be heard underneath the sound, especially at low sound levels. The higher quality the tape, the lower the overall hiss, but there is always some present because of the random tape magnetization. Good quality tape has strength, resistance to stretching, good flexibility, and smoothness. The quality of the metal surface varies. "Pure metal" coated tapes are often claimed to have the highest fidelity performance. The speeds with which tapes are drawn through the mechanism also vary. The drive mechanism needs to move the tape past the head at constant speed in order that constant pitch be maintained. Speed inconsistencies lead to wow and flutter in the played-back music, which can be the fault of the drive mechanism or a stretched or worn-out tape. Tape units operate at various tape speeds, such as $7\frac{1}{2}$ ips (inches per second), $3\frac{3}{4}$ ips, or $1\frac{7}{8}$ ips (for cassette tapes). The higher the tape speed (with shorter play time) and the wider the tracks, the higher the overall fidelity. Some professional studio reel-type tape machines operate at 15 ips.

Tape heads both record and play back music. The lowest number of heads a deck can have is two, one that serves both to record and play back, and one to erase (thus randomizing the tape particles). Higher quality decks have separate record and playback heads. One reason for separating them out is that the optimal head gap differs for recording (ideally 4 μ m wide) and playback (ideally 1 μ m wide) to avoid losses. A dual-purpose record/playback head must strike a compromise between the two head gaps. Some systems employ a third head, used to monitor the signal being recorded just after it is transferred to the tape.

16.4.4 Noise Reduction – Dolby

Dolby noise reduction was developed to greatly reduce the effects of tape noise through electronic means. You'll recall that when pre-emphasis was applied to the recording of LP records, the effect was to reduce the noise of crackles and pops from small particles and scratches in the record grooves. This was achieved by first, during the recording process, raising the sound level of high frequency tones considerably before recording them into the record groove. In this way the wiggles in the groove for these high frequencies ended up much larger than they normally would have been, and when played back would therefore produce excessive volume on playback, were it not for the appropriate reduction of volume activated in the playback process. As the playback reduction for high tones is applied, the crackles and pops of the record imperfections were also greatly reduced in volume, thus achieving a level of noise reduction.

Two common types of noise reduction for tape systems are Dolby B, which achieves around 10 db of noise reduction, and Dolby C, which consists of two two Dolby B applications in sequence, which achieves around 20 db of noise reduction.

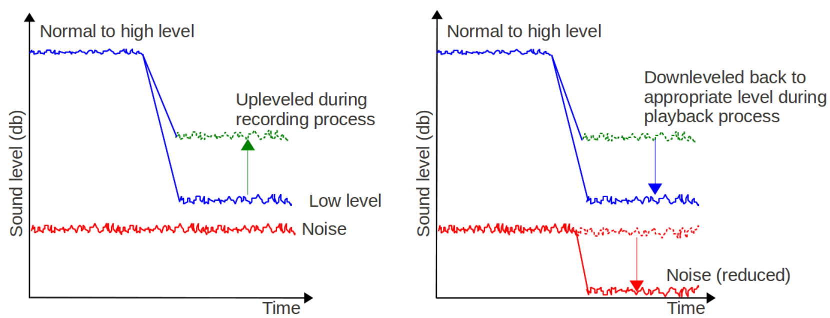


Figure 16.15: (***For color version see Appendix E***) The effects of Dolby noise reduction technology. Low intensity level sounds are up-leveled during the recording process, and down-leveled in the playback process. The net effect of this is to reduce the noise by the same down-level amount during playback.

The Dolby process for tapes is analogous to the process of pre-emphasis for records. Here's how it works. Tape noise is barely noticeable when the volume of music is very high. Thus no noise reduction is really necessary as long as the music maintains strong volume level. But as soon as the music quiets down, the hiss can be heard much better. In order to reduce the effects of this hiss, the low-intensity signals are up-leveled (*i.e.* increased in intensity) as they are recorded – see figure 16.15. Electronic circuitry automatically up-levels the sound as soon as it drops below a certain intensity level, and on playback the opposite is applied at the appropriate level of intensity. No correction is made for sound levels above the threshold value of intensity. Note that in order for this noise reduction to work properly, it must be employed during both the recording and playback processes.

16.4.5 Sound Compression - DBX

Recall that the range of human hearing is approximately 120 db, meaning that the sound level difference between the faintest sounds we can detect and the loudest sounds that begin to produce pain is 120 db.

As we will see in chapter 18, the dynamic range of compact disc (CD) technology is 96 db (which is sufficient to capture live music with very high fidelity). The dynamic range of an LP record technology is around 70 db, and that of a cassette tape is only around 60 db.

What are the limitations behind the tape's low dynamic range? On the low end of volume, the tape is limited by noise from the random magnetization. On the high end, it is limited by what is called non-linear saturation, where the magnetization of the tape reaches its highest value – when all of the magnetic particles are aligned, and they can't align any further – the magnetization "saturates." Sound at levels higher than this saturation limit are "clipped," meaning that the highest peaks and lowest valleys of the waveform become "flattened" since the magnetization cannot exceed its saturated value to faithfully record the highest and lowest points in the waveform curve.

A technology that has addressed this dynamic range limitation of tape is called DBX, and in brief, operates as follows. Dolby noise reduction is really a form of audio *compression*, in the following sense. The *dynamic range* of the music, when recorded using Dolby noise reduction, ends up lower than the original source's dynamic range, since the lowest volume tones have been increased to higher levels. DBX technology takes this approach, but applies the sound level adjustment across the entire dynamic range of music. The low sound levels are up-leveled and the highest sound levels are down-leveled, and the algorithm is continuously applied across all levels in the dynamic range.

What starts out with a dynamic range of perhaps 95 db (typical of high fidelity music) is "compressed" down to 65 db by up-leveling sound levels below an average level, and down-leveling sound levels above that average level. Figure 16.16 illustrates the process. On the left, the entire dynamic range of perhaps 95 db is squeezed down to approximately 60 db by a continuous level adjustment applied across the entire dynamic range. This reduced range is recorded onto the tape. On the right, during playback, decompression is applied to bring the music back up to its original dynamic range. As with all noise reduction and compression technologies, it is important that a DBX tape is played with DBX decompression, or the it will not sound correct.

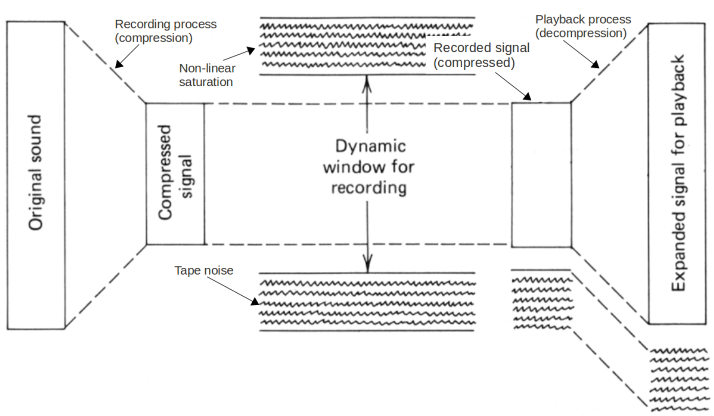


Figure 16.16: DBX technology involves both up-leveling the lowest volume passages of sound as well as down-leveling the highest sound levels of sound. The net compression allows the dynamic range of the music to be adjusted to a level able to be recorded onto the tape. The dynamic range is decompressed back to its original dynamic range upon playback.[1]

16.5 Radio Transmission and Reception

The phenomena of radio transmission and reception are also based on electromagnetic principles similar to the ones we've discussed so far. There are no transducers involved (at least not in the transmission and reception of radio waves) – everything is electronic. The only transducer to note is the speaker that creates the final sound waves.

16.5.1 Production of Radio Waves

Let's consider how radio waves are produced. Recall that a current carrying wire produces a magnetic field. When the current changes, the field correspondingly changes. The faster the current varies, the more quickly the magnetic field changes as well. However, there comes a point when the frequency of current variations brings about a new and "magical" phenomenon. When the current begins fluctuating at a rate of something like 10,000 Hz or more, the magnetic field fluctuations begin *inducing electric fields*, which in turn induce magnetic fields, *etc.*, and an *electromagnet wave* is born, which propagates away at the speed of light. This is how radio waves are created, by very quickly alternating electric currents. The electromagnetic wave consists of interlocked electric and magnetic fields oscillating together as they travel through space, much like a flapping butterfly. And because they consist of electric and magnetic fields, radio waves passing by a conductor affect electric charges in it.

Radio waves, microwaves, infrared, light, ultraviolet, X-rays, and gamma rays are all examples of electromagnetic waves. They are all manifestations of the same phenomenon, but at very different frequencies. All electromagnetic waves travel at the speed of light, which is 3.0×10^8 m/s, or 186,000 miles per second. And they are all created by moving charges. The full band of AM radio signals consists of frequencies extending from 535,000 Hz to 1,605,000 Hz (or 535 kHz to 1.605 MHz) with wavelengths ranging from 561 m to 187 m. The full band of FM radio signals consists of frequencies extending from 88,000,000 to 108,000,000 Hz (or 88 MHz to 108 MHz), with wavelengths ranging from 3.41 m to 2.78 m. The FCC (Federal Communications Commission) allocates to each radio station a small "band" of frequencies in which to broadcast.

On a related note . . .

Radio Waves from Space

The universe is filled with electromagnetic radiation in the form of radio waves. Very large dish antennas around the world are used in radio astronomy research to study these transmissions and to try and determine their origin. Astrophysical objects identified as sources of radio waves are numerous and include stars, galaxies, as well as a whole new class of very strange and mysterious objects including radio galaxies, quasars, and black holes. Astronomical radio waves are produced when large numbers of charges circulate in stellar and galactic magnetic fields. The entire universe is also bathed in a microwave background radiation, believed to be the leftover remnant of the Big Bang explosion. Since radio waves have such long wavelengths (some are much larger than the earth) one needs a very large telescope to resolve detailed images of the sources. For this reason, radio telescopes around the world are used in sync with one another (a very substantial technological feat) to form very large baseline telescopes with effective sizes as large as the earth!

16.5.2 Reception of Radio Waves

So far we see that quickly alternating currents can produce electromagnet waves, and so a natural question to ask is whether the reverse can also happen – if it is possible for electromagnetic waves to produce alternating electric currents. The answer is yes. Radio waves traveling through the air cause the creation of currents in metals. Charges in conducting materials that are free to move respond to passing radio waves by oscillating at the same frequency as the wave, producing an alternating current. This is the principle behind the reception of radio waves by an antenna. The antenna is very simply a conducting rod filled with charges able to move freely throughout its interior. When these charges are set into motion by a passing radio wave, the current produced in the antenna can be sent to the appropriate circuitry to filter and amplify the signal to produce sound.

So the question naturally arises, how can radio waves be *encoded* to carry information in the form of music (and talk shows, news, *etc.*). After all, the frequencies of AM and FM transmission are *far* in excess of the frequencies associated with sound ($\sim$ 20-15,000 Hz). In the case of LP records, the wiggles in the record grooves contain a pattern reminiscent of the musical waveform recorded onto it. Likewise, the pattern of atomic magnetization in the metal layer of a recording tape contains a pattern reminiscent of the musical waveform recorded onto it. When played back, the variations in the grooves of the record and in the magnetization patterns of the tape are already at the frequencies of sound and can be directly amplified and sent to the speakers. So in the case of radio waves and sound, the question boils down to where and how the wiggles are stored in the electromagnet wave and retrieved by the tuner.

16.5.3 AM – Amplitude Modulation

Both AM (amplitude modulation) and FM (frequency modulation) radio waves consist of a *carrier signal* that contains the encoded sound waveform. The idea for transmission of audio signals is somehow to *superimpose* the audio signal onto this carrier wave. The process of *modulating* a signal involves varying its amplitude, frequency, or phase, or any combination of them. To produce an AM radio wave, the carrier signal is sent through an electric modulator circuit that causes its amplitude to vary in a way that is reflective of the sound waveform with which it is being encoded. Figure 16.17 illustrates the process of signal encoding schematically.

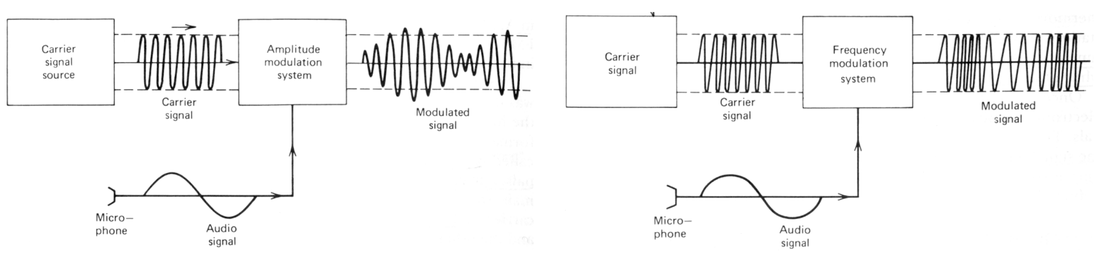


Figure 16.17: The idea behind amplitude and frequency modulation to encode a carrier signal with an audio waveform.[1]

A carrier wave of constant amplitude and frequency is fed into an amplitude modulation system. This modulator also takes as an input the audio signal to be encoded. The modulator adjusts the amplitude of the carrier signal according to the audio signal pattern. In the figure, a sinusoidal audio signal turns into a sinusoidal variation of the carrier signal amplitude. More complicated audio waveforms would accordingly result in the same complicated amplitude variation – see figure 16.18.

16.5.4 FM – Frequency Modulation

The same basic principle applies to the production of FM radio waves, except that instead of varying the amplitude of the carrier wave, the frequency is modulated. Since a typical FM frequency is around

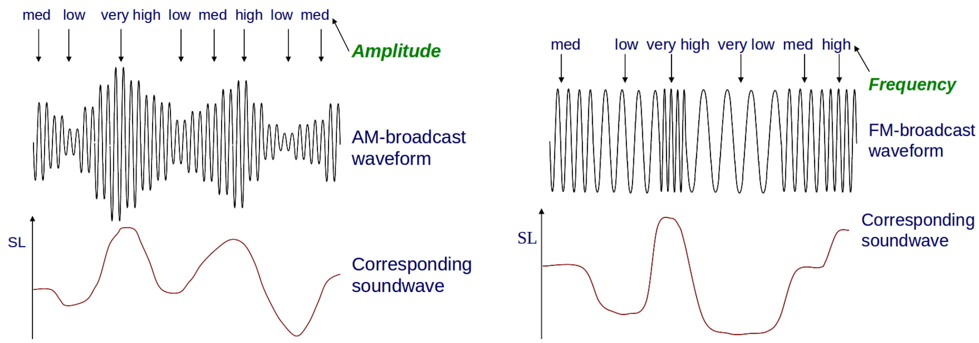


Figure 16.18: Amplitude-modulated (AM), and a frequency modulated (FM) electromagnetic waves. On the left, variations in the sound curve encoded in this radio wave appear as changes in the *amplitude* of the carrier wave. On the right, variations in the sound curve appear as changes in the *frequency* of the carrier wave.

100 MHz, variations in frequency corresponding to those present in audio signals are very small in comparison. Audio signals vary in frequency from about 50 Hz to 15,000 Hz. Compared with the frequency of the carrier signal, this means that variations in the signal produced by the modulator circuit will end up with variations of around 0.00015, or about 0.015%. Thus it is entirely feasible for a radio station to stay well within its FCC allocated bandwidth when encoding the carrier signal.

Unlike AM radio, which is monophonic, FM radio has the ability to broadcast stereo music. An FM station sends out three separate signals, a main carrier frequency, a "subcarrier" frequency that differs from the main carrier by 38 kHz, and a pilot frequency that differs from the main carrier by 19 kHz. The left (L) and right (R) stereo channels are encoded into these three in the following way. The main carrier contains the sum of both the L and R channel signals, *i.e.* the waveform consisting of $L + R$. The subcarrier frequency carries the signal corresponding to the difference between the two stereo channels, $L - R$. A radio that can only play monophonic music will utilize only the main carrier signal, which contains all of the music. A stereo radio will add and subtract the main and subcarrier frequencies to obtain the two separated L and R channels,

$$(L + R) + (L - R) = 2L \quad \rightarrow \quad \text{directed to the left channel speaker}$$

$$(L + R) - (L - R) = 2R \quad \rightarrow \quad \text{directed to the left channel speaker.}$$

The pilot frequency activates the circuitry in the radio tuner and helps with matching the frequencies and phases of the signals.

16.5.5 AM/FM Tuner

And now we come to the final topic on radio transmission and reception – the tuner. The air is *filled* with electromagnetic waves from countless sources at many different frequencies and intensities. There are many AM and FM stations, amateur radio signals, shortwave broadcasts, citizen's band radio, television, aviation communications, police bands, cellphones, GPS signals, *etc.* (not to mention the numerous radio waves from space). The job of any tuner is to filter out all of the signals except the one desired. The electrons in the radio antenna respond to any and all of these broadcasts. There is no differentiation; they don't know the difference, and will respond to any electromagnetic wave that passes in their vicinity, and there are many indeed. The signal they pass on to the tuner consists of everything that makes the electrons move. This signal is one huge superposition of all signals affecting the antenna. And it is the tuner's job to filter out just the one desired.

The tuner consists of a special filtering circuit that amplifies one small band of frequencies out of the multitude it receives from the antenna – see figure 16.19.

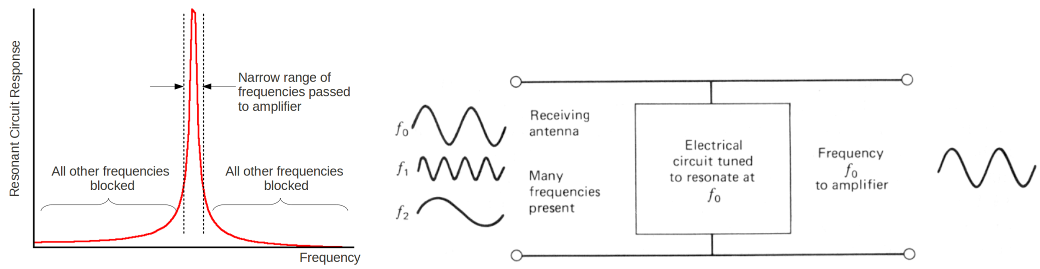


Figure 16.19: The basic operation of a radio tuning circuit. Out of the many frequencies present in the signal from the antenna, it filters out one band of frequencies corresponding to a single station.[1]

The way it accomplishes this is through resonance. An electronic circuit can be constructed so that it resonates at a particular frequency, and over a very narrow range. When the circuit is tuned to a particular frequency, all frequencies except those near resonance are blocked, and only those in the close vicinity of the resonance are passed on to the amplifier. The knob or button that serves to change the station simply alters the resonance character of the circuit, allowing one to change the station by simply changing the frequency at which the tuner resonates.

16.6 High Fidelity Playback

The process of transduction from a moving magnet to electrical current (for example, from the wiggles of the phonograph stylus to the induced current in the cartridge) produces *low* currents. On the other hand, transduction works the other way around, using electrical current to produce motion (for example, driving a speaker cone) requires *high* current. Therefore, in between a turntable and a speaker system, an amplifier is needed that can take the low current from the phonograph cartridge and amplify it sufficiently to drive the speakers. A typical arrangement consists of a pre-amplifier followed by a power amplifier.

16.6.1 Pre-Amplifier

The pre-amplifier takes the weak signals from the various program sources (labeled, for example, auxiliary input, CD player, phono, AM-FM tuner, cassette, *etc.*) and amplifies them appropriately. Figure 16.20 shows the typical arrangement for the pre-amplifier and how it fits into the whole sound system.

