Chapter 16

Fourier Analysis, Normal Modes and Sound

In Chapter 15 we discussed the principle of superposition—the idea that waves add, producing a composite wave that is the sum of the component waves. As a result, quite complex wave structures can be built from relatively simple wave forms. In this chapter our focus will be on the analysis of complex wave forms, finding ways to determine what simple waves went into constructing a complex wave.

As an example of this process, consider what happens when white sunlight passes through a prism. White light is a mixture of all the colors, all the wavelengths of the visible spectrum. When white light passes through a prism, a rainbow of colors appears on the other side. The prism separates the individual wavelengths so that we can study the composition of the white light. If you look carefully at the spectrum of sunlight, you will observe certain dark lines; some very specific wavelengths of light are missing in light from the sun. These wavelengths were absorbed by elements in the outer atmosphere of the sun. By noticing what wavelengths are missing, one can determine what chemical elements are in the sun's atmosphere. This is how the element Helium (named after Helios, Greek for sun) was discovered.

This example demonstrates how the ability to separate a complex waveform (in this case white sunlight) into its component wavelengths or frequencies, can be a powerful research tool. The sounds we hear, like those produced by an orchestra, are also a complex mixture of waves. Even individual instruments produce complex wave forms. Our ears are very sensitive to these wave forms. We can distinguish between a note played on a Stradivarius violin and the same note played by the same person on an inexpensive violin. The only difference between the two sounds is a slight difference in the mixture of the component waves, the harmonics present in the sound. You could not tell which was the better violin by looking at the waveform on an oscilloscope, but your ear can easily tell.

You can hear these subtle differences because the ear is designed in such a way that it separates the complex incoming sound wave into its component frequencies. The information your brain receives is not what the shape of the complex sound wave is, but how much of each component wave is present. In effect, your ear is acting like a prism for sound waves.

When we study sound in the laboratory, the usual technique is to record the sound wave amplitude using a microphone, and display the resulting waveform on an oscilloscope or computer screen. If you want to, you can generate more or less pure tones that look like sine waves on the screen. Whistling is one of the best ways to do this. But if you record the sound of almost any instrument, you will not get a sine wave shape. The sound from virtually all instruments is some mixture of different frequency waves. To understand the subtle differences in the quality of sound of different instruments, and to begin to understand why these differences occur, you need to be able to decompose the complex waves you see on the oscilloscope screen into the individual component waves. You need something like an ear or a prism for these waves.

A way to analyze complex waveforms was discovered by the French mathematician and physicist Jean Baptiste Fourier, who lived from 1786 to 1830. Fourier was studying the way heat was transmitted through solids and in the process discovered a remarkable mathematical result. He discovered that any continuous, repetitive wave shape could be built up out of harmonic sine waves. His discovery included a mathematical technique for determining how much of each harmonic was present in any given repetitive wave. This decomposition of an arbitrary repetitive wave shape into its component harmonics is known as **Fourier analysis**. *We can think of Fourier analysis effectively serving as a mathematical prism.*

The techniques of Fourier analysis are not difficult to understand. Appendix A of this chapter is a lecture on Fourier analysis developed for high school students with no calculus background (explicitly for my daughter's high school physics class). To apply Fourier analysis you have to be able to determine the area under a curve, a process known in calculus as integration. While the idea of measuring the area under a curve is not a difficult concept to grasp, the actual process of doing this, particularly for complex wave shapes, can be difficult. Everyone who takes a calculus course knows that integration can be hard. The integrals involved in Fourier analysis, particularly the analysis of experimental data are much too hard to do by hand or by analytical means.