Note that the phono are directed to the phono equalization circuit that applies the appropriate up-leveling of the bass tones and down-leveling of the high tones referred to earlier. All program source inputs are directed to the amplification circuit, which appropriately boosts the strength of the signals without adding any additional noise or distortion. These signals are then sent to the control center, which is controlled by several knobs (shown in the bottom of the figure), referenced below:

1. *Function Selector* – which enables the user to select which program source to use, *i.e.* phono, tape, AM or FM tuner.
2. *Speakers Switch* – used to direct the output to a particular set of speakers.

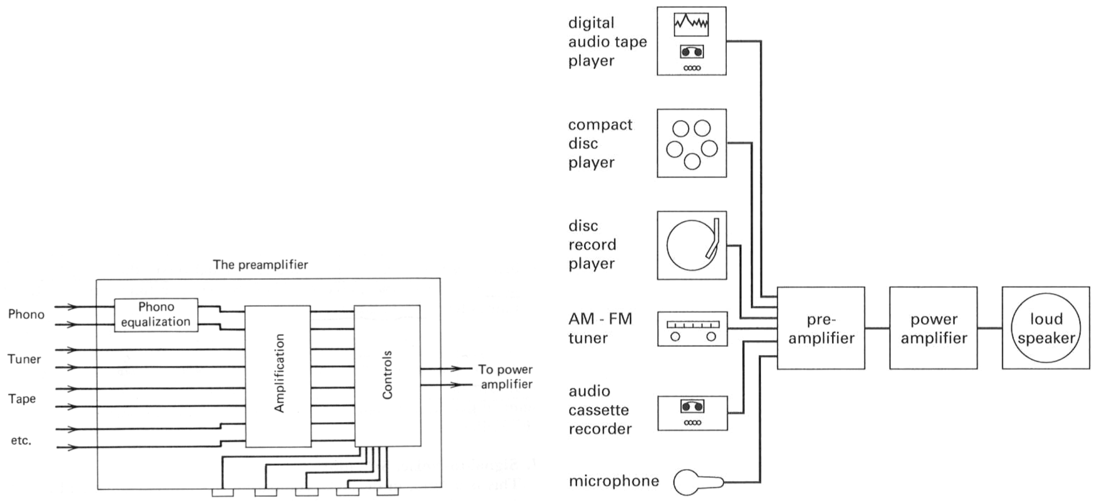


Figure 16.20: The basic function of the pre-amplifier is to take the weak signals from the program sources and amplify and shape them before sending them to the power amplifier and ultimately to the speakers.[1][7]

3. *Volume Control* – to set the desired output volume for the speakers.
4. *Balance Control* – to set the relative volume between the left and right speakers.
5. *Tone Controls* – to shape the perceived sound in various frequency ranges such as bass, midrange, or high frequency.
6. *Loudness Control* – on some older systems, used to boost the volume of the bass and treble tones relative to the midtones when the system is played at low volume. This is intended to compensate for the ear's difficulty in hearing low and high frequencies when the overall volume is low (see chapter 7 to remind yourself of the ear's frequency response).
7. *Filter Controls* – to activate high and low frequency filters to cut out the sound spectrum above and below specific frequencies. The purpose of the high frequency filter is to cut out hissing and scratching noises from records and tapes, and the low frequency filter can help reduce the sound of hum and rumble from the turntable.

16.6.2 Power Amplifier

In order to adequately drive a large speaker cone, the power amplifier needs to deliver electrical power to the speakers ranging from a few watts to tens of watts or more. The signals coming from the pre-amplifier are relatively weak by comparison, and so it is the job of the power amplifier to take these signals and significantly amplify them with as little alteration in the waveform shape as possible. Any change of the signal shape represents a decrease in fidelity. The power amplifier needs to be able to deliver power *uniformly* over the entire frequency range. For high fidelity performance, overall harmonic distortion added by the power amplifier must be kept to minimum.

Finally, the output of an amplifier needs to be matched to the speakers. The "load impedance" of the speakers, a quantity that is measured in "Ohms," needs to be matched by the amplifier so that power is effectively transferred to the speakers and none is reflected back to the amplifier. Typical load impedance for speakers is 4 to 8 Ohms.

16.6.3 Loudspeakers

Inside every speaker, including headphones, earbuds, *etc.* is a transducer, whose job it is to translate electric current into motion of a speaker cone. This process of transduction can be achieved using various methods, but the most common type for loudspeakers involves the interaction between electric current and magnetic fields. Sound is produced by a vibrating speaker cone which sets the air into motion. On the back end of the cone is positioned a coil of wire that is slipped over the pole of a permanent magnet. See figure 16.21 for a typical arrangement of speaker coil and magnet.

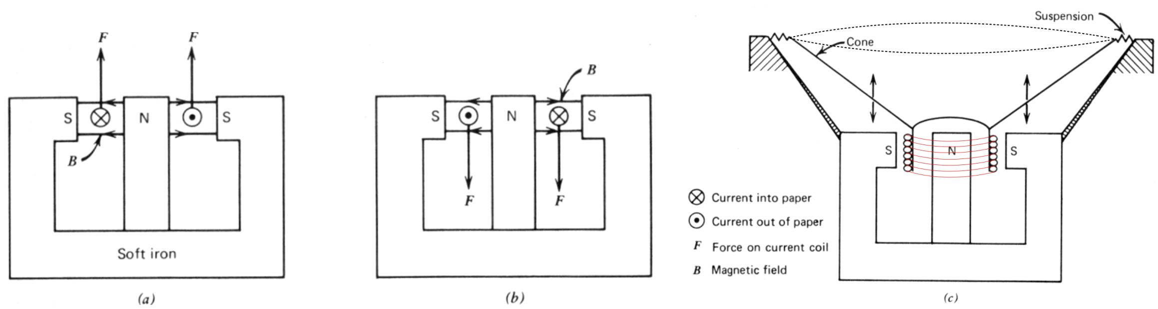


Figure 16.21: In (a) and (b), the magnetic field lines point from the North pole to the South poles, and the sense of current flow by the $\times$ or $\bullet$ as going "into the page" or "out of the page." The net force on the current in (a) is up, and in (b) is down. In (c), a cutaway view of the speaker cone attached to the coil. Force on the coil causes motion of the speaker cone.[1]

When current passes through the coil, a force from the magnetic field is applied to the wire coil that is perpendicular both to the magnetic field and to the current, so that the net force on the coil, and therefore the speaker cone, is up or down, shown by the force arrows. When current passes through the coil in one direction, a force is exerted on the coil tending to push the coil, and therefore the speaker cone, up, while for current in the opposite direction, the force pulls the coil and speaker cone down. Thus, speaker cone motion is dependent on the strength and direction of the current in the coil. Convince yourself of the force directions shown in the figure by applying the right-hand-rule to the directions of the coil current and the magnetic field. In part (c) the coil attached to the speaker cone translates this current-magnet force into motion of the cone, the direction of which depend on the current flow in the coil. The higher the current, the higher the speed of the cone and hence the louder the sound. A sound with a complicated waveform corresponds to a current that changes in a similarly complicated fashion. The motion of the speaker cone will therefore be complex as well, producing a soundwave in the air with the same complex waveform as the one that was originally recorded.

A speaker that is isolated produces very little sound (see figure 16.22). The reason for this is that, at the edges of the cone, the forward movement of the cone pushes air forward (a) at front of the edge, and pulls air forward at the rear of the edge of the cone. This allows for the establishment of air currents shown at the edges of the cone. When the cone moves backwards (b), the opposite current of air can be established around the edge of the cone. The net effect of this is that sound energy is lost to air flow. When the speaker is located in an enclosure (c) the air flow is prohibited, and more power can be established in sound waves that propagate into the room.

An *air suspension* enclosure allows for no air movement from inside the sealed speaker cabinet. This cushion of air behind the speaker cone can act as a sort of "spring," enabling it to have greater efficiency in its power output. Because speaker enclosures have resonant properties of their own (recall our discussion of sound resonance in a small enclosure where dimensions along each of the three axes give rise to specific harmonic families of resonant frequencies), they are often filled with absorbing material to inhibit the establishment of resonant vibration – see figure 16.22.

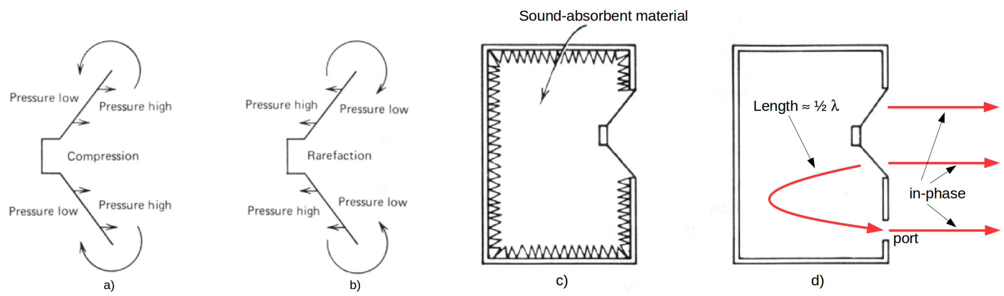


Figure 16.22: An isolated speaker cone cannot produce much power output since near the edges of the cone air can circulate around the edge of the cone in such a way as to partially cancel the air vibrations. Speakers located in an enclosure have no such cancellation and can operate with much greater efficiency and higher power output. [1]

A *bass reflex* speaker employs a small port or opening in one of the speaker walls, usually in the front wall. For this bass port to assist with sound production at low frequencies, it must be located an optimal distance away from the back of the speaker cone. Imagine that a speaker cone is producing a low bass tone of frequency f with wavelength λ. When the speaker cone is in the process of moving forward, its front surface pushes air into the room and its rear surface pulls air from inside the cabinet. A half-cycle later the front of the cone is pulling air at its front surface and pushing air at its rear. If the bass port is located approximately one-half wavelength away from the rear of the speaker cone (see the path of air vibrations in the figure *d*)), sound vibrations emerging from the port, displaced as they are by one-half cycle, are *in phase* with the sound at the cones front surface. Thus the presence of the port can enable vibrations at the cone's rear surface to be added (via the bass port) to those emerging from the front of the cone in order to support and strengthen the sound for low frequency tones. Because the distance from the rear of the cone to the bass port (red curve in figure *d*)) is designed to be approximately $\frac{1}{2}\lambda$, the port is often referred to as a "tuned" bass port. The port has little influence on tones much higher than the bass tones to which it is tuned.

Large cones intended for bass tones are called "woofers," medium cones well-suited for mid-range frequencies are called mid-range, and small cones for high frequency are called "tweeters". Horns are sometimes employed on tweeters in order to assist with the directional spreading of high frequency of sound throughout the room. This is not necessary for low frequencies since the wavelengths are more on the order of the room size and spread out naturally.

16.7 Chapter Summary

Key Notes

- A transducer is a device that changes a signal from one type to another. The most important transducers in musical recording and playback are electromagnetic, changing motion into an electrical signal or *vice versa*.
- Electric currents produce magnetic fields surrounding the wire.
- A magnetic field can exert a force on a current-carrying wire causing it to move.
- A moving magnet can induce currents in a wire. Currents can also be induced when the wire moves relative to the magnet.

- The phonograph cartridge is a transducer that changes the moving needle (as it rides in the groove of an LP record) into electrical signals that can be amplified and sent to speakers.
- The shape of record grooves are a reflection of the waveform of the music recorded onto them. Stereo records are encoded by groove shapes that produce left-right and up-down movement of the needle.
- Bass equalization is the process by which bass tones are reduced in amplitude before recording onto a vinyl disc, and re-amplified when played back from the disc. The purpose is to be able to record more music onto the disc since narrower tracks can used.
- Pre-emphasis is the process by which high tones are amplified before recording onto a vinyl disc, and then reduced in amplitude when played back. The purpose is to help reduce high frequency noise coming from imperfections in the record grooves.
- Tape recording and playback technology involve the storage and retrieval of patterns of magnetization induced in an iron oxide layer of a thin plastic tape. Alignment of the iron particles is accomplished by running the tape near to a magnetic head whose magnetization changes with the waveform of the music being recorded. Playback involves allowing the tape magnetization to induce changing magnetic fields that are turned into electrical currents and amplified before being sent to speakers.
- Dolby reduction involves enhancement of low-level sound in the recording process, and corresponding reduction of this low-level sound in the playback process.
- DBX technology involves "compressing" sound by reducing the volume of loud passages and enhancing soft passages of music in the recording process, and applying the opposite during playback. The purpose is to store music of a large dynamic range onto a medium with less dynamic range capability, such as the cassette tape.
- Radio transmission and reception comes in two varieties, FM (frequency modulation) and AM (amplitude modulation). The waveform of the music being broadcast is encoded in the amplitude of the carrier wave for AM, and in the frequency of the carrier wave for FM.
- High-fidelity playback of music involves a suite of electronic components including pre-amplifiers, amplifiers, and speakers.
- The operation of the speaker is based on the transduction of electrical signals into movement of the speaker cone.

Exercises

Questions

1) How is information conveyed by an FM broadcasting station?

2) How is information stored on a magnetic tape? How is it retrieved?

3) A stereo record has information stored upon it. Briefly, how is that information retrieved by a stereo sound system?

4) Briefly describe what a transducer is, and how it functions in the recording of music.

5) Describe briefly what adjustments need be

made to the low and high frequencies during the recording of a record album, and why.

6) Explain briefly how electrical sound is changed into the motion of a speaker cone in order to produce sound.

7) Explain briefly how electrical signals corresponding to sound are transferred to a cassette tape.

8) What is the purpose of a bass port on a loudspeaker?

Problems

1. Distinguish between *turnover frequency* and *transition frequency.*

2. What are the differences between MM and MC cartridges?

3. Suppose there is a slight warp that goes across the diameter of a record disk. The tone arm will receive two equally spaced vertical bounces with each revolution of the turntable. What is the frequency of this disturbance?

4. What are pre-emphasis and bass equalization, and why are they needed in LP recording?

5. For a phonograph record groove of radius 10 cm, on a disc rotating at $33\frac{1}{3}$ rotations per minute, what is the spacing of successive wiggles along the 62.8 cm circumference for signals of frequency a) 100 Hz, and b) 10 kHz?

6. When a phonograph record is created from a master disk, suppose the tracks end up being created off-center by 1 mm. When the stylus is tracking a groove of radius 10 cm, how much pitch variation results, and how often does it recur? (Recall that a 6% frequency change corresponds to a pitch difference of one semitone.)

7. The speed with which a cassette tape moves past the playback head is $1\frac{7}{8}$ ips (inches per second). If one side of the tape is to play 45 minutes of music, how long does the tape need to be?

8. Suppose a tape recording is made with the tape moving at 4.75 cm/s ($1\frac{7}{8}$ ips), and played back at 19cm/s ($7\frac{1}{2}$ ips). What happens to the a) pitch, and b) the tempo of the music?

9. Consider an FM station broadcasting at 88.7 MHz. When the carrier wave passes by an antenna, the charges in the antenna begin to oscillate at the same frequency as the wave.

   a) What is the period of oscillation for the charges?

   b) What is the wavelength of the transmission? (The speed of light $c = 2.998 \times 10^8$ m/s.)

10. At what frequency does the wavelength of sound equal the diameter of a 15 in woofer? a 2 in tweeter?

References

[1] http://www.dreamstime.com/stock-images-lp-record-image8080234

[2] http://www.ramelectronics.net/template-resources/images/RAM/Static/turntable.jpg

[3] http://www.needlengroove.com/

[4] Encyclopedia Britannica

[5] http://www.visualphotos.com/image/1x6058619/stylus_in_groove_of_lp_record_colored_sem

[6] Rigden, Physics and the Sound of Music, Wiley (1985)

[7] Berg and Stork, The Physics of Sound, Second Edition, Prentice Hall, pg. 180 (1995)

References

Chapter 17

Musical Synthesis

All sound, as we've come to learn, no matter how complex, has a single waveform associated with it. When an entire Beethoven symphony is recorded, each microphone receives one long, continuous waveform, from beginning to end. You recognize, of course, that this waveform will be very complex, much more so than for a single instrument, since it consists of the superposition of all the instruments, each with its own family of harmonics, various pitches, percussion sounds, changing dynamics, tempo variations, *etc.* It's hard to imagine that all of that complex sound can be encoded into a single waveform, but it can! After all, when we sit in the audience and listen to the performance, each of our eardrums is vibrating in the complex way dictated by the wave form impinging upon it. The detailed motion is very complex, since the pattern of its movement is dictated by the single superposition curve of countless simultaneous sounds.

17.1 Synthesis of Waveforms

A theorem in a branch of mathematics called Fourier Analysis states that *any* complex waveform can be created from the superposition of an appropriate collection of simple sine waves with different frequencies and amplitudes. The specific frequencies and amplitudes needed to create a particular waveform can be calculated by a procedure that is beyond the scope of this course. But it is important to know that such calculations can be done, for any waveform, no matter how complex. The more complex the waveform, the larger the collection of sine waves needed to superimpose in order to produce the waveform.

The process of musical synthesis basically boils down to superimposing several sine waves of varying amplitudes and frequencies to produce "imitative" waveforms, or those intended to sound like physical instruments, or "new" waveforms with unique or unusual timbres not realized in physical instruments. Some like to use synthesis in order to create electronic versions of common instruments. The process of composing on a computer is made a lot easier if waveforms corresponding to real instruments can be drawn upon to make sound, and often these waveforms are the result of synthesis. On the other hand, others like to explore new sounds by creating instruments never before heard. In principle, since any complex waveform can be constructed out of the appropriate sine waves, *any* known sound can be synthesized. In both cases, the recipe is fairly straightforward – one starts with a desired waveform, and calculates the recipe for which specific frequency and amplitude sine waves need to be added, and the other creates new and interesting sounds by experimentally adding sinusoidal components of various frequencies and amplitudes together.

So how are waveforms produced? How do we decide which waves need to be added in order to produce a desired waveform? Some common waveforms produced by synthesizers, some of which are also used in electronic music, are depicted in figure 17.1. Each of these can be produced with the proper selection of different amplitude and frequency sine waves.

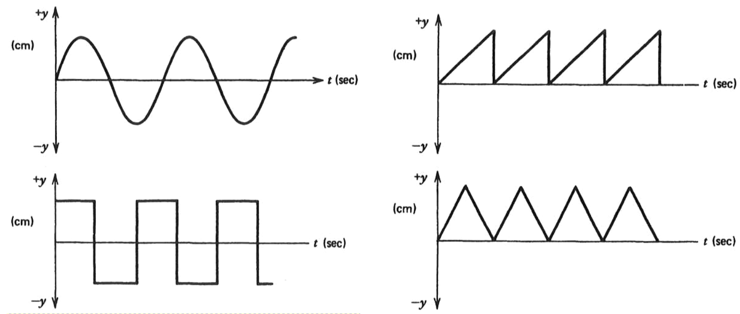


Figure 17.1: Some common waveforms produced by synthesizers and used in music – the pure tone, the sawtooth wave, the square wave, and the triangle wave. [1]

17.1.1 Example – the Square Wave

Consider a simple square wave – a waveform that is often used in electronic music. The mathematical calculation needed to compute the frequencies and amplitudes of the sine waves required to produce a square wave by superposition is beyond the scope of this course. The sequence of sine waves necessary to synthesize the square wave is shown in equation 17.1,

$$\text{Square wave } = \frac{4}{\pi}\left[\sin(\omega t) + \frac{\sin(3\omega t)}{3} + \frac{\sin(5\omega t)}{5} + \frac{\sin(7\omega t)}{7} + \ldots\right] \tag{17.1}$$

where t is time, f is the frequency of the wave, and $\omega = 2\pi f$. It consists of an infinite sum of sine waves of different amplitudes and frequencies. The sound spectrum of the first five components is depicted in figure 17.2, and the superposition curves for the first 1, 3, 5, and 12 components of this series are shown in figure 17.3.

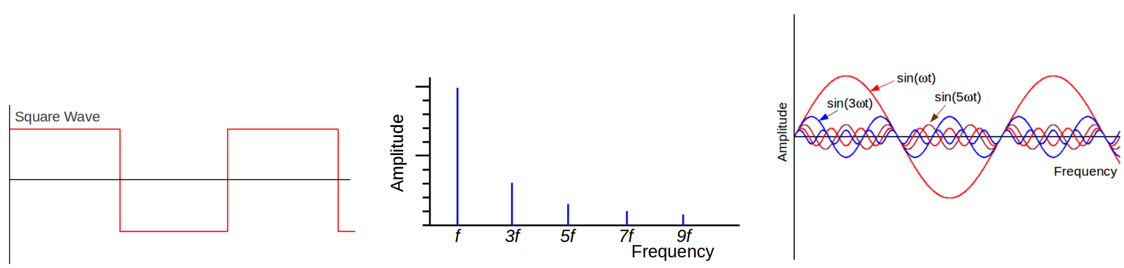


Figure 17.2: (***For color version see Appendix E***) The ideal square wave, the first five components of its frequency spectrum, and the sinusoidal waves corresponding to these five components, of differing frequency and amplitudes.

As seen in equation 17.1, the calculation results in an infinite series of sine waves of differing amplitudes and frequencies. Even so, we don't require a particularly large number of the terms in this series to see the shape of the square wave taking form. The first pane of figure 17.3 shows the very first term in the mathematical sequence of sine waves. Note that it does not particularly resemble the square wave, nor should it, being only the first term of an infinite series!

But at least we can see that the first term is high where the square wave is high, and low where the square wave is low. The second pane shows the result of superimposing the first three terms, and already we can see the square wave taking shape. The next pane of 5 terms looks even closer, and

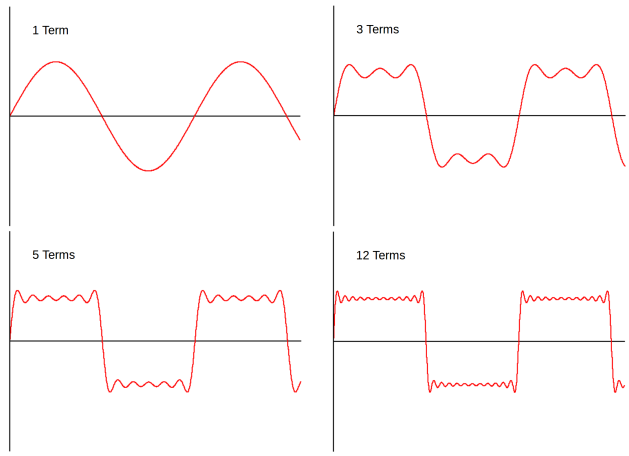


Figure 17.3: Synthesis of a square wave using sine waves, ranging from 1 sine wave to 12. Note the progression toward what begins to look distinctly like a square wave. The more terms in the superposition, the better the approximation.

the fourth pane with 12 terms is already looking very much like a square wave, albeit with some extra wiggles. The wiggles would continue to disappear as the number of superimposed waves becomes larger.

17.2 Electronic Synthesis

A physical synthesizer contains several internal oscillators, each of which can produce a sine wave with any amplitude and frequency. A 12-oscillator bank synthesizer has 12 internal oscillators, and would be able to produce the square waveform shown in the fourth pane of figure 17.3. Synthesizers can have a much larger number of oscillators, meaning that they can produce a wider variety of waveforms with great precision and fidelity. Oscillators are typically capable of producing waveforms relevant to the production of music, with frequencies ranging from the subsonic region (fractions of a Hertz) up to 10,000 Hz or more.

17.2.1 Electronic Frequency Filters

In addition to oscillators, synthesizers also employ a variety of *filters*. Filters operate on a basis similar to what we found for the formants of the human voice. When they are applied to a set of frequencies coming from a bank of oscillators, they can emphasize certain frequencies and de-emphasize others. Different sound spectrum components can be shaped in such a way that some can be reduced or eliminated while others are passed at full amplitude, ultimately for the purpose of shaping timbre. There are several types of filters used in a synthesizer, the most common of which include the low-pass, high-pass, band-pass, and band-reject filters (see figures 17.4 and 17.5).

The idea behind the low-pass filter is that it "passes," or accepts frequencies below a certain cutoff frequency, and rejects those above it. The high-pass filter does the opposite – it passes frequencies above the cutoff frequency and rejects those below it. The location of the "cutoff frequency," which determines which frequencies will be passed or rejected, is adjustable for both these filters.

Band-pass filters are created by passing the output of a low-pass into a high-pass filter. The result of

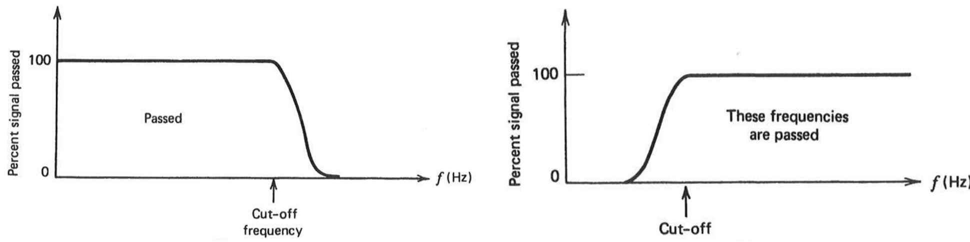


Figure 17.4: A low-pass and a high-pass filter, used to shape the sound from a bank of oscillators. [1]

both filters in sequence is that they pass all frequencies within the frequency range that the low-pass and high-pass filters have in common. Those that fall below the cutoff frequency of the low-pass filter (*i.e.* are *passed* as a result) and also fall above the cutoff for the high-pass filter will be passed on to the system. All others falling outside this range of frequencies will be rejected. Note that for a band-pass filter, the cutoff frequency for the high-pass filter needs to be *below* the cutoff for the low-pass filter.

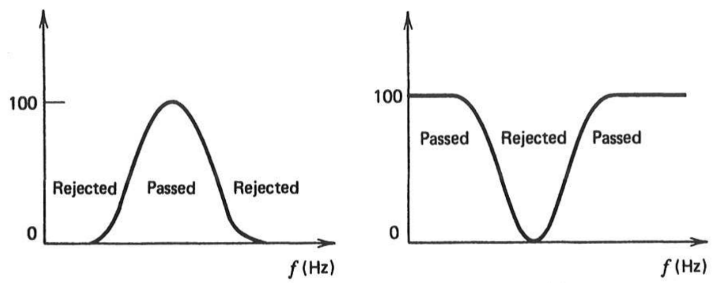


Figure 17.5: A band-pass filter (left) and a band-reject filter (right), used to shape the sound from a bank of oscillators.[1]

If the low-pass and high-pass filters have no frequency in common, no frequencies will pass the combination, as long as the output of the first is directed to the input of the second. However, if the sound from the oscillator bank is split into two, one half to be sent through the low-pass filter and the other half through the high-pass filter, and if these two outputs are then recombined, we have a band-reject filter, which will pass all frequencies below the low-pass cutoff and all frequencies above the high-pass cutoff. All frequencies in between these two are rejected.

Look again at figure 13.4 in the Human Voice chapter and see how the human vocal tract serves as a type of complicated band-pass filter. It emphasizes the support of frequencies that fall within the formants (which are somewhat similar to band-pass filters) and de-emphasizes other frequencies outside these formant regions (which is a little like rejecting them). The purpose of this vocal tract shaping is to form vowels, which correspond to distinct timbres of the voice. Likewise, band-pass and band-reject filters shape the output sound by emphasizing certain frequencies coming from the oscillators and rejecting others, with the ultimate purpose of forming distinct timbres of the sound.

17.2.2 Electronic Amplitude Filters

Synthesizers are also able to shape the *amplitude* of the signal with the use of filters, allowing the output volume of sound to be shaped as a function of time. An *envelope generator* consists of a sequence of amplitude filters that operate for short intervals of time, in order to shape the attack, decay, sustain, and release characteristics of a tone. Figure 17.6 shows, schematically, the operation of the more simple "AR" (attack-release) filter, which shapes the beginning attack and the ending release of the sound. The more complex ADSR filter adds shaping ability to the decay and sustain characteristics of the sound. The attack of an instrument corresponds to how quickly the sound is established. You'll recall for the piano, a percussive instrument based on the sound produced by a hammer hitting the string, the attack was very distinct and sharp (see figure 12.6). On the other hand, the attack of a flute is gentle by comparison, the amplitude building up over a longer period of time before the full note is established. The decay registers how the sound dies out initially before sustaining a steady sound. The sustain governs how long and how loud the note stays at volume before being released, and the release corresponds to how the sound dies out when the note is ended.

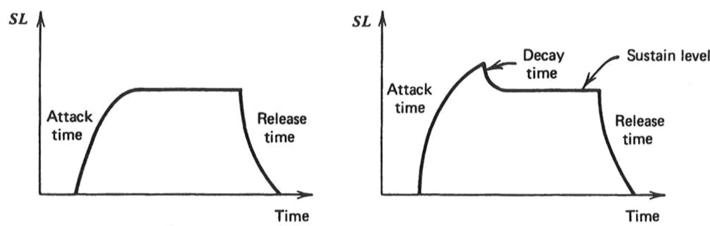


Figure 17.6: Schematic representation of the AR (left) and ADSR filter (right) as a function of time. [1]

17.2.3 Additive and Subtractive Synthesis

With a sufficiently large bank of oscillators, each able to produce a pure tone with any frequency, we could, in principle, synthesize any waveform by simply adjusting the relative amplitudes and phases of the oscillator outputs. Adding together the outputs of several oscillators in this fashion is called *additive synthesis*. In order to produce rich and distinct timbres, a large number of oscillators is necessary. On the other hand, if we were to begin with a complex tone that already consists of a large number of harmonics with varying frequencies and amplitudes, we could apply filters to selectively remove certain harmonics. Consider the square wave of section 17.1.1. If the fundamental of this tone were 500 Hz, and we were to send this tone through a low pass filter with a cutoff of 10,000 Hz, the first 20 harmonics of the tone would be passed and those above 10,000 Hz would be rejected. If we were to lower the cutoff frequency of this low-pass filter, higher harmonics would be successively removed until the cutoff reaches around 600 Hz, at which point only the fundamental remains, constituting a pure tone. This application of filters, with the intent of removing harmonics from an already established tone to change the timbre, is called *subtractive synthesis*. With both of these techniques, new waveforms and therefore new sounds and timbres can be created.

17.3 Chapter Summary

Key Notes

- An oscillator is an electronic device that generates a periodic signal. Several waveforms can be generated by oscillators. In principle, any waveform shape can be generated electronically, and thus a wide variety of new sounds (i.e. timbres) can be generated by electronic means, enabling the creation of "new instruments."
- **Filters**: With electronic filters, specific harmonic information can be reduced or eliminated, thus enabling the changing of the timbre of sounds. There are several types of filters used in a synthesizer, including low-pass, high-pass, band-pass and band-reject filters. The location

where the filter cuts off, passes, or rejects frequencies can be adjusted to achieve whatever musical objective is desired.

- Amplifiers enable control to be exercised over the output amplitude. Envelope generators control the attack, decay, sustain, and release characteristics of a tone. These envelope generators can take advantage of the obvious differences in the way different instruments achieve their characteristic sounds. A flute, for example, is characterized by an attack during which the tone gradually builds up, whereas the guitar has a very sudden percussive attack. Thus if a synthesizer is to faithfully produce sounds similar to real instruments, it needs to shape the tone appropriately.
- **Additive and subtractive synthesis**: With this technique, new waveforms can be created merely by adjusting the relative amplitudes and phases of appropriate frequencies (additive synthesis) to produce new sounds. Alternatively, one can start with a waveform that already exists and low-pass filters can be used to subtract out higher harmonics, for example, to change the timbre of the tone.
- **Sample Patches**: When components are put together in a specific fashion to form a new musical "instrument," it requires an oscillator, which creates the basic waveform, connected to a filter that adjusts its shape. This combination is then called a patch. Setting up a patch means connecting components together, which provides conducting paths for the electrical signals that eventually are transformed into an audio output signal.

Exercises

Questions

1) Explain the basic idea behind synthesis of a musical waveform.
2) Does a synthesizer with a larger number of oscillators produce waveforms with better fidelity, or does it matter? Explain briefly.
3) Describe the function of a low-pass filter. What is the significance of its cutoff frequency?
4) Briefly, how is a band-pass filter constructed?
5) What is the difference between frequency filters and amplitude filters? For what application is each intended?
6) Differentiate between additive and subtractive synthesis.
7) Can any waveform be synthesized? Explain briefly.

Problems

1. Draw the superposition curve of a square wave and a simple harmonic (sinusoidal) wave with half the amplitude and twice the frequency of the square wave. Draw the curve for about 2 periods of the square wave.
2. Square waves of 200 and 301 Hz are sounded together. Identify the first three harmonics above the fundamental for each tone and indicate which harmonics will form beats with the others. How many beats will be heard?

3. A bedside radio is designed to tune in one frequency out of many frequencies that are incident on the antenna. What kind of filter is used to tune in this frequency, and why?

4. A broadband of frequencies form the input to a low-pass filter. The output of the low-pass filter then leads to a high-pass filter. If the low-pass filter has a cutoff frequency of 2000 Hz and the high-pass filter has a cutoff frequency of 1800 Hz, what is the output of the combination? Draw a figure showing the low-pass and high-pass filters and the result of their combination.

5. Design a band-pass filter to pass frequencies in the range of 1200 to 1500 Hz. To design this filter you have a low-pass and a high-pass filter whose cutoff frequencies are variable.

6. Make a drawing (to scale) of an ADSR envelope with a maximum amplitude of 2V and the following parameters: $A = 50$ ms, $D = 20$ ms, $S = 1.5$ V, $R = 100$ ms. (Note that the drawing in figure 17.6 plots the sound level on the y-axis, whereas you will be plotting the voltage signal on the y axis. Otherwise the plots will be very similar.) Please remember to label and include values on your axes.

References

[1] John S. Rigden, Physics and the Sound of Music, Second Ed., Wiley (1985)

References

Chapter 18

Digital Music

18.1 Analog *vs.* Digital

The musical technologies we've considered so far have been *analog* in character. An analog quantity, such as the amplitude of a musical waveform, is one that varies *continuously* in time. It is characterized by a curve that increases or decreases *smoothly* as it evolves. In contrast, a digital quantity is one that increases or decreases *discretely*, in steps that are characterized by integral numerical values. The small rise in elevation just outside the library is connected by a stairway, and the individual steps represent discrete changes in elevation (digital), whereas the smooth ramp alongside the stairs negotiates the same change in elevation continuously (analog). Hence, the three basic technologies we studied in chapter 16, the LP record, the magnetic tape, and AM and FM transmission, are analog technologies since the means by which they record the analog musical waveform is by a smoothly varying quantity.

The groove variations in the LP record change smoothly as the wall shape pushes the needle to varying degrees left and right, up and down. When the motion of the needle causes the magnet inside the cartridge to induce currents in the pickup coils, the resulting value of the current is likewise a smoothly varying quantity, following the variations of the original recorded waveform. The same notion applies to the analog magnetization variations on the recorded tape or in the modulation of the AM or FM radio signals – the recorded waveform appears as a continuously changing quantity of magnetization in the thin iron oxide layer, or of the amplitude and frequency variations of the carrier wave of the AM and FM signals.

The value of an analog sound technology is in its ability to store and retrieve the small and continuous variations of the musical waveform with consistent accuracy. To the extent that the groove variations in the record are an accurate reflection of the detailed waveform of the original musical source, the LP has high fidelity. To the extent that the amplified signal from the induced currents in the LP cartridge produce motion in the speaker cone that creates a pattern of propagating pressure variations in the air identical to those of the original recorded source, the system has high fidelity. But analog technologies also have a serious drawback. The high fidelity of the recorded waveform decreases over time as the device is used.

18.1.1 Analog and High Fidelity

Consider the LP record. As the record is played, the needle rides in the groove, guided by the bumps and wiggles of the walls. As it gets pushed this way and that by the groove variations, the peaks of the sharp bumps get slightly worn down as they slide by the hard needle. Likewise, as the needle is pushed up and falls back down with the narrowing and widening of the groove, the weight of the needle and tonearm combination serves to wear down the fine details of the walls and floor of the groove. The next time the record is played, the sharpness of the groove peaks are somewhat rounded, and the subtle shape variations of the walls are changed. As a result, the fidelity of the recording is reduced.

The fidelity continues to diminish with each playing since the shape of the groove can only *depart* from the original waveform, and not magically gain back any of its original integrity. Both record and tape recordings suffer a reduction in fidelity over time, accompanied by an increase in distortion and white noise in the background.

As another example, consider AM transmission and reception. In chapter 17 we pointed out that since the waveform is encoded in the amplitude of the carrier wave, the signal can be diminished in integrity by reflections and absorptions in the surrounding countryside. As the amplitude variations are affected by the propagation of the signal from the transmission tower to the antenna, the fidelity of the original signal is diminished. Both AM and FM transmission and reception are analog technologies. The method of signal encoding used by FM, however, is superior to that of AM for the ability to maintain fidelity. As a FM signal propagates from the tower to the antenna, any reflections and absorptions that occur will affect the amplitude of the carrier wave, much as in the case of AM radio. However, the encoded waveform in FM radio is "hidden" from the degradation of the signal, since it is encoded in the frequency modulation, which is unaffected by changes in amplitude. The only effect that amplitude variations will have on the FM signal is in its overall strength, which will translate into small variations in loudness. But radios can partly compensate for this by adjusting the volume output as it monitors variations in signal strength.

18.1.2 Digital Signal and Fidelity

Digital technologies are superior to analog technologies in their ability to maintain fidelity over time, for reasons that are somewhat analogous to the superiority of FM over AM. A digital signal stores the original waveform information in a way that is "hidden" from the inevitable degradations of signal integrity that occur as the signal is sent through a long cable or over an optical fiber, or through circuits and manipulated by electronic devices. The detailed shape of the waveform is encoded in *bits* of information that are either "on" or "off" – "1's" or "0's." The *information* of a sound signal is encoded in a pattern of 1's and 0's, and not in the size of the signal amplitude or in its detailed shape. When a digital pattern suffers moderate degradation in its signal shape or amplitude, the playback device can still retrieve a faithful record of the original waveform variations which is encoded digitally. This is because even when a signal becomes degraded in shape, it is still possible to tell the difference between the peaks (1's) and valleys (0's) – see figure 18.1.

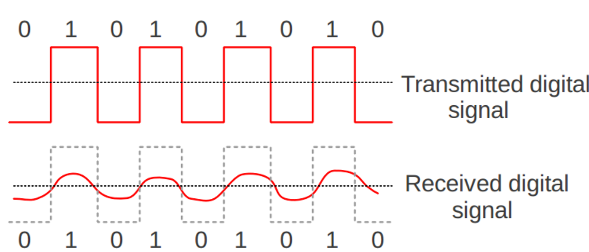


Figure 18.1: (***For color version see Appendix E***) A transmitted and received digital signal. The original 1's and 0's of the transmitted signal are greatly degraded in shape by the time they reach the receiver, but since the encoding of the waveform is in the peaks and valleys of the signal, the receiver can still interpret the difference between the 1's and 0's above and below the dotted line, even though the signal is greatly rounded and reduced in amplitude.