People find integration hard to do, but computers don't. With a computer one can integrate any experimental wave shape accurately and rapidly. As a result, Fourier analysis using a computer is very easy to do. A particularly fast way of doing Fourier analysis on the computer was discovered by Cooley and Tukey in the 1950s. Their computer technique or algorithm is known as the **Fast Fourier Transform** *or* **FFT** *for short. This algorithm is so commonly used that one often refers to a Fourier transform as an FFT.*

The ability to analyze data using a computer, to do things like Fourier analysis, has become such an important part of experimental work that older techniques of acquiring data with devices like strip chart recorders and stand alone oscilloscopes have become obsolete. With modern computer interfacing techniques, important data is best recorded in a computer for display and analysis. We have developed the **MacScope™** *program, which will be used often in this text, for recording and displaying experimental data. The main reason for writing the program was to make it a simple and intuitive process to apply Fourier analysis. In this chapter you will be shown how to use this program. With the computer doing all the work of the analysis, it is not necessary to know the mathematical processes behind the analysis, the steps are discussed in the appendix. But a quick reading of the appendix should give you a feeling for how the process works.*

HARMONIC SERIES

We begin our discussion with a review of the standing waves on a guitar string, shown in Figure (15-15) reproduced here. We saw that the wavelengths λ_n of the allowed standing waves are given by the formula

$$\lambda_n = \frac{2L}{n} \qquad (15\text{-}37)$$

Where L is the length of the string, and n takes on integer values n = 1, 2, 3, ...

Each of these standing waves has a definite frequency of oscillation that was given in Equation 15-38 as

$$\begin{aligned} f_n \frac{cycles}{sec} &= \frac{v}{\lambda_n} \frac{meters/sec}{meters/cycle} \\ &= \frac{v}{\lambda_n} \frac{cycles}{sec} \end{aligned} \qquad (15\text{-}38)$$

If we substitute the value of λ_n from Equation 15-37 into Equation 15-38, we get as the formula for the corresponding frequency of vibration

$$f_n = \frac{v_{wave}}{\lambda_n} = \frac{v_{wave}}{2L/n} = n\left(\frac{v_{wave}}{2L}\right) \qquad (1)$$

For n = 1, we get

$$f_1 = \frac{v_{wave}}{2L} \qquad (2)$$

All the other frequencies are given by

$$\boxed{f_n = nf_1} \quad \textit{harmonic series} \qquad (3)$$

This set of frequencies is called a ***harmonic series***. The ***fundamental frequency*** or ***first harmonic*** is the frequency f_1. The ***second harmonic*** f_2 has twice the frequency of the first. The third harmonic f_3 has a frequency three times that of the first, etc. Note also that the fundamental has the longest wavelength, the second harmonic has half the wavelength of the fundamental, the third harmonic one third the wavelength, etc.

It was Fourier's discovery that any continuous repetitive wave could be built up by adding together waves from a harmonic series. The correct harmonic series is the one where the fundamental wavelength λ_1 is equal to the period over which the waveform repeats.

To begin our discussion of Fourier analysis and the building up of waveforms from a harmonic series, we will first study the motion of two air carts connected by springs and riding on an air track. We will see that these ***coupled air carts*** have several distinct ***modes of motion***. Two of the modes of motion are purely sinusoidal, with precise frequencies. But any other kind of motion appears quite complex.

However, when we record the complex motion, we discover that the velocity of either cart is repetitive. A graph of the velocity as a function of time gives us a continuous repetitive wave. According to Fourier's theorem, this waveform can be built up from sinusoidal waves of the harmonic series whose fundamental frequency is equal to the repetition frequency of the wave. When we use Fourier analysis to see what harmonics are involved in the motion, we will see that the apparent complex motion of the carts is not so complex after all.