18.2 Digital Sound

18.2.1 Digital Encoding of an Analog Waveform

The basic idea behind digital encoding is to translate the continuously varying musical waveform (sometimes called a sound curve) into a series of digital numbers that can be recorded onto a storage medium, such as a CD, a hard-drive, etc. Figure 18.2 depicts schematically how this is done. The "sampling rate" of CD technology is 44,100 Hz, meaning that the recording circuitry measures the waveform amplitude every $1/44{,}100^{th}$ of a second and turns its value into a digital number. The analog waveform is in this way translated into a series of integer numbers.

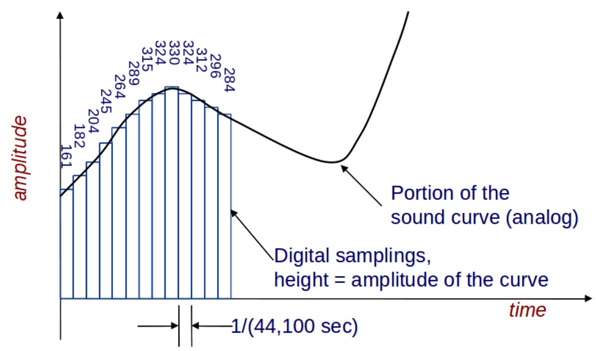


Figure 18.2: (***For color version see Appendix E***) Digital encoding of an analog waveform. Each $44{,}100^{th}$ of a second, the waveform amplitude is converted into a number. The stream of numbers then represents an approximation of the waveform.

18.2.2 Computers and Information

The numbers appearing above the sound curve in figure 18.2 are shown in decimal representation for illustrative purposes, but in order for them to be read and processed by a computer, they need to be converted to *binary* format. Computers and electronic devices (such as CD players) can only distinguish between two types of signals – "on" and "off." Any numerical information corresponding to the variations in the waveform need to be converted to binary format, whose only digits are 1 and 0.

18.2.3 Decimal and Binary Numerical Representation

A multi-digit number expressed in the decimal representation (also called "base 10") is to be interpreted in the following way. Each digit, depending on its position in the numerical representation, is to be multiplied by a particular power of 10, after which they are all added together. The last digit on the right is to be multiplied by $10^0 = 1$, and added to the next digit multiplied by $10^1 = 10$, and added to the next digit multiplied by $10^2 = 100$, etc. For example, the number 432 (processed from the lowest to the highest digit) is

$$432 = \left(2 \times 10^0\right) + \left(3 \times 10^1\right) + \left(4 \times 10^2\right) = 2 + 30 + 400 = 432.$$

Each digit can be filled with one of 10 symbols, which are the familiar 0-9.

Similarly, in binary format (also referred to as base 2), each digit corresponding to a certain position in the number is to be multiplied by a particular power of 2, after which they are added. For example, the same number 432 expressed in binary is 110110000, interpreted as follows (again starting from the digit on the right):

$$110110000 = \left(0 \times 2^0\right) + \left(0 \times 2^1\right) + \left(0 \times 2^2\right) + \left(0 \times 2^3\right) + \left(1 \times 2^4\right) + \left(1 \times 2^5\right) + \left(0 \times 2^6\right)$$

$$+ \left(1 \times 2^7\right) + \left(1 \times 2^8\right) = \text{ (in decimal) } 0 + 0 + 0 + 0 + 16 + 32 + 128 + 256 = 432.$$

Each digit can only be represented by one of two symbols, 0-1, since it is base 2, and therefore numbers expressed in binary format have many more digits than the same numbers in decimal format. Table 18.1 lists the numbers from 1 to 20 in binary format. These numbers can all be expressed with 5 bits (digits) or less.

The total number of bits used to express a number determines the "numerical range" of that representation. If we are limited to using three digits for a number in decimal, what is the numerical range of this 3-digit representation, *i.e* how many numbers can be represented using only three digits? A three-digit decimal number is able to represent 1000 different numbers raging from 0 to 999. The numerical range of a decimal number consisting of n digits is 10^n. In the case of a 10 digit decimal number, the numerical range is $10^{10} = 10$ billion.

Table 18.1: The decimal numbers 1 to 20, expressed in binary format

Decimal Number	Five-Bit Binary Number	Decimal Number	Five-Bit Binary Number
1	00001	11	01011
2	00010	12	01100
3	00011	13	01101
4	00100	14	01110
5	00101	15	01111
6	00110	16	10000
7	00111	17	10001
8	01000	18	10010
9	01001	19	10011
10	01010	20	10100

The same basic reasoning can be applied to the binary counting system. A binary number with 3 digits can represent $2^3 = 8$ different numbers, able to express numbers between 0 (000) to 7 (111). With 5 bits we can represent $2^5 = 32$ separate numbers, between 0 (00000) to 31 (11111), and with 16 bits (the number used in CD recordings) we can represent 65536 separate numbers from 0 (0000000000000000) to 65535 (1111111111111111).

18.2.4 Sound Level Dynamic Range

As a musical waveform is encoded digitally, the generated stream of binary numbers reflects the variations of the waveform amplitude in time. An important point to note is that the intensity of a sound is proportional to the *square* of the waveform amplitude. This means that in order to understand the *dynamic range* of sound intensity able to be represented by a numerical system of n binary digits, we need to *square* the numerical range. Since the numerical range of a binary number containing n digits is 2^n, then the dynamic range of sound intensity represented by this number is the square of its numerical range, or $(2^n)^2 = 2^{2n}$, and the sound intensity level, defined as $10 \log\left(\frac{I}{I_0}\right)$ will be $10 \log\left(2^{2n}\right)$.

Example 18.1

Dynamic Range of 8-Bit Sound *What is the dynamic range of sound represented by an 8-bit binary system?*

Solution: The numerical range of an 8-bit number is $2^8 = 256$, meaning that 256 different numbers are able to be expressed using this 8-digit binary number. The range of sound intensities that can be represented is the square of its numerical range, or $256^2 = 65,536$. This means that the ratio in intensity of the quietest to the loudest sound for an 8-bit system is 1:65,536. The range of sound level for this 8-bit system is therefore $10 \log(65,536) = 48$ db.

18.2.5 Digital Disk: The CD

We are now in a position to understand how information is stored on a Compact Disc (CD). The CD consists of a thin circular metal layer encased in clear plastic. Tiny pits are etched with a laser into the surface of this metal layer, forming a circular spiral starting at the inner radius of the metal layer and running to the outer. Note that this layout is very similar to that of the LP, with two exceptions. The data on the CD is read starting from the inner radius and moving out to the outer, and the CD rotates at a variable rate depending on which radius the current track is being read. The CD rotates fastest when reading the inner radius and slowest when reading the outer. Figure 18.3 shows some of the basic dimensions of the CD.

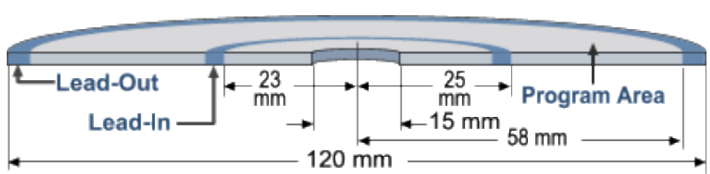


Figure 18.3: Cross section of a compact disc showing some of the dimensions and locations of the lead-in an lead-out radii.[1]

The pits etched into the metal layer encode the binary information corresponding to the digitized musical waveform curve. Figure 18.4 shows the basic dimension of a CD pit. Individual pits are very small, having a depth of around $0.2\mu m = 0.2 \times 10^{-6}$ m, or 0.0000002 m and a width of around $0.5\mu m$. The width of the tracks in which the pits are placed is $1.6\mu m$. Suppose that a smooth portion of the metal layer represents a "1" and a pit represents a "0." We then see how the pattern can represent binary numbers. For the purpose of simplicity, imagine that a CD is encoded with 5-bit words of information. If a particular portion of the waveform has amplitude variation that corresponds to the sequence 13, 14, 13, 11, 9, 8, then the pattern of pits would look as illustrated in figure 18.5.

The CD operates on a 16-bit system, meaning that one grouping of digits corresponding to one amplitude on the sound waveform consists of 16 separate bits. A 16-bit binary number can represent the

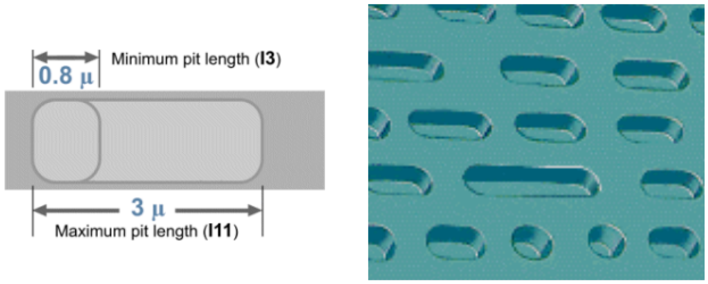


Figure 18.4: Schematic showing the basic shape of the CD pits and their dimensions.[1]

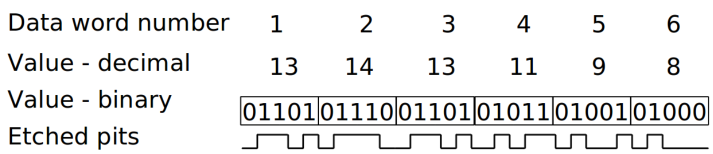


Figure 18.5: A representative section of CD bits for a 5-bit system. Note that real CD systems use 16-bit words for a sound level dynamic range of 96 db.

numbers from 0 to 65,535, which can therefore represent this many variations in the sound curve amplitude. As mentioned earlier, the dynamic range of intensity is proportional to the square of the numerical range. Hence, the dynamic range able to be represented by a CD is $\Delta SL = 10 \log \left[(2^{16})^2\right] = 96$ db. This is a significant portion of the range of human hearing, and represents the appropriate dynamic range for musical performance with high fidelity.

When there are two "0's" in a row, the pit is twice as long as for one. Similarly, when there are three "1's" in a row, the smooth surface is longer then when there are two, *etc.* The pits are quite small – if the smallest pit on the CD were enlarged to be 1 inch in length, the diameter of the CD would be over 2 miles.

The series of pits is read by a laser that is focused on the track as it rotates. When a smooth portion of the metal layer passes under the scanning laser, the light is reflected back to a sensor and read as a "1." When a pit passes underneath the laser, the light is transmitted through the layer and scattered, not to return to the sensor and read as a "0." The distance between bits is constant throughout the entire CD. Since the inner radius of the CD has a smaller circumference than the outer, this means that the inner tracks of the CD contain less information per rotation than the outer tracks. Therefore in order for the information rate (bits/second) retrieved from the disk to be constant, the CD spins faster when the inner track is being read, slows down gently as the reading proceeds along the spiral to outer tracks. The scanning velocity is between 1.2 and 1.4 m/s. This means that the rotational speed of the disk must vary from a high of 500 rpm (when the inner edge is being read) to a low of 200 rpm (at the outer edge). The series of binary numbers read in this fashion of 1's and 0's is then sent to a circuit that translates the digital signal into an analog signal that can be amplified and sent to the speakers.

18.3 Chapter Summary

Key Notes

- The basic idea of digital encoding is to change the continuous waveform of a piece of music, which is "analog" in nature, i.e. a continuously and smoothly changing amplitude (sometimes called a sound curve), into a series of digital numbers that can be recorded onto a storage medium, such as a CD or hard-drive.
- Computers can only distinguish two types of signals: "on" and "off." Therefore, information that is encoded needs to be in binary format, where the only digits are 1 and 0.
- In the decimal system (i.e. base 10), each digit contains a number that is to be multiplied by a power of 10 and added. For example, the number 432 is $2 \times 100 + 3 \times 101 + 4 \times 102 = 2 + 30 + 400$. Each digit can be represented by one of 10 different symbols, 0 - 9. In binary format (i.e. base 2), each digit contains a number that is to be multiplied by a power of 2 and added. For example, the number $10110 = 0 \times 20 + 1 \times 21 + 1 \times 22 + 0 \times 23 + 1 \times 24 =$ (in decimal) $0 + 2 + 4 + 0 + 16 = 22$. Each digit can be represented by one of two symbols, 0 - 1. See table 17.2 for the binary representations of numbers from 1 to 20.
- The Compact Disc consists of a thin circular metal layer encased in clear plastic. Tiny pits are etched with a laser in a circular spiral starting at the inner radius of the metal layer and running to the outer. This layout is very similar to the analog groove that is cut into a LP record, except that in the record the groove starts at the outer radius and spirals toward the inner.
- The pits etched into the metal layer encode the binary information corresponding to the digitized music curve. The rate at which a sound curve for a single channel is digitized for storage onto a CD is 44,100 samples/sec.
- The frequency of this sampling is above the human frequency range of hearing.
- The CD operates on a 16-bit system, meaning that each number characterizing the amplitude of the sound waveform when it is sampled every $1/44{,}100^{th}$ of a second is expressed as a 16-bit binary number.
- A 16-bit binary number can represent the numbers from 0 to 65,535, which can therefore represent this many variations in the sound curve amplitude. Since the intensity of the sound wave is represented by the square of the amplitude of the sound curve, this means that the CD can produce a sound level difference between its lowest and highest possible level output of 10log(65,5362) = 96 dB.
- Since the inner radius of the CD has a smaller circumference than the outer, and since the distance separation between bits is constant, this means that the inner tracks of the CD contain less information per rotation than the outer tracks. Therefore in order for the information rate (bits/second) retrieved from the disk to be constant, the CD spins faster when the inner track is being read, slows down gently as the reading proceeds along the spiral to outer tracks.

Exercises

Questions

1) What are some advantages of digital over analog technologies?

2) What are some advantages of analog over digital technologies?

3) Briefly, how is a waveform converted to digital format?

4) Why is binary the base of choice for digital technologies?

5) How is the intensity of a signal related to its waveform amplitude?

6) What is the basic idea behind the binary counting system – that is, how do we interpret the digits of a binary number?

7) Why does the CD rotate faster when the inner tracks are read compared with when the outer tracks are read?

8) Why are CD's encoded with words of 16-bit length?

9) How does a laser retrieve the information from a CD?

10) Briefly, how is information encoded on CDs?

Problems

1. What is the range of decimal numbers that can be represented by a 6-bit binary system?

2. Write the binary representations of the following decimal numbers: a) 12 b) 22 c) 36.

3. Write the decimal representations of the following binary numbers: a) 1101001 b) 100110 c) 1010101.

4. Try to multiply binary numbers 101 times 110, using the same procedure that you use to multiply ordinary decimal numbers. Check your answer by multiplying the corresponding decimal numbers.

5. How many different wave amplitudes can be expressed with an 8-bit system?

   Solution: An 8-bit system can express a total of 256 values of amplitude, ranging from 00000000 to 11111111, which in decimal is from 0 to 255.

6. How many different levels can be represented by a 13-bit binary code? What is the possible dynamic range for this signal?

7. Why is it necessary for the rotational rate of the compact disk to be faster when the information is being read on the inner track than when it is being read on the outer track?

8. Consider CD-quality stereo sound requiring 44,100 numbers per second, with 16 bits for each number, for each of two channels. How many bits are required to specify a piece of music lasting 3 minutes?

References

[1] http://www.disctronics.co.uk/technology/cdaudio/

Appendix A

Units and Conversions

The following are a few of the common units we will be encountering in this course, along with their relationship to the standard MKS units of meters, kilograms, and seconds.

Length, MKS unit meter, m:

1 millimeter	= 1 mm	$= 10^{-6}$ m
1 centimeter	= 1 cm	$= 10^{-2}$ m
1 kilometer	= 1 km	$= 10^{3}$ m
1 inch	= 2.54 cm	= 25.4 mm
1 ft	= 30.48 cm	= 0.3048 m
1 mile	= 1 mi	= 1610 m

Mass, MKS unit kilogram, kg:

1 gram	= 1 g	$= 10^{-3}$ kg
1 milligram	= 1 mg	$= 10^{-6}$ kg

Time, MKS unit second, s:

1 nanosecond	= 1 ns	$= 10^{-9}$ s
1 microsecond	= 1 μs	$= 10^{-6}$ s
1 millisecond	= 1 ms	$= 10^{-3}$ s
1 minute	= 1 min	= 60 s
1 hour	= 1 h	= 3600 s

Force, MKS unit Newton , N $= \dfrac{\mathbf{kg \cdot m}}{\mathbf{s^2}}$:

1 pound	= 1 lb	= 4.448 N

Pressure, MKS Pascal, Pa = 1 N/m^2:

Air pressure	= 1 atm	$= 1.013 \times 10^5$ N/m^2
1 lb/ft^2	= 1 lb/ft^2	= 6891 Pa
Air pressure	= 1 atm	= 14.7 lbs/in^2

Energy, MKS unit Joule, J:

1 kilowatt·hour	= 1 kW·h	$= 3.6 \times 10^6$ J
1 calorie	= ca	= 4.186 J
1 kcal (food calorie)	= Ca	= 4186 J

Power, MKS unit Watt, W:

1 kilowatt	= 1kW	= 1000 W
1 horsepower	= 1 hp	= 746 W

Appendix B

Problem Solving Strategy

The purpose of the following is to offer some good suggestions to assist you in working on homework problems and for studying the course material. Of course, this approach is not going to be applicable to every sort of problem we will encounter, but it can serve as a helpful template from which to develop good problem solving techniques.

- Find a quiet, well-lit place. Bring lots of blank paper for scratch work, and a nice writing pen (I find nice pens inspirational).
- After reading the problem *all the way through*,
  1. Write down what you know ... all the variables that are relevant to the problem, both those you are given and those you don't yet know. Draw a picture of the situation. You can almost always draw *some* diagram or figure that will help you assemble your thinking and clarify things.
  2. Think about what core, underlying *physical concept or principle* is at the heart of the problem. This core principle can always be stated in words without the use of equations. It helps clarify the problem statement by removing some of the extra "trappings" from the problem.
  3. Look through the relevant chapter material and identify equations you will need for the problem. Write them down, putting a check mark above each variable for which you have a value. Identify the variable(s) you are trying to solve. If you know all the variables except for one, you need only one key equation to solve for it. If there are two or more unknowns, you will need as many equations as there are unknown values to solve for them. This can be a clue as to whether you still need an additional equation or two.
  4. Perform the calculations needed to arrive at values for the unknown(s). This will usually mean performing algebraic calculations or interpreting graphs.
- After you have worked out the problem, write a neat, concise version of it as your final product. You'll need to show your relevant work, put a box around your answer, and make sure to include units on all your physical quantities. A number corresponding to a physical quantity is meaningless without an appropriate unit!
- If you like doing so, I encourage working on homework with friends, where everyone contributes equally to the discussion. It helps to talk through the material with each other, and provides a great place to ask questions, no matter how basic, and offers you the opportunity to strengthen each others' weaknesses.
- You then need to write up your own, unique version of each of the problem solutions. Do not lend your homework solutions to or borrow from a friend in order to be "reminded" of how to go about doing the problems.
- Remember always to work in Standard International (SI) units: MKS. This means converting all of your given quantities to MKS *before starting the problem.* That way all your answers will automatically be in MKS units as well.

Appendix C

Notes and Frequencies of the Tempered Scale

The Western musical scale is divided into 12 semitones per octave. The traditional way of identifying them is through the letters A through G with sharps (♯) and flats (♭) for denoting semitones between some of them. Figures C.1 and C.2 show the notes as they correspond to the keys of the piano. Frequency values for notes in the tempered scale (see chapter 14) are listed in table C.1, using the frequency of $A_4 = 440$ Hz ("concert A") as reference point.