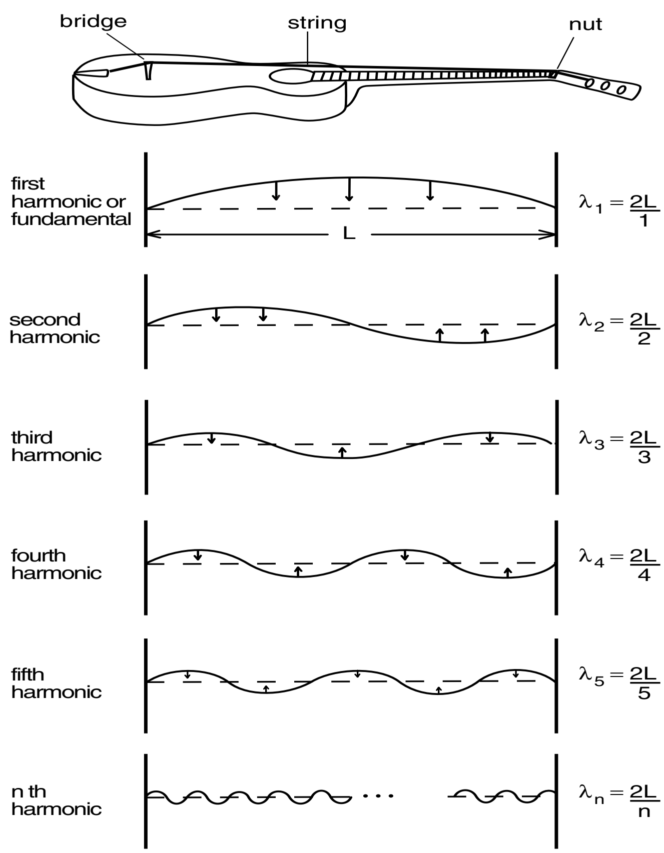


Figure 15-15
Standing waves on a guitar string.

NORMAL MODES OF OSCILLATION

The reason musical instruments generally produce complex sound waves containing various frequency components is that the instrument has various ways to undergo a resonant oscillation. Which resonant oscillations are excited with what amplitudes depends upon how the instrument is played.

In Chapter 14 we studied the resonant oscillation of a mass suspended from a spring, or equivalently, of a cart on an air track with springs attached to the end of the cart, as shown in Figure (1). This turns out to be a very simple system—there is only one resonant frequency, given by $\omega = \sqrt{k/m}$. The only natural motion of the mass is purely sinusoidal at the resonant frequency. This system does not have the complexity found in most musical instruments.

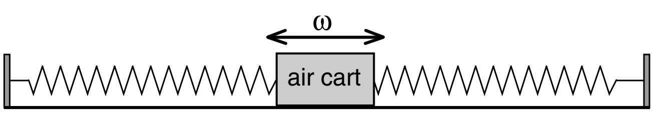


Figure 1
Cart and springs on an air track.

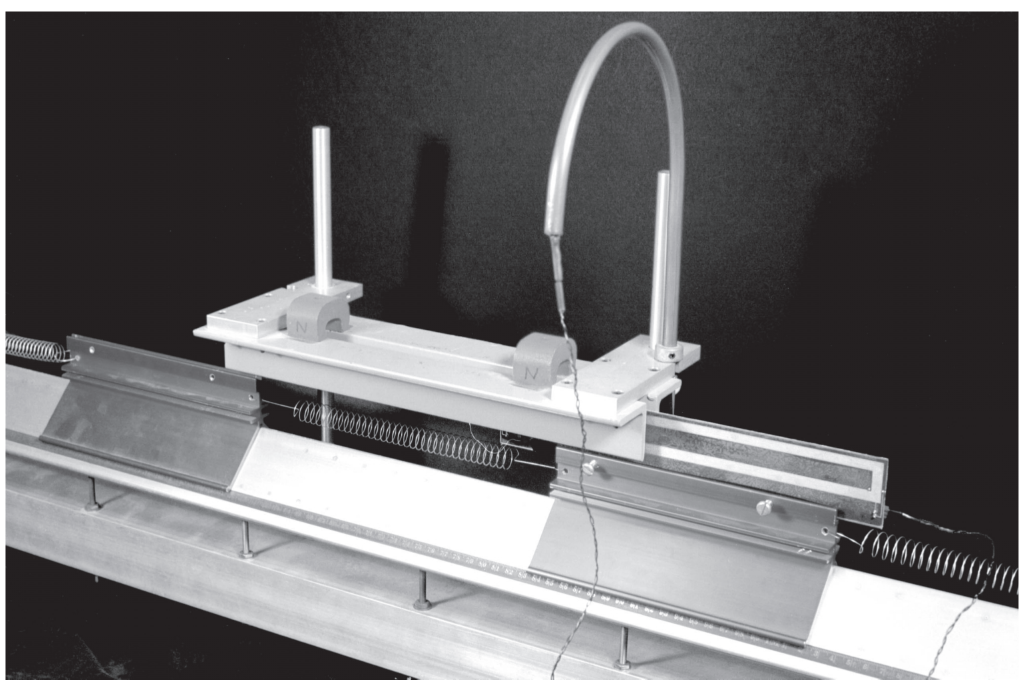
Figure2
System of coupled air carts.