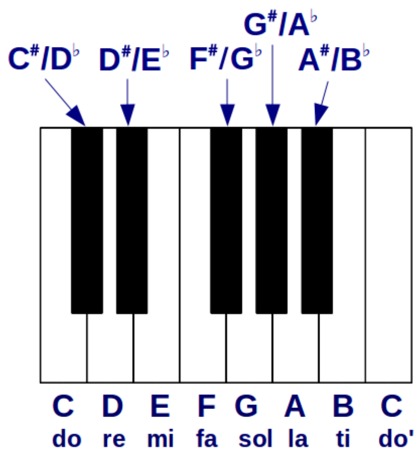


Figure C.1: A depiction of the full diatonic scale on keyboard, including the solfège notes.

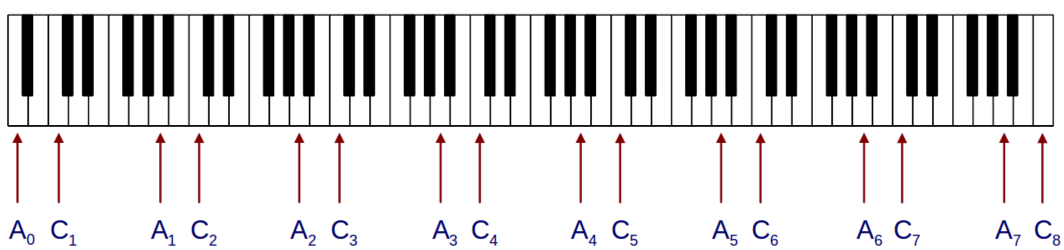


Figure C.2: The full piano keyboard with some key notes identified

Table C.1: Notes of the full tempered scale, with their frequencies noted (in Hz).

C_0	16.352	C_3	130.81	C_6	1046.5
$C^\sharp_0/D^\flat_0$	17.324	$C^\sharp_3/D^\flat_3$	138.59	$C^\sharp_6/D^\flat_6$	1108.7
D_0	18.354	D_3	146.83	D_6	1174.7
$D^\sharp_0/E^\flat_0$	19.445	$D^\sharp_3/E^\flat_3$	155.56	$D^\sharp_6/E^\flat_6$	1244.5
E_0	20.602	E_3	164.81	E_6	1318.5
F_0	21.827	F_3	174.61	F_6	1396.6
$F^\sharp_0/G^\flat_0$	23.125	$F^\sharp_3/G^\flat_3$	185.00	$F^\sharp_6/G^\flat_6$	1480.0
G_0	24.500	G_3	196.00	G_6	1568.0
$G^\sharp_0/A^\flat_0$	25.957	$G^\sharp_3/A^\flat_3$	207.65	$G^\sharp_6/A^\flat_6$	1661.2
A_0	27.500	A_3	220.00	A_6	1760.0
$A^\sharp_0/B^\flat_0$	29.135	$A^\sharp_3/B^\flat_3$	233.08	$A^\sharp_6/B^\flat_6$	1864.7
B_0	30.868	B_3	246.94	B_6	1975.5
C_1	32.703	C_4	261.63	C_7	2093.0
$C^\sharp_1/D^\flat_1$	34.648	$C^\sharp_4/D^\flat_4$	277.18	$C^\sharp_7/D^\flat_7$	2217.5
D_1	36.708	D_4	293.66	D_7	2349.3
$D^\sharp_1/E^\flat_1$	38.891	$D^\sharp_4/E^\flat_4$	311.13	$D^\sharp_7/E^\flat_7$	2489.0
E_1	41.203	E_4	329.63	E_7	2637.0
F_1	43.654	F_4	349.23	F_7	2793.8
$F^\sharp_1/G^\flat_1$	46.249	$F^\sharp_4/G^\flat_4$	369.99	$F^\sharp_7/G^\flat_7$	2960.0
G_1	48.999	G_4	392.00	G_7	3136.0
$G^\sharp_1/A^\flat_1$	51.913	$G^\sharp_4/A^\flat_4$	415.30	$G^\sharp_7/A^\flat_7$	3322.4
A_1	55.000	A_4	440.00	A_7	3520.0
$A^\sharp_1/B^\flat_1$	58.270	$A^\sharp_4/B^\flat_4$	466.16	$A^\sharp_7/B^\flat_7$	3729.3
B_1	61.735	B_4	493.88	B_7	3951.1
C_2	65.406	C_5	523.25	C_8	4186.0
$C^\sharp_2/D^\flat_2$	69.296	$C^\sharp_5/D^\flat_5$	554.37	$C^\sharp_8/D^\flat_8$	4434.9
D_2	73.416	D_5	587.33	D_8	4698.6
$D^\sharp_2/E^\flat_2$	77.782	$D^\sharp_5/E^\flat_5$	622.25	$D^\sharp_8/E^\flat_8$	4978.0
E_2	82.407	E_5	659.26	E_8	5274.0
F_2	87.307	F_5	698.46	F_8	5587.7
$F^\sharp_2/G^\flat_2$	92.499	$F^\sharp_5/G^\flat_5$	739.99	$F^\sharp_8/G^\flat_8$	5919.9
G_2	94.999	G_5	783.99	G_8	6271.9
$G^\sharp_2/A^\flat_2$	103.83	$G^\sharp_5/A^\flat_5$	830.61	$G^\sharp_8/A^\flat_8$	6644.9
A_2	110.00	A_5	880.00	A_8	7040.0
$A^\sharp_2/B^\flat_2$	116.54	$A^\sharp_5/B^\flat_5$	932.33	$A^\sharp_8/B^\flat_8$	7458.6
B_2	123.47	B_5	987.77	B_8	7902.1

Appendix D

Glossary of Terms

Absolute Temperature, T – The temperature (in kelvins) on a scale that has its zero at the lowest attainable temperature (-273°C); absolute temperature is found by adding 273 to the Celsius temperature.

Absorption, A – The partial loss of a sound wave when reflecting from a surface. Part of the soundwave is lost to heat in the material and is absorbed. The absorption (measured in *sabins*) is equal to the surface area multiplied by the material's absorption coefficient, which ranges from 0 (perfectly reflecting) to 1 (perfectly absorbing).

Additive Synthesis – Sound synthesis based on adding together many simple waveforms, such as sinewaves at various frequencies, amplitudes, and phase offsets.

ADSR Envelope Generator – A, D, S, and R refer to the parameters of an envelope: A = attack time; D = (initial) decay time; S = sustain level; R = (final) release time.

Air Suspension Speaker – A loudspeaker mounted in the front of an airtight box so that the pressure of the enclosed air furnishes a major part of the force that restores the speaker cone to its equilibrium position.

Amplifier – A device in which a small amount of input power controls a larger amount of output power.

Amplitude, A – the largest value of motion, or "distance," from its equilibrium position for an object in periodic motion. For example, if a mass on a spring is executing periodic motion and its largest distance from the equilibrium position is ± 3 cm, then its amplitude is 3 cm and its motion is between these two extreme values.

Amplitude Modulation (AM) – The method of radio broadcasting in which the amplitude of a carrier wave is determined by the audio signal. The AM band extends from 540 to 1600 kHz.

Analog to Digital Converter (ADC) – A circuit that converts numbers from an analog to a digital representation.

Anechoic – Echo-free; a term applied to a specially designed room with highly absorbing walls.

Antinode – see *Node*.

Auditory Canal – A tube in the outer ear that transmits sound from the external pinna to the tympanic membrane, or eardrum.

Auditory Membrane – See *Eardrum*.

Aural Harmonic – A harmonic that is generated in the auditory system by a non-linear response.

Axon – That part of a neuron or nerve cell that transmits neural pulses to other neurons.

Basilar Membrane – A thin membrane spanning the entire length and width of the cochlea. It contains 20,000 - 30,000 thin fibers extending across the width of the cochlea with varying lengths and stiffnesses. Each fiber has a resonant frequency that can be set into vibration by individual sound spectrum components of a sound wave traveling the length of the cochlea.

Bass – Sounds with very low pitch, generally below about G_3, corresponding to frequencies below 200 Hz.

Bass Bar – The wood strip that stiffens the top plate of a violin or other string instrument and distributes the vibrations of the bridge up and down the plate.

Beats – Periodic variations in amplitude that result from the superposition or addition of two tones with nearly the same frequency.

Bell – The flared section that terminates all brass instruments and determines their radiation characteristics.

Bernoulli's Law – An increase in the speed with which a fluid (or air) flows corresponds to a decrease in pressure.

Binaural– Sound reproduction using two microphones (usually in a "dummy" head) and played through headphones, so that the listener hears the sound he or she would have heard at the recording location.

Bit – A binary digit, used in the sense of a minimum unit of computer memory as well as a small, two-valued unit of information.

Boundary Conditions – A physical system that supports vibrational modes has a certain physical size and material makeup. The way in which these attributes translate into allowed and disallowed modes of vibration constitute boundary conditions. The boundary conditions enable us to determine which vibrational modes are supported by the medium and which are not. A good example of such a system is a string under tension. They physical distance between the tied-down ends of the string dictates which wavelengths of vibration will be supported along the string's length. The allowed frequencies of vibration will then be related to the allowed wavelengths and the speed with which waves move through the medium. Boundary conditions often result in *discrete* or *quantized* solutions to the allowed frequencies and wavelengths supported by the system.

Brass – Wind instruments consisting of resonating tubes using the lips as the source of vibration; often made of brass, but not necessarily, and often employing a flare or bell at the open end.

Bridge – The wood piece that transmits string vibrations to the sound board or top plate.

Byte – An ordered collection of 8 bits of computer memory.

Carrier Wave – A high-frequency electromagnetic wave capable of being modulated to transmit a signal of lower frequency.

Cent, ¢ – $1/100^{th}$ of a semitone.

Chladni Plate – A means for studying vibrational modes of a plate by making nodal lines visible with powder.

Chroma – All tones in the 12-tone chromatic scale are said to have different chroma, or pitch "color."

Chroma corresponds to a continuous sense of pitch variation. Each of the tones in the chromatic scale have distinct chroma, even though it is hard to pin down exactly what it is in essence (much the same as it is with visual color).

Chromatic Scale – The 12-tone scale complete with evenly spaced semitones spanning one complete octave.

Combination Tones – A secondary tone heard when two primary tones are played. Combination tones are usually different tones than the original two.

Compression – A region in a medium where the pressure is above the nominal pressure of the medium by virtue of a disturbance propagating through the medium. See also *rarefaction*.

Condenser Microphone – A microphone in which the diaphragm serves as one plate of a small capacitor or condenser. As the diaphragm moves, the electrical charge on the condenser varies.

Conical Tube – A tube geometry that supports vibrational resonances similar to the open tube, *i.e.* all even and odd harmonics of the fundamental. If it is truncated at the small end, the resonant structure is the same as long as the portion removed represents no more than 25% of the original cone length. If more than 25% is removed, the harmonics become overtones, not integral multiples of the fundamental. The oboe, the bassoon, and the saxophone are woodwind instruments based on this geometry. See also *Open Tube* and *Semiclosed Tube*.

Consonance – The pleasant character associated with certain musical intervals or chords. Consonance is associated with a peaceful or restful feeling, and is dependent both on the particular intervals played and the context in which they are heard. See *Dissonance*.

Critical Band, $\mathbf{\Delta f_{CB}}$ – See figure 8.1. The value of frequency difference between two tones necessary for the auditory system to process the two tones with roughness or with clarity. This band is a function of the frequency of the average frequency of two tones. See *Discrimination Band*.

Crossover Network – A network designed to send high frequencies to the tweeter and low frequencies to the woofer. The crossover frequency is the approximate point of division between high and low frequencies.

Current, *I* – Rate of flow of electric charge in a wire. Measured in amperes, A.

Damping – The gradual dissipative loss of vibrational energy in an oscillating system, usual to friction or sound radiation.

Decibels, *dB* – A dimensionless unit used to compare the ratio of two quantities (such as sound pressure, power, or intensity), or to express the ratio of one such quantity to an appropriate reference. The unit for sound level or sound level differences.

Dendrite – That part of a neuron that receives neural pulses from other neurons.

Diaphragm – The dome-shaped muscle that forms a floor for the chest cavity.

Diatonic Scale – The Western scale of 7 notes spanning the octave, consisting of 5 whole tone and 2 semitone intervals in the following sequence: WWSWWWS where W is for whole tone and S is for semitone.

Difference Tone – When two tones having frequencies f_1 and f_2 are sounded together a difference tone with frequency $f_2 - f_1$ is often heard.

Diffraction – The spreading out of waves when they encounter a barrier or pass through a narrow opening.

Digital to Analog Converter – A circuit that converts numbers from a digital to an analog representation.

Diphthong – A combination of two or more vowels into one phoneme.

Direct Sound – Sound that reaches the listener without being reflected.

Discrimination Band – See figure 8.1. The value of frequency difference between two tones necessary for the auditory system to process the two as separate tones. This band is a function of the average frequency of the two tones. When the difference between the two tones is below the discrimination band value, they are perceived as a single "fused" tone with roughness. Above this value, the two are perceived as separate tones with roughness or clarity, depending on this separation. (See *Critical Band*).

Displacement – A distance away from an equilibrium point. When a mass on a spring is stretched a certain distance from equilibrium, it is said to have a displacement equal to this distance.

Dissonance – The unpleasant sound or "restlessness" associated with certain musical intervals or chords. Dissonance is often associated with tension, with a perceived need to move toward a peaceful resolution, or consonance.

Distortion – An undesired change in waveform that creates unpleasant sound. Two common examples are harmonic distortion and intermodulation distortion.

Dolby System – A widely used noise-reduction system that boosts low-level high-frequency signals when recorded and reduces them in playback.

Dynamic Microphone – A microphone that generates an electrical voltage by the movement of a coil of wire in a magnetic field.

Eardrum – A fibrous membrane that terminates the auditory canal and is caused to vibrate by incoming sound waves.

Elasticity – When a disturbance passes through a medium, it causes displacement of individual particles of the medium from their equilibrium positions. It is the *elasticity* or *springiness* of the medium that serves to restore the particles back to their equilibrium positions. This *restoring force* is always directed toward the equilibrium position and is therefore characterized as a negative force (i.e. it is always points in the opposite direction to the particle's displacement). See also *Inertia*.

Electromagnetic Force – The force that results from the interaction of an electric current with a magnetic field.

Electromagnetic Induction – Generation of electrical voltage in a wire that moves across a magnetic field.

Embouchure – The lip position used in playing a wind instrument. The embouchure hole in a flute is the hole through which the lips blow air.

Energy, *E* – Energy comes in many forms, and can be transformed from one form to another. While we don't know exactly what energy *is*, we do know that in all processes in nature, the total amount of energy in the system is conserved.

Envelope – Time variation of the amplitude (or energy) of a vibration.

Equalization – Boosting some frequencies and attenuating others to meet some prescribed recipe for sound. Equalization is usually applied during both recording and playback.

Equilibrium – The state of a system "at rest" when it has no net force acting on it. A mass on a spring is in equilibrium when it is at rest in its "home" position where the string is relaxed. When it is moved away from this point, it is no longer in equilibrium but will move as a result of the spring force.

Equilibrium Position – The position of an object when it is at rest. When set into oscillation, the object will oscillate about this equilibrium position. For example, the equilibrium position of a pendulum corresponds to the position of the bob when the pendulum is at rest. When the bob is pulled to one side and released, the bob proceeds to oscillate about this equilibrium position. Eventually when the motion dies down, the bob will once again come to rest at the equilibrium position.

Eustachian Tube – A tube connecting the middle ear to the oral cavity that allows the average pressure in the middle ear to equal atmospheric pressure.

Exponent – The number expressing the power to which 10 or some other number is raised.

Faraday's Law – A moving or changing magnetic field induces a current in nearby conductors, such as a coil of wire.

Filters (High-pass, Low-pass, and Band-pass) – Acoustic elements that allow certain frequencies to be transmitted while attenuating others. A high-pass filter allows all components above a cutoff frequency to be transmitted; a low-pass filter allows all components below a cutoff frequency to be transmitted; a band-pass filter allows frequencies within a certain band to pass.

f-holes – The openings in the top plate of a string instrument shaped like the letter f.

Flow Rate – The volume of air that flows past a certain point and is measured per second.

Force, $\boldsymbol{F}$ – In simple terms, a force is a push or a pull. It always involves two bodies interacting with one another. We say that one body *exerts a force on another*. The unit of force is the *Newton*, *N*. One Newton is defined as the force necessary to cause a body with a mass of 1 kg to accelerate at 1 m/s^2, that is, to cause the body to have a speed that increases by an additional 1 m/s every second.

Formant – A resonance in the vocal tract that serves to support and strengthen vocal fold harmonics in the region of that resonance.

Fourier Analysis – The determination of the component tones that make up a complex tone or waveform.

Fourier Synthesis – The creation of a complex tone or waveform by combining spectral components.

Free Field – A reflection-free environment, such as exists outdoors or in an anechoic room, in which sound pressure varies inversely with distance ($p \propto 1/r$).

Frequency, $\boldsymbol{f}$ – The number of cycles a periodic system executes in one second, measured in units of Hertz (Hz), or cycles/second.

Frequency Modulation – The method of radio broadcasting in which the frequency of a carrier wave is modulated (altered slightly) by the audio signal. The FM band, which extends from 88 to 108 MHz, allows stations sufficient bandwidth to transmit high-fidelity stereophonic sound.

Fricatives – Consonants that are formed by constricting air flow in the vocal tract (such as f, v, s, z, th, sh, *etc.*).

Fundamental – The lowest frequency in a family of harmonics produced by a tonal instrument.

Fundamental Tracking – The auditory system's tendency to associate the pitch of a complex tone with the fundamental harmonic, even if the fundamental is missing. Combination tones between the harmonics support the vibration of the fundamental.

Glottis – The V-shaped opening between the vocal folds.

Hair Cells – The tiny sensors of sound in the cochlea.

Harmonic Distortion – Distortion of sound by the introduction of extra harmonics produced by non-linearities in the ear or sound equipment.

Harmonic Singing – Tuning vocal tract resonances to a harmonic of the vocal-fold vibration frequency to produce a single or multiple tones.

Harmonics – Vibrations in a complex tone whose frequencies are integral multiples of the fundamental tone, or lowest frequency of the tone.

Harmonic Distortion – The creation of harmonics (frequency multiples) of the original signal by some type of nonlinearity in the system (the most common cause is overdriving some component).

Helmholtz Resonator – A vibrator consisting of a volume of enclosed air with an open neck or port.

Hertz, *Hz* – A unit of frequency, named after Heinrich Hertz, corresponding to 1 cycle per second.

High-fidelity Sound – Sound that reproduces much of the spectrum, dynamic range, and spatial characteristics of the original sound and adds minimal distortion.

Inertia – That quality of a medium that "resists" being moved. When a disturbance propagates through a medium, it causes the particles of the medium to be displaced from their equilibrium positions, and the elasticity of the medium then serves to turn the displaced particles around and force them back toward equilibrium. Because a particle is in motion when it approaches its equilibrium point, its inertia causes it to "overshoot" that point (see *Principle of Inertia*), after which the elasticity of the medium once again turns it around, forcing it back toward equilibrium. This oscillation will continue as long as the wave continues to displace the particles of the medium, after which damping forces cause all oscillation to to cease.

Infrasonic – Having a frequency below the audible range.

Inharmonicity – The departure of the frequencies from those of a harmonic series.