Things become more interesting if we place two carts on the air track, connected by springs as shown in Figure (2). We will call this a system of two ***coupled*** air carts.

When we analyzed the one cart system, we found that the force on the cart was simply $F = -kx$, where x is the displacement of the cart from its equilibrium position. With two carts, the force on one cart depends not only on the position of that cart, but also on how far away the other cart is. A full analysis of this coupled cart system, using Newton's second law, leads to a pair of coupled differential equations whose solution involves matrices and eigenvalues. In this text we do not want to get into that particular branch of mathematics. Instead we will study the motion of the carts experimentally, and find that the motion, which at first appears complex, can be explained in simple terms.

In order to record the motion of the aircarts, we have mounted the velocity detector apparatus shown in Figure (3). The apparatus consists of a 10 turn wire coil mounted on top of the cart, that moves through the magnetic field of the iron bars suspended above the coil. The operation of the velocity detector apparatus depends upon Faraday's law of induction which will be discussed in detail in Chapter 30 on Faraday's law. For now all we need to know is that a voltage is induced in the wire coil, a voltage whose magnitude is proportional to the velocity of the cart. This voltage signal

Figure 3
Recording the velocity of one of the aircarts. A 10 turn coil is mounted on top of one of the carts. The coil moves through the magnetic field between the angle irons, and produces a voltage proportional to the velocity of the cart. This voltage is then recorded by the Macintosh oscilloscope.



from the wire coil is carried by a cable to the Macintosh oscilloscope where it is displayed on the computer screen.

When you first start observing the motion of the coupled aircarts, it appears chaotic. One cart will stop and reverse direction while the other is moving toward or away from it, and there is no obvious pattern. But after a while, you may discover a simple pattern. If you pull both carts apart and let go in just the right way, the carts come together and go apart as if one cart were the mirror image of the other. This motion of the carts is illustrated in Figure (4a). We will call this the ***vibrational mode*** of motion.

In Figure (4b) we have used the velocity detector to record the motion of one of the coupled air carts when the carts are moving in the vibrational mode. You can see that the curve closely resembles a sine wave.

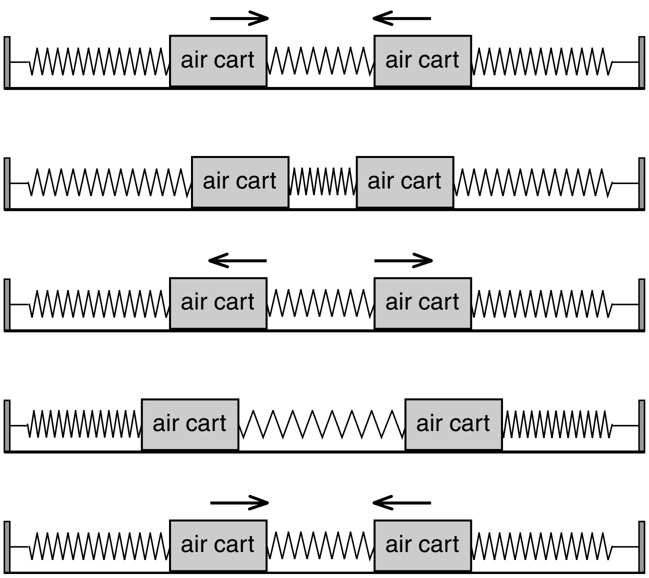


Figure4a
Vibrational motion of the coupled air carts.

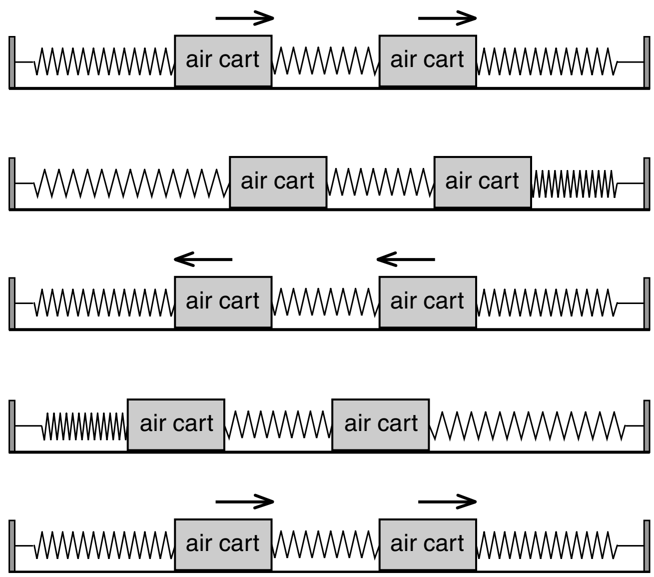


Figure 5a
Sloshing mode of motion of the coupled aircarts.

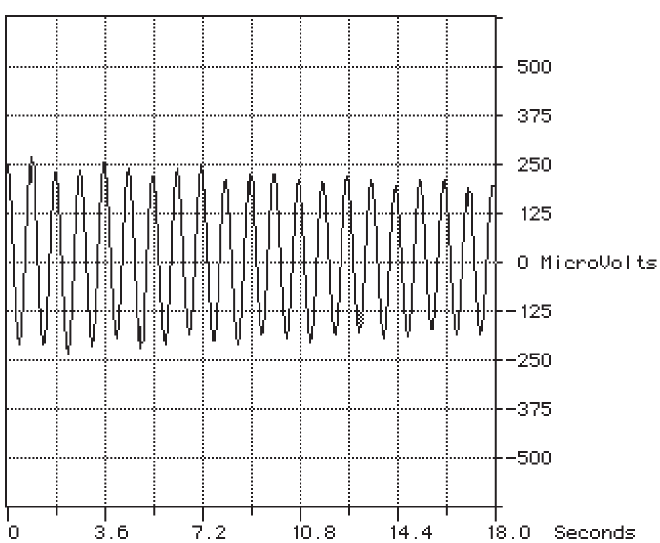


Figure 4b
Vibrational mode of oscillation of the coupled aircarts. The voltage signal is proportional to the velocity of the cart that has the coil on top.

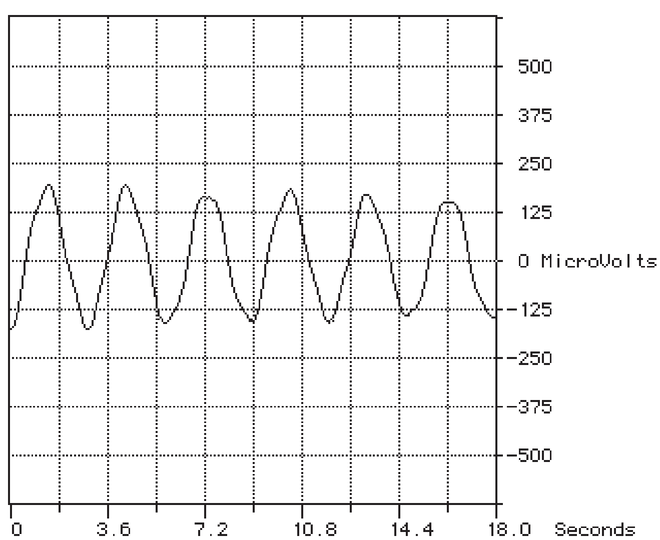


Figure 5b
A pure sloshing mode is harder to get. Here we came close, but it is not quite a pure sine wave.

If you play around with the carts for a while longer, you will discover another way to get a simple sinusoidal motion. If you pull