Initial Time Delay Gap – The time difference between the arrival of the direct sound and the first reflected sound at the ear.

Intensity – The sound intensity is a measure of how much power a wave carries per unit area, and is measured in Watts/m^2.

Interference – The interaction of two or more identical waves, which may support (constructive interference) or cancel (destructive interference) each other.

Interval – A separation between two pitches, used in music to form consonant and dissonant combinations.

Intonation – Refers to the degree or accuracy with which pitches are produced.

Joule, *J* – A unit of energy or work; one joule is equal to one newton-meter, also one watt-second.

Just Temperament – A system of tuning that attempts to make thirds, fourths, and fifths as consonant as possible; it is based on major triads with frequency ratios 4:5:6.

Kinetic Energy, *K* – A form of energy associated with *motion*. The kinetic energy of a mass m moving at speed v is $K = \frac{1}{2}mv^2$, and the units of kinetic energy (as for energy in general) are *Newtons*.

Larynx – The section of the vocal system, composed mainly of cartilage, that contains the vocal folds.

Linear Scale – A scale in which moving a given distance right or left adds or subtracts a given increment.

Localization – The ability to determine the location or direction of a sound source.

Logarithm – The power to which 10 (or some other base) must be raised to give the desired number.

Logarithmic Scale – A scale on which moving a given distance right or left multiplies or divides by a given factor.

Longitudinal Wave – A wave for which the particles of the medium move along a line parallel to the wave's direction of propagation through the medium. An example would be a sound wave, for which the particles of the air move back and forth in the same direction as the wave's propagation. See also *Transverse Wave*.

Loudness, *L* – The loudness, expressed in units of sones, corresponds to the ear's ability to discern different levels of loudness for a particular frequency. A 1000 Hz pure tone with SL = 40 db is defined to have a loudness of 1 sone. The scale is then set up so that, for example, a loudness of 8 sones is judged by the average listener to be twice as loud as a loudness of 4 sones. This measure is *subjective* in nature, having to do with the particular response of the human ear to various levels of sound intensity.

Loudness Level, *LL* – A subjective measure of loudness as heard by the human ear. Tones of different frequency need to be set at different intensities in order for the ear to perceive them as having the same loudness level, owing to the non-linear frequency response of the ear. The loudness level is measured in *phons*, and is defined to be equal in value to the sound level (measured in *decibels*) for the standard tone of frequency 1000 Hz.

Main Wood Resonance, *MWR* – Normally in reference to one of the instruments of the string family (violin, viola, cello, double bass), this resonance corresponds to the lowest resonant frequency of the of the body of the instrument. The standing wave supported by the box of the instrument, consisting of the top and bottom plates and the sides, has a natural frequency whose value depends on the specific dimensions of the body as well as the speed with which waves travel in the wood plates. The placement of this resonant frequency in relation to the tuning of the open strings helps determine the quality of the instrument. See also *Main Air Resonance*.

Main Air Resonance, *MAR* – Normally in reference to one of the instruments of the string family (violin, viola, cello, double bass), this resonance corresponds to the lowest resonant frequency of the enclosed air volume of the instrument. The standing wave supported by the air in the volume has a natural frequency whose value results from the combination of the specific dimensions of the volume and the speed of sound in air. The placement of this resonant frequency value in relation to the tuning of the open strings helps determine the quality of the instrument. See also *Main Wood Resonance*.

Major Diatonic Scale – A scale of seven notes with the following sequence of intervals: two whole notes, one semitone, three whole tones, one semitone.

Major Triad – A chord of three notes having intervals of a major third and a minor third, respectively (as: C:E:G).

Masking – The obscuring of one sound by another.

Mass, M – A measure of resistance to change in motion; equal to the force divided by acceleration.

Meantone Temperament – A system of tuning which raises or lowers various notes from their Pythagorean values by quarters of a syntonic comma.

Medium – Material through which a wave passes, such as air. A medium is necessary for the support of wave propagation from one place to another. Two essential characteristics of any medium that supports wave propagation are *Elasticity* and *Inertia*.

Microtone – Any interval smaller than a semitone.

MIDI – The musical instrument digital interface communications standard adopted widely in the music synthesizer industry in the early 1980s.

Minor Scale – A scale with one to three notes lowered by a semitone from the corresponding major scale.

Minor Triad – A chord of three notes having intervals of a minor third and a major third, respectively (as: C:E$^\flat$:G).

Missing Fundamental – Pitch judgment can focus on a fundamental tone that may not be physically present. Harmonics of the fundamental produce combination tones that support the fundamental such that even if it is not present, pitch perception perceives it to be.

Modulate – To change some parameter (usually amplitude or frequency) of one signal in proportion to another signal. In music, the transition from one musical key to another.

Monaural – Sound reproduction using one microphone to feed a single headphone, such as is used in telephone communication.

Mouthpiece – The part of a brass or wind instrument that couples the vibrating lips or reed to the air column.

MP3 – The common name given to digital audio that is encoded according to a particular standard known as the motion picture expert's group (MPEG)-1, Layer III. The perception-based encoding process compresses the digital representation greatly, allowing it to be readily transmitted or stored.

Natural Frequency, F_0 – Physical objects have an elastic character, so that when they undergo a sudden deformation, perhaps from being struck or bumped, they oscillate with definite frequencies characteristic of the object itself. These frequencies are called the object's *natural frequencies*.

Neuron, or Nerve Cell – Building block of the nervous system that both transmits and processes neural pulses.

Newton, N – A unit of force

Nibble – An ordered collection of 4 bits of computer memory.

Node – In a vibrating system, a node is defined as a point at which the vibrational amplitude is zero.

An example would be the ends of a stretched string where it is tied down, at which it cannot vibrate. The ends are therefore nodes of vibration. Higher harmonic modes of vibration have nodes located along the length of the string. Alternatively, a point in a vibrating system at which the vibrational amplitude is a maximum is defined as an *antinode*.

Nut – The strip of hard material that supports the string at the head end of a violin or guitar or other string instrument.

Octave – A particular musical interval corresponding to a frequency ratio of two. The perception is of two tones with the same chroma but different tone height.

Open Tube – Sometimes referred to as an open-open tube, a cylindrical tube that is open on both ends. The flute is a classic woodwind instrument based on this geometry. The open tube geometry supports the vibration of all even and odd harmonics of the fundamental. See also *Semiclosed Tube* and *Conical Tube*.

Organ of Corti – The part of the cochlea containing the hair-cells; the "seat of hearing."

Ossicles – Three small bones of the middle ear that transmit vibrations from the eardrum to the cochlea.

Overtone – A vibration in a complex tone whose frequency is equal to a non-integral multiple of the fundamental tone, or lowest frequency of the family.

Palate – The roof of the mouth.

Pascal, *Pa* – A unit of pressure equivalent to one N/m^2. See *Pressure*.

Patch – (verb) To interconnect; or (noun) a set of interconnections that causes a synthesizer to produce certain types of sounds.

Pentatonic Scale – A scale of five notes used in several musical cultures, such as Chinese, Native American, and Celtic cultures.

Period, *T* – The time it takes for a periodic system to complete one cycle of motion, measured in units of seconds.

Periodic Motion – A motion that repeats itself with the same pattern of motion, however complex it may be.

Pharynx – The lower part of the vocal tract connecting the larynx and the oral cavity.

Phase, ϕ – The relationship between two signals of the same frequency but different points in their cyclical variation. "In phase" means that they both peak together, for example, whereas "out of phase" means that one has a peak where the other has a valley. Phase is a continuous quantity ranging from 0 to 360^o or 0 to 2π radians.

Phase Difference, $\Delta\phi$ – A measure of the relative position of two vibrating objects at a given time; also the relative positions, in a vibrating cycle, of a vibrating object and a driving force.

Phon – Unit denoting the loudness level of any sound. Defined such that the numerical value of phons for a particular sound at 1000 Hz is equal to the numerical value of sound level for that tone, measured in db.

Phonemes – Individual units of sound that make up speech.

Phonetics – The study of speech sounds.

Physics – The study of matter and its interaction with energy. The basic goal of physics is an understanding of the physical universe in which we live, on all size scales and through all time.

Physics of Music – The study of music (or more broadly sound) from a physical science standpoint, including the origin of sound (vibrations, resonances), the travel of sound (as a propagating wave in a medium), and the reception of sound (by the human ear or recording instrumentation). It includes (but is not limited to) an analysis of complex waveforms and frequency spectra resulting from combinations of fundamental tones and harmonics, a detailed analysis of the human ear and voice, the perception of harmony and dissonance, the nature of musical instruments and the various timbres resulting from their unique designs, the behavior of sound in an enclosure and the contribution from room acoustics, and the recording and playback of music.

Pink Noise – Low-pass-filtered random noise for which the energy contained in each octave band is the same.

Pinna – The external part of the ear.

Pitch – The psychological sensation of the *chroma* of a sound and its tone height. Pitches are assigned note names to denote their position in the musical scale.

Plosives – Consonants that are produced by suddenly removing a constriction in the vocal tract (p,b,t,d,k,g).

Potential Energy, U – A form of "stored" energy able to be turned into kinetic energy. A mass on a spring has both kinetic and potential energy as it oscillates. When it passes through its equilibrium point, it has maximum kinetic energy and zero potential energy. When it reaches its extreme amplitude positions (when the spring is maximally stretched or compressed) it briefly stops (at which point the kinetic energy is zero), and has maximum potential energy. The stored energy is contained within the spring.

Power, P – The rate of doing work or expending energy, equal to energy divided by time, measured in the MKS unit Watt (W).

Power Density – A measure of how much power is spread over a unit area.

Pressure, P – Pressure is defined as *force per unit area.* It corresponds to a force that is spread over a certain area. For constant force, the larger the area over which it is applied, the lower the pressure. The smaller the area over which it is applied, the higher the pressure. Measured in the MKS unit N/m^2, which is is defined as one Pascal (Pa).

Principle of Inertia – States that a body at rest tends to stay at rest, and a body in motion tends to stay in motion, unless acted on by an external force. Both Galileo Galilei and Isaac Newton formulated versions of this principle. A body's *inertial* is what causes it to preserve its state of motion, whether moving or at rest. See *Inertia.*

Prosodic Feature – A characteristic of speech, such as pitch, rhythm, and accent, that is used to convey meaning, emphasis, and emotion.

Pythagorean Comma – The small difference between two kinds of semitones (chromatic and diatonic) in the Pythagorean tuning; a frequency ratio 1.0136 corresponding to 23.5¢.

Pythagorean Temperament – A system of pitches based on perfect fifths and fourths.

Rarefaction – A region in a medium where the pressure is below the nominal pressure of the medium, by virtue of a disturbance propagating through the medium. See also *compression.*

Real-time Spectrum Analyzer – An instrument that rapidly creates a spectrum of a sound.

Reflection – An abrupt change in the direction of wave propagation at a change of medium (by waves that remain in the first medium).

Refraction – A bending of waves when the speed of propagation changes, either abruptly (at a change in medium) or gradually (*e.g.* sound waves in a wind of varying speed).

Register Hole – A hole that can be opened in order to cause an instrument to play in a higher register.

Resonance – A phenomenon that occurs when the frequency of a periodic force applied to an object equals the natural frequency of the object. The result is large response from the object, such that a small force produces large amplitude response. Consider, for example, a tuning fork attached to a hollow wooden box. If the wooden box is designed so that the air inside has a natural oscillation frequency equal to that of the tuning fork, then when the tuning fork is struck, it provides the periodic force that drives the air in the box into resonance, resulting in large-amplitude oscillations and amplified sound.

Restoring Force, $\boldsymbol{f}_R$ – A force that always points toward an equilibrium point and therefore serves to return an object displaced from equilibrium.

Reverberant Sound – Sound that builds up and decays gradually and can be "stored" in a room for an appreciable time.

Reverberation – Caused by reflections of sound in a room, and gives rise to a smoothing of the sound both in its buildup and in its decay.

Reverberation Time – Defined as the time it takes for the intensity of a sound to die out to 1 part in 10^6 of its initial intensity, corresponding to a 60 db drop.

Sabin – Unit for measuring absorption of sound; the sabin is equivalent to one square meter of open window.

Scale – Succession of notes arranged in ascending or descending order.

Semiclosed Tube – A cylindrical tube open on one end and closed on the other. Able to support the vibration of an odd number of quarter wavelengths, this tube geometry only supports the odd integral harmonics of the fundamental. See also *Open Tube* and *Conical Tube.*

Semitone – One step on a chromatic scale. Normally $\frac{1}{12}$ of an octave.

Signal to Noise Ratio – The ratio (usually expressed in dB) of the average recorded signal to the background noise.

Simple Harmonic Motion – Motion that results from a mass under the influence of a linear restoring force. Linear means that as the mass is displaced from its equilibrium position, the force exerted on it by the restoring force is linearly proportional to its displacement, and always directed back toward the equilibrium position. The position of an object, relative to its equilibrium position, as it undergoes simple harmonic motion, is sinusoidal in character. The motion is symmetric about its equilibrium position, it repeats itself in a definite time interval called the period, and the speed of the oscillator varies through out the cycle in a regular fashion, taking on its maximum value as the mass moves through the equilibrium, and zero when it reaches the extremes of its motion. Examples of simple harmonic motion are a mass on a spring and a pendulum bob swinging with low angular amplitude.

Sine Wave – A waveform that is characteristic of a pure tone (that is, a tone without harmonics or overtones) and also simple harmonic motion.

Singer's Formant – A resonance around 2500 to 3000 Hz in male (and low female) voices that adds brilliance to the tone.

Sone – A unit used to express subjective loudness; doubling the number of sones should describe sound twice as loud.

Sonic Boom – Pressure transient that occurs during the flyover of an aircraft traveling faster than the speed of sound.

Sound Level, *SL* – A logarithmic representation of sound intensity, defined as $10\log\left(I_2/I_1\right)$ and is measured in decibels. The sound level is a *relative* measurement, meaning that it corresponds to the difference between two separate sounds, i.e. we talk about the sound level of one sound relative to another. The *absolute* sound level is the value relative to the threshold of hearing, which has an intensity of 10^{-12} Watts/cm^2.

Sound Post – The short round stick (of spruce) connecting the top and back plates of a violin or other string instrument.

Sound Power Level – $L_W = 10\log W/W_0$, where W is sound power and $W_0 = 10^{-12}$ W (abbreviated PWL or L_W).

Sound Pressure Level – $L_p = 20\log p/p_0$, where p is sound pressure and $p_0 = 2\times10^{-5}$ N/m^2 (or 20 micro-pascals) (abbreviated SPL or L_p).

Sound Spectrograph – An instrument that displays sound level as a function of frequency and time for a brief sample of speech.

Sound Waves – Disturbances that move through a medium as zones of compression and rarefaction. Sound waves propagate through the air as *longitudinal waves*, for which the particles of the medium are forced into motion along the same direction as the wave's propagation.

Spectrogram – A graph of sound level versus frequency and time as recorded on a sound spectrograph or similar instrument.

Spectrum – A plot that serves as a "recipe" that gives the frequency and amplitude of each component of a complex vibration.

Speech Synthesis – Creating speech like sounds artificially.

Speed of Sound, v_s – Sound is broadly defined as a compression disturbance that propagates through any medium, whether gas, liquid, or solid. The speed with which the disturbance moves is the speed of sound. In air, the speed of sound is given by $v_s = 20.1\sqrt{T_A}$. In terms of the wavelength and frequency of the disturbance, $v = f\lambda$.

Spring Constant, k – Characterizes the strength of the force exerted by a spring as it is stretched from equilibrium. The force that a spring exerts on an agent stretching it from its equilibrium position force is described by $F = -kx$, where F is the force, x is the displacement from equilibrium, and k is the spring constant. The units of the spring constant are *Newtons/m*.

Springiness – See *Elasticity*.

Standing Wave – A wavelike pattern that results from the interference of two or more waves; a standing wave has regions of minimum and maximum amplitude called nodes and antinodes that stay in place. The wave has the appearance of "standing" still.

Stereophonic – Sound reproduction using two microphones to feed two loudspeakers.

Stretch Tuning – Tuning octaves slightly larger than a 2:1 ratio.

Subtractive Synthesis – Sound synthesis based on the controlled attenuation or removal (by filtering) of components from a multicomponent waveform, such as a sawtooth or triangular waveform.

Subwoofer – A loudspeaker designed to produce extra-low frequency sound, below the woofer.

Superposition – The process of adding two or more waves together to form one. When a speaker cone vibrates under the simultaneous application of two different sine waves of different frequency, it can't follow each sine wave individually (since it can only move in one direction at a time) but rather vibrates with the motion of their superposition. The superposition is found by adding the amplitudes of the various waves together at any particular point along the x-axis. By finding this sum amplitude at several points on the axis, one can find the overall shape of the superposition curve.

Supersonic – Having a speed greater than that of sound (approximately 340 m/s or 770 mph).

Sustaining Pedal – Right-hand pedal of a piano which raises all the dampers, allowing the strings to continue vibrating after the keys are released.

Sympathetic Vibration – One vibrator causing another to vibrate at the same frequency (which may or may not be a resonant frequency). An example is a piano string causing the bridge and soundboard to vibrate at the string's frequency.

Syntonic Comma – The small difference between a major or minor third in the Pythagorean and just tunings.

Temperament – System of tuning in which intervals deviate from acoustically "pure" (Pythagorean) intervals.

Tension, T – The force applied to the two ends of a string, or around the periphery of a membrane, that provides a restoring force during vibration.

Timbre – An attribute of auditory sensation by which two sounds with the same loudness and pitch can be judged as dissimilar.

Tone Height – Notes with the same pitch, or chroma, that are separated by one more more octaves are said to have different tone height.

Transducer – A device that converts one form of energy into another.

Transverse Wave – A wave for which the particles of the medium move along a line perpendicular to the wave's direction of propagation through the medium. An example would be a wave on a guitar string, for which the particles of the string move up and down while the wave moves back and forth along the string. See also *Transverse Wave*.

Triad – A chord of three notes; in the just tuning, a major third has frequency ratios 4:5:6, and a minor third has ratios 10:12:15. See *Major Triad* and *Minor Triad*.

Tuning – An adjustment of pitch in any instrument so that it corresponds to an accepted norm.

Tweeter – A loudspeaker designed to produce high-frequency sound.

Tympanic Membrane – See *Eardrum*.

Ultrasonic – Having a frequency above the audible range.

VCA (Voltage Controlled Amplifier) – An amplifier whose gain varies linearly or exponentially in proportion to a control voltage.

VCF (Voltage Controlled Filter) – A filter whose cutoff frequency (or center frequency and bandwidth) varies linearly or exponentially in proportion to one or more control voltages.

VCO (Voltage Controlled Oscillator – An oscillator whose frequency varies linearly or exponentially in proportion to one or more control voltages.

Vibrato – Frequency modulation that may or may not have amplitude modulation associated with it.

Vocal Folds – Folds of ligament extending across the larynx that interrupt the flow of air to produce sound.

Vocal Tract – The tube connecting the larynx to the mouth consisting of the pharynx and the oral cavity.

Voiceprints – Speech spectrograms from which a speaker's identity may be determined.

Waveform – Graph of some variable (*e.g.* position of an oscillating mass or sound pressure) versus time.

Watt, $\boldsymbol{W}$ – A unit of power equal to 1 Joule per second, J/s.

Wave – A disturbance that propagates through a medium. Wave carries energy, and displace particles of the medium as they pass through. See *longitudinal* and *transverse* waves.

Wavelength – The distance between corresponding points on two successive cycles of a wave.

White Noise – Noise whose amplitude is constant throughout the audible frequency range.

Woofer – A loudspeaker designed to produce low-frequency sound.

Work, $\boldsymbol{W}$ – Work is defined as something that is done on an object when a force is applied to it and it moves in the same direction as the applied force. The work is the product of the force and the distance moved, $W = F \cdot d$ and is measured in Joules, the MKS unit for energy.

Appendix E

Color Figures

The following color versions of selected chapter figures are reproduced here as an additional aid to study. Color has been included in order to help highlight important features that may be less clear in their greyscale chapter versions.

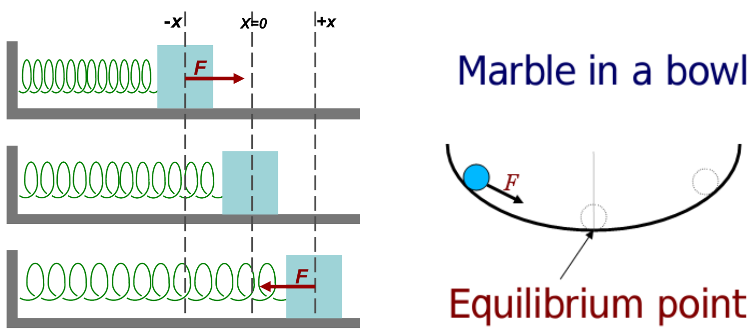


Figure 3.5: Left: A mass attached to a spring. In the middle figure the mass is at rest at the equilibrium position. Note that the force F exerted by the spring is always directed toward the equilibrium point. At the top the mass is pushed to the left (compressing the spring), to which the spring responds with a force to the right (toward the equilibrium point). At the bottom, the mass is pulled to the right (stretching the spring), to which the spring responds with a force to the left (again toward the equilibrium point). Right: A marble in a bowl. At the center of the bowl the marble is in equilibrium - no net force acts on it. Here the restoring force is provided by gravity. As soon as the marble is displaced from the center, gravity applies a force back toward the center. When let go, the resulting motion of the marble is simple harmonic motion.

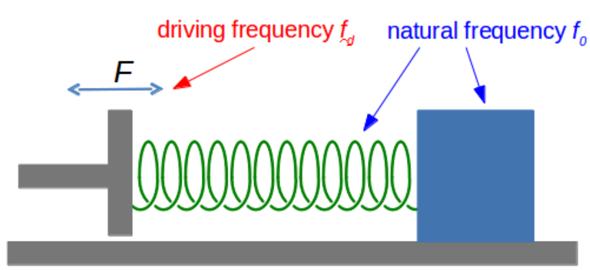


Figure 3.7: Mass on a spring driven by a periodic force F. When the piston moves back and forth, the connecting spring causes the mass to oscillate as well. The amplitude of the mass will depend on how the driving force frequency f_d compares with the natural oscillation frequency of the mass-spring system f_0.

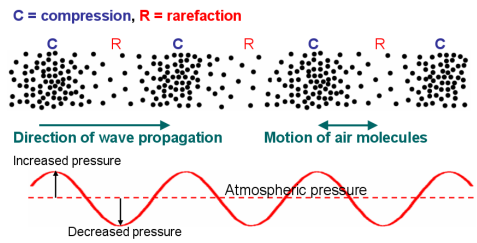


Figure 4.7: The medium for sound is air. Waves are longitudinal, meaning that the molecules of air move back and forth in the direction of the wave propagation. The physical distance between neighboring compression zones is called the wavelength (see section 4.1.4). The above is a graphical representation of the pressure variations in the medium corresponding to the moment the wave is depicted. Compression regions are *above* atmospheric pressure, and rarefaction regions are *below* atmospheric pressure.

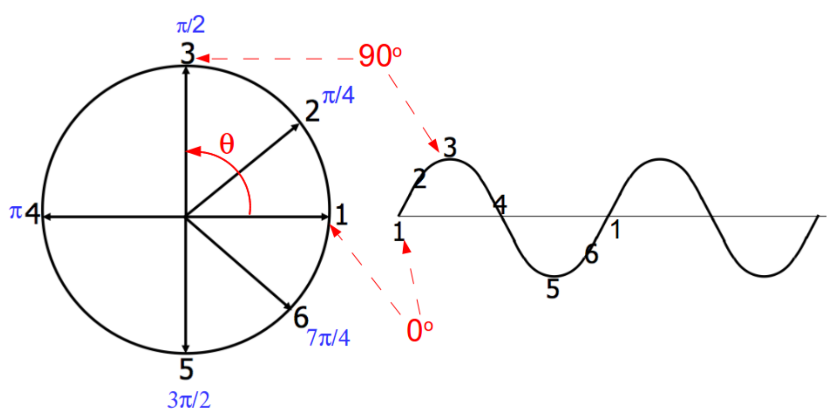


Figure 6.4: The unit circle with points identified on its perimeter corresponding to certain phases. The corresponding points are shown on the sine curve to indicate phase and its relationship to points along the cycle of an oscillatory system. The circumference of the circle is $2\pi r$, and once around corresponds to a total angular journey of 2π radians = 360°.

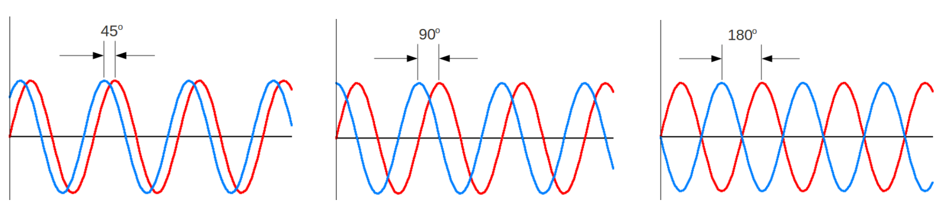


Figure 6.5: Three situations where two identical oscillating systems differ from one another by a net phase.

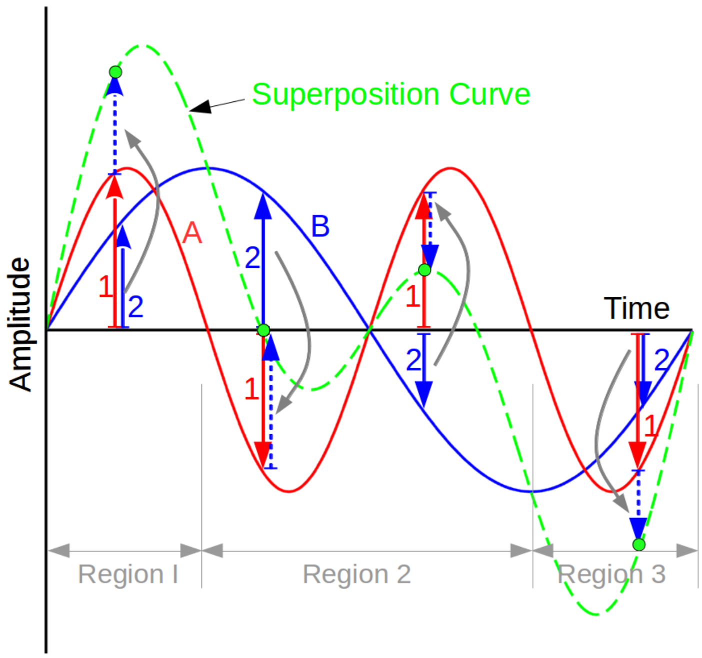


Figure 6.6: Example of adding two curves by superposition. Arrows show the graphical sum of the two waves in four sample time locations. Arrows for the amplitudes of curves A and B are added head-to-tail to give the result for the dotted superposition curve amplitude. In Region 1 both curves A and B have positive amplitude so that the superposition is above both. In Region 2 the curves A and B have opposite sign amplitude so that the superposition curve is between the two. In Region 3 both A and B have negative amplitude meaning that the superposition curve is below both.

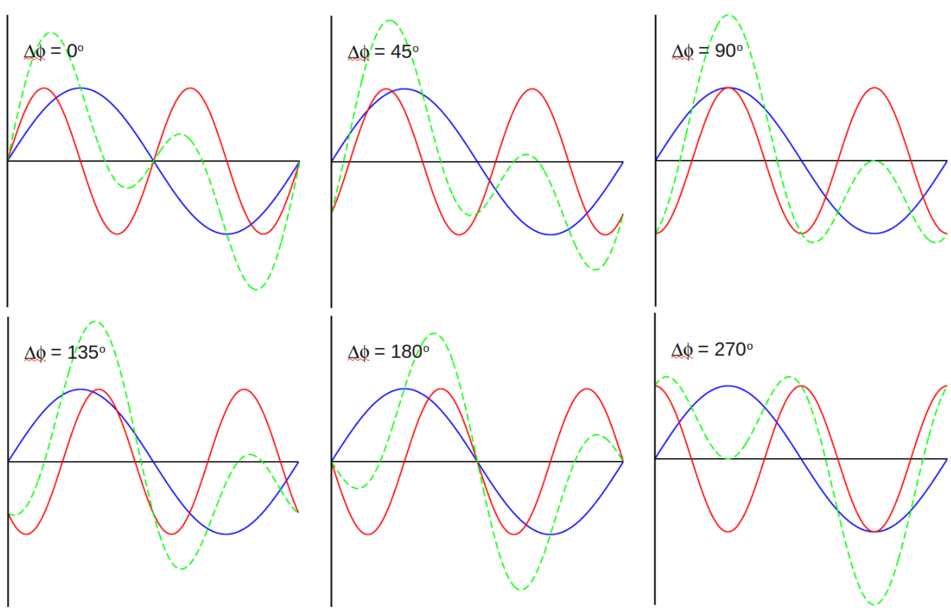


Figure 6.7: Superposition of two pure tones at various relative phases between the two solid curves. The phases $\Delta\phi$ marked on each pane correspond to the phase difference between the two curves at the origin. Note how the superposition curve changes shape as the relative phase between the two changes.

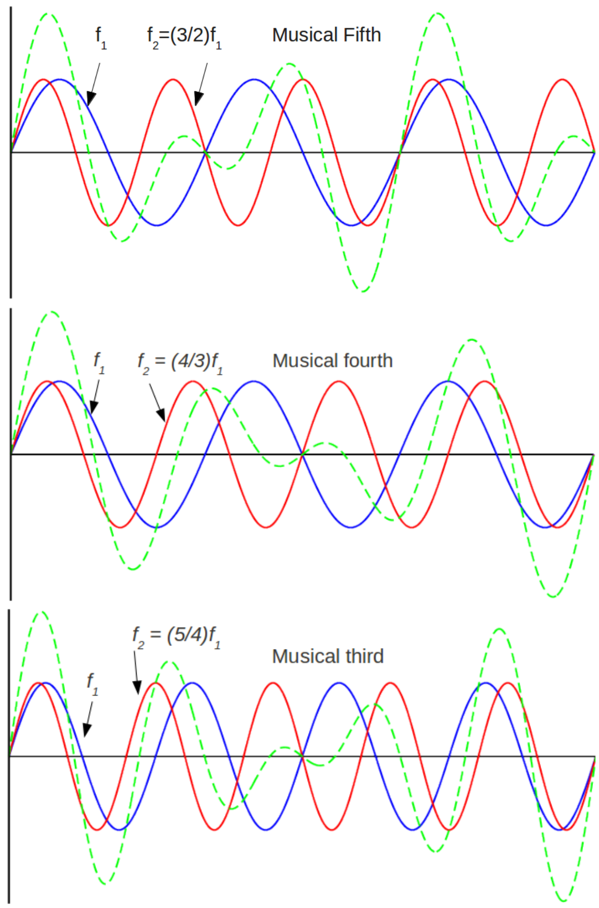


Figure 6.8: Superposition curves for three musical intervals, the fifth, fourth, and third.

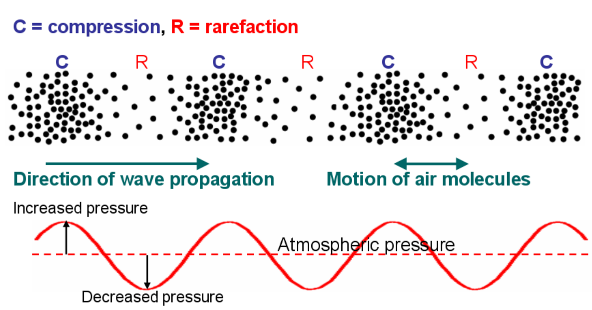


Figure 6.13: A "picture" of a sound wave frozen in time. The physical distance between neighboring compression zones corresponds to the wavelength. At the bottom is a graphical representation of the pressure variations in the medium, which has a similar shape to the vibrational pattern in the vibrational source of the wave.

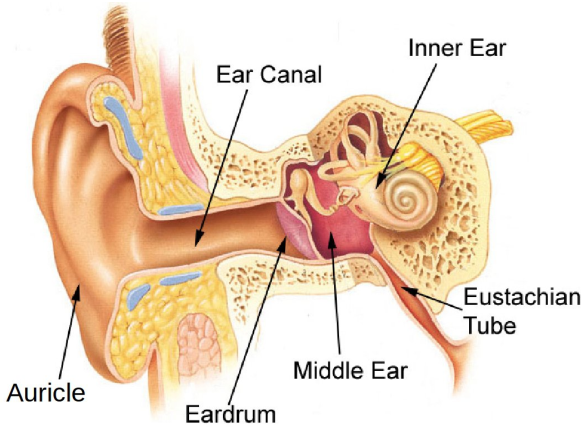


Figure 7.1: Schematic of the human ear, consisting of the outer ear (auricle and ear canal), the middle ear (consisting of the chain of three bones, or ossicles, the opening of the Eustachian tube), and the inner ear (consisting of the cochlea, which passes information on to the brain via the auditory nerve, and the semicircular canals, used for maintaining balance).[1]

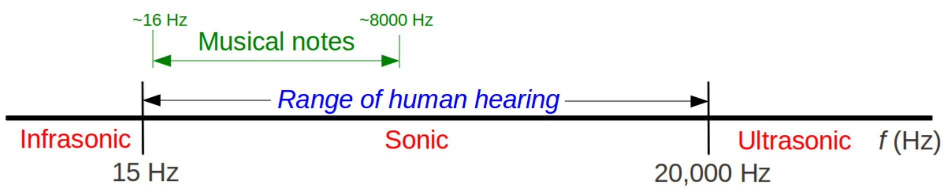


Figure 7.4: The range of human hearing extends from about 15 to 20,000 Hz. The range of musical notes produced by instruments extends to as high as around 8000 Hz, but harmonics up to 20,000 Hz are very important for timbre and fidelity.

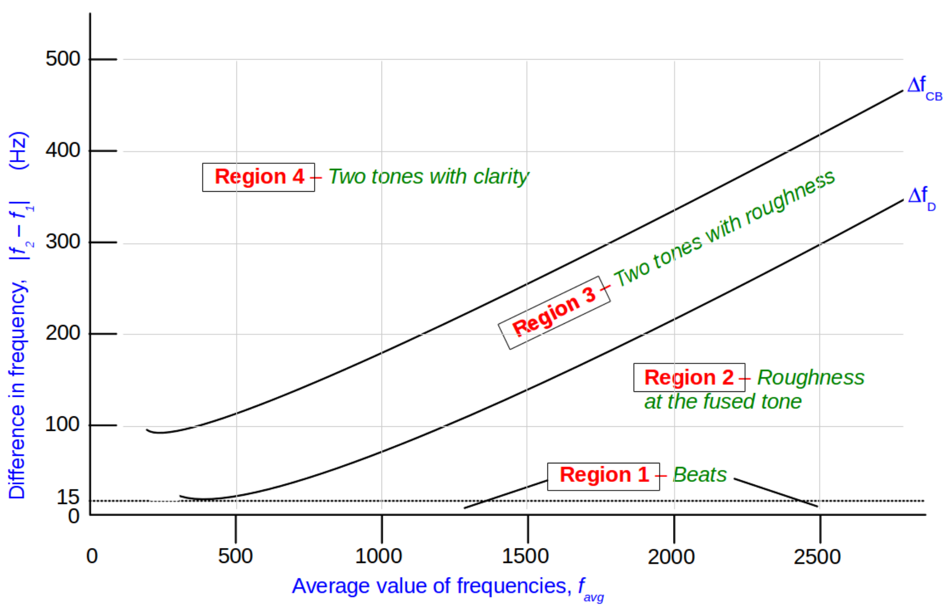


Figure 8.1: Critical band plot, summarizing the auditory response from perception of two simultaneous pure tones.

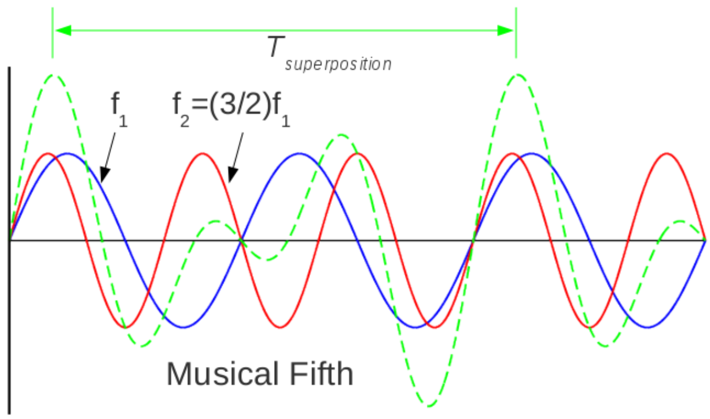


Figure 8.2: Superposition of two tones constituting the interval of the musical fifth. The periodicity of the superposition curve is longer than both component curves, having period twice that of f_1.

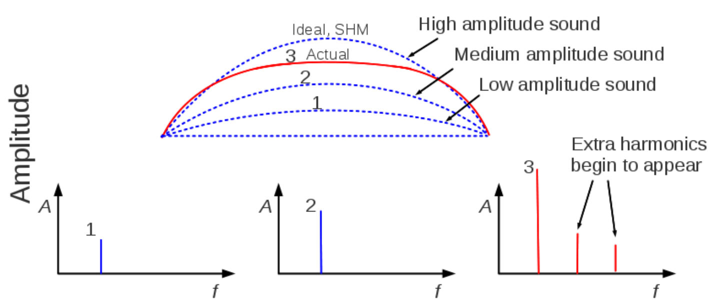


Figure 8.4: Illustration of the eardrum shape when being driven into oscillation by low (1), medium (2), and high (3) amplitude pure tones. When driven beyond its elastic limit (3), its shape becomes distorted from sinusoidal, and as a result extra harmonics begin to appear in the sound spectrum which are passed on to ossicles of the middle ear. These extra harmonics are not part of the original tone driving the eardrum into oscillation, but are rather added by the eardrum as a result of being stretched beyond its elastic limit.

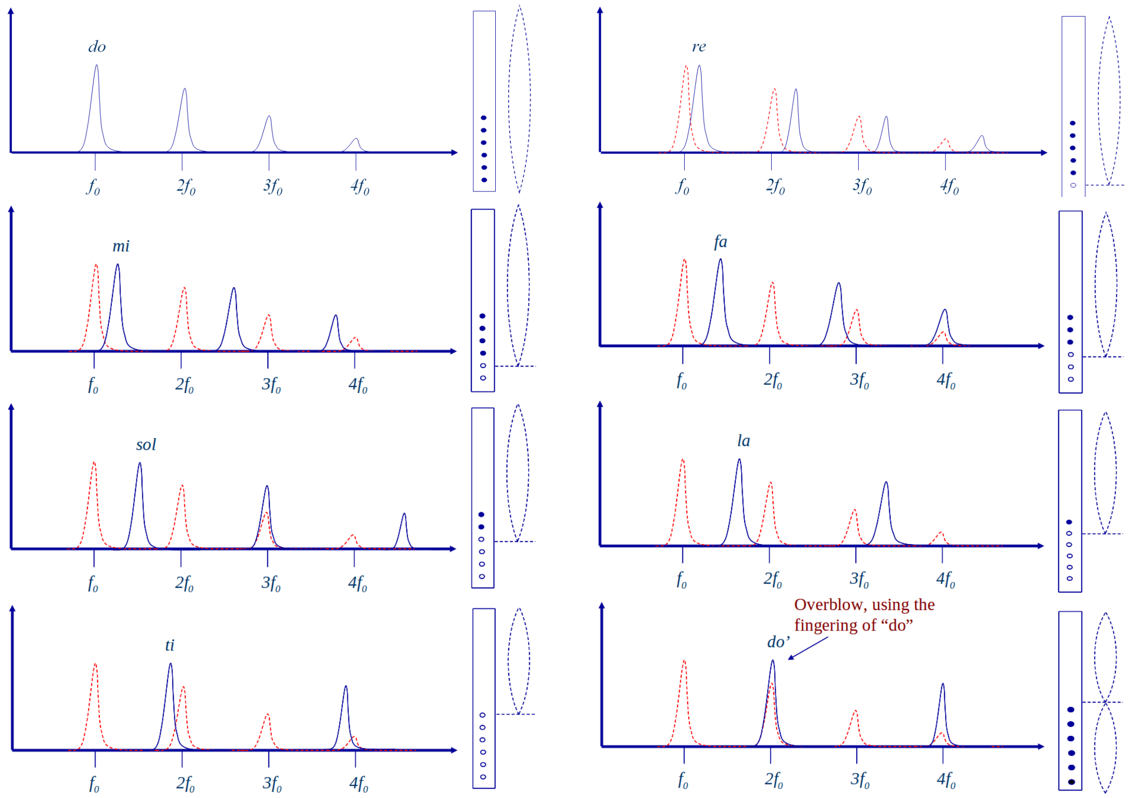


Figure 10.4: Resonances in the flute as a function of key hole coverage. In each of the eight panes is shown a sound spectrum on the left, which includes the fundamental and several harmonics, and a schematic of the flute with open and/or closed holes on the right, with a representative standing wave showing the resonance for that particular closed hole configuration.

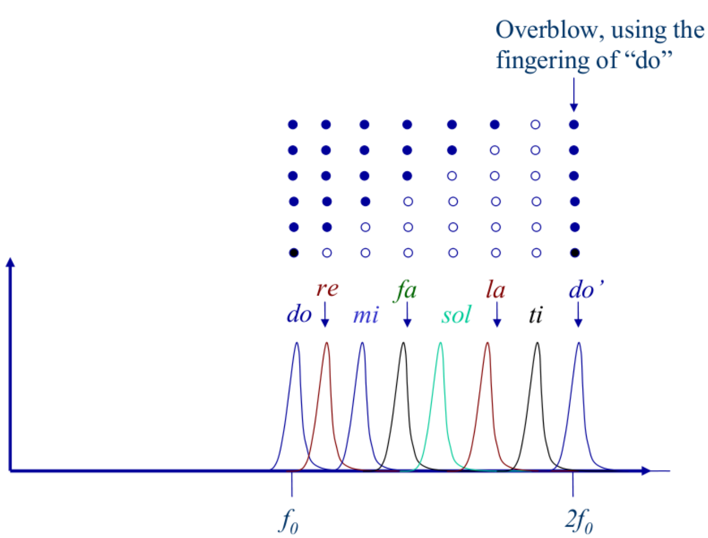


Figure 10.5: The notes and key hole closings for the flute, starting from all holes closed and back again with use of the register key

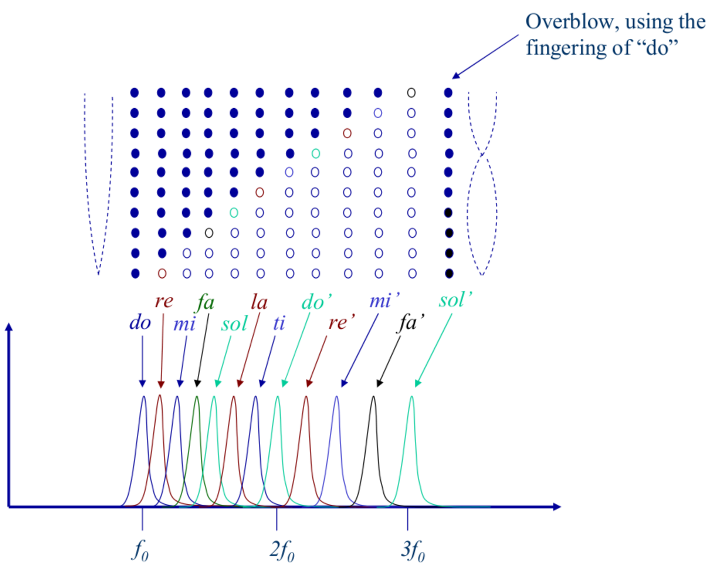


Figure 10.6: The notes and key hole closings for the clarinet, starting from all holes closed and back again with use of the register key.

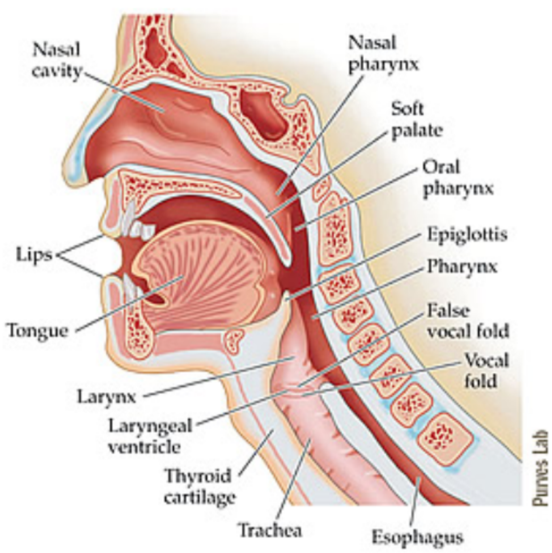


Figure 13.1: The anatomy of the human vocal system. Air supplied by the lungs is forced up the trachea, and the vocal fold opens and closes at high frequency to modulate the moving air into sound waves. The resonant character of the vocal tract (consisting of the larynx, pharynx, mouth, and nasal cavity) determines the specific timbre of the spoken sound. Changes in the shape of the vocal tract cause changes in the resonant character and hence in the timbre of the spoken sound.

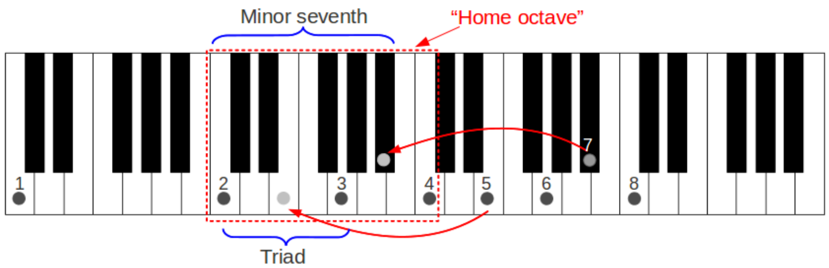


Figure 14.3: Within the harmonic structure of the sound from a tonal instrument are already several notes that are part of the Western diatonic scale. Some higher notes (the E and the B$^\flat$) are moved down by an octave to place them in a "home octave" and identify them as members of the diatonic scale. Also present within the first 8 harmonics are the notes C, E, G, and B$^\flat$.

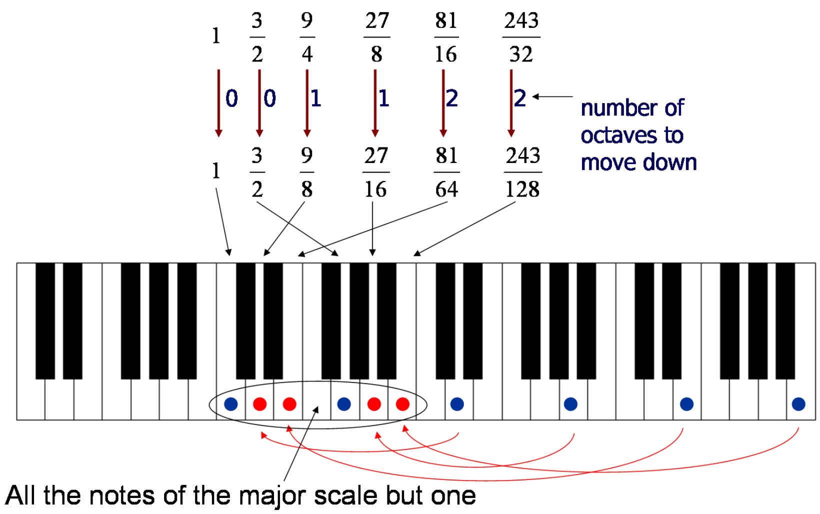


Figure 14.10: Moving the acquired notes into a common octave, we have 6 out of the 7 required notes: C, D, E, G, A, and B.

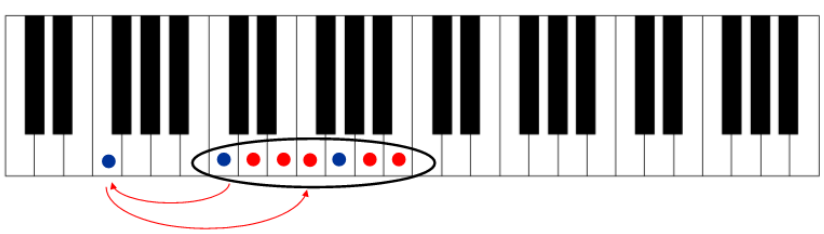


Figure 14.11: Moving down one fifth and then up one octave identifies the last note of the major diatonic scale in the Pythagorean temperament.

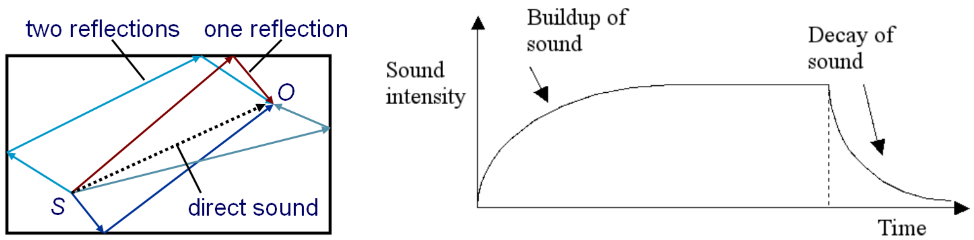


Figure 15.8: Sound reaching an observer consists of direct plus reflected sound. The buildup of sound results from the slight delay between the direct and reflected sound.

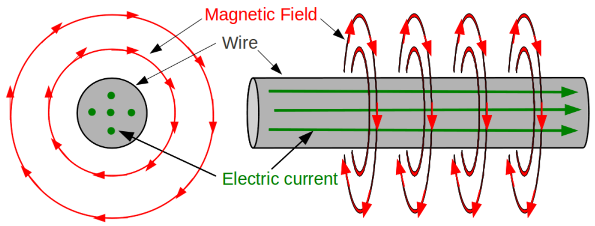


Figure 16.1: The magnetic field associated with moving charges in a wire. Note that the field is "circumferential," meaning that it encircles the wire with a direction consistent with the "right hand rule."

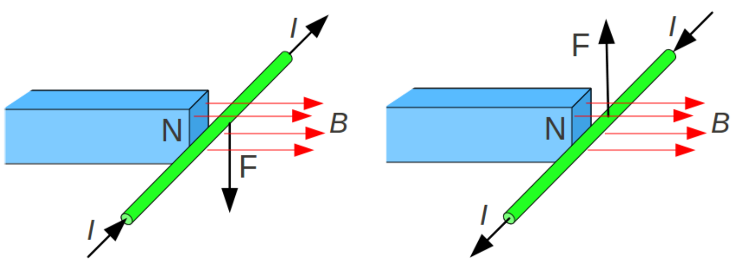


Figure 16.2: The force exerted by an external magnetic field on a current-carrying wire. The direction of the force depends on the direction of the magnetic field and the current, and can be found by the right-hand-rule. If the fingers of the right hand are pointed along the current I, and then curl toward the direction of the magnet field B, the thumb points in the direction of the induced force.

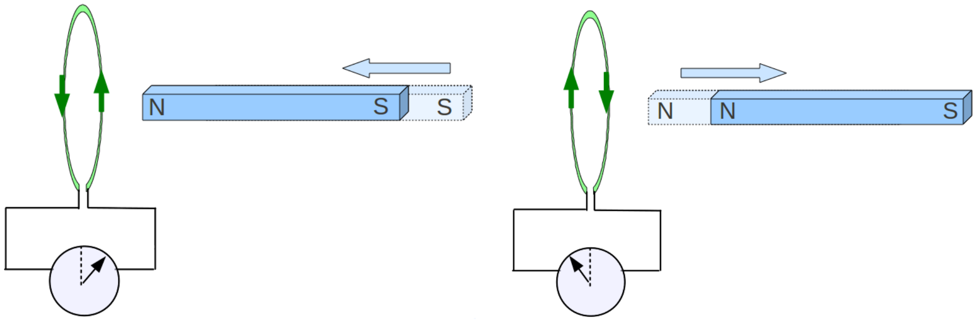


Figure 16.3: A magnet moving in the vicinity of a wire loop generates current in the loop as long as the magnet is moving. Once it stops, the current ceases. The current changes direction when the magnet is moved in the opposite direction. The faster the magnet or the coil moves, the higher the current generated.

Figure 16.4: A vinyl LP record, a turntable, and a closeup view of the cartridge and stylus sitting on the surface of the rotating record.[1][2][3]

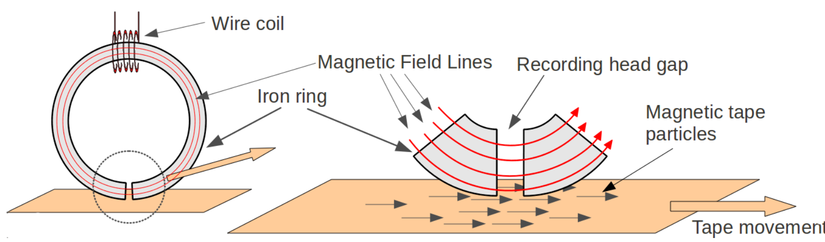


Figure 16.13: The tape head, capable of both recording and playback. When recording, variable currents in the wire coil produce variable magnetic fields in the ring. As the tape moves under the head, the fluctuating magnetic field aligns the magnetic particles in the tape. During playback, the magnetic particles in the tape create a magnetic field in the ring. This fluctuating field induces currents in the coil that can be amplified and sent to the speakers.

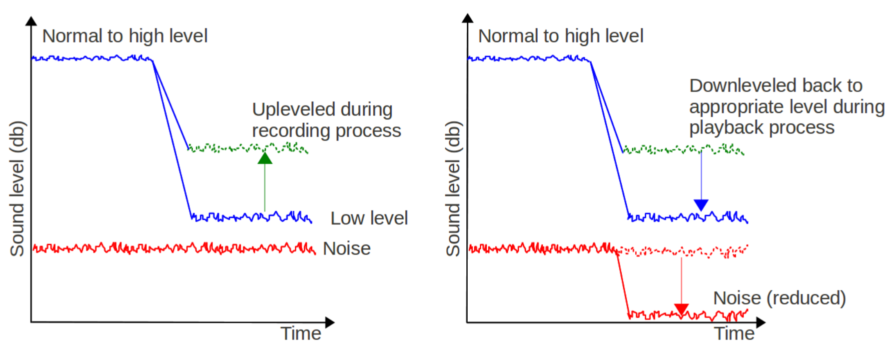


Figure 16.15: The effects of Dolby noise reduction technology. Low intensity level sounds are up-leveled during the recording process, and down-leveled in the playback process. The net effect of this is to reduce the noise by the same down-level amount during playback.

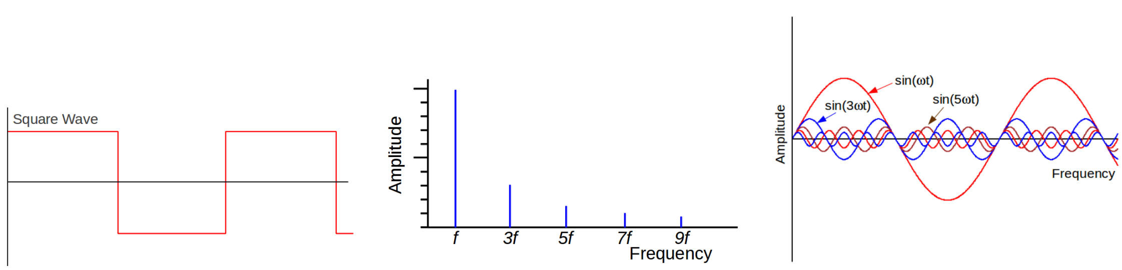


Figure 17.2: The ideal square wave, the first five components of its frequency spectrum, and the sinusoidal waves corresponding to these five components, of differing frequency and amplitudes.

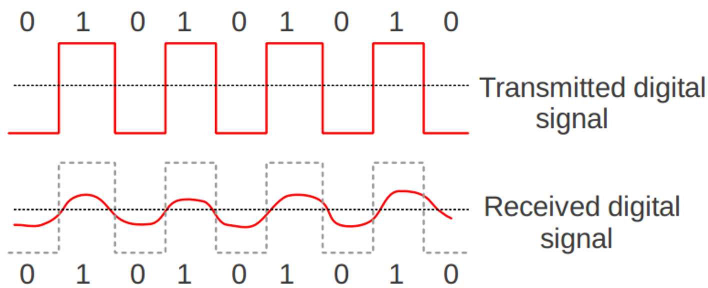


Figure 18.1: A transmitted and received digital signal. The original 1's and 0's of the transmitted signal are greatly degraded in shape by the time they reach the receiver, but since the encoding of the waveform is in the peaks and valleys of the signal, the receiver can still interpret the difference between the 1's and 0's above and below the dotted line, even though the signal is greatly rounded and reduced in amplitude.

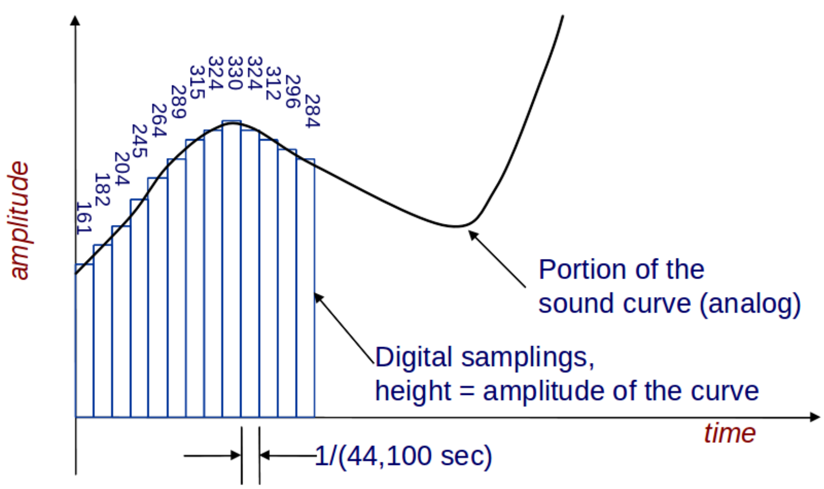


Figure 18.2: Digital encoding of an analog waveform. Each 44,100th of a second, the waveform amplitude is converted into a number. The stream of numbers then represents an approximation of the waveform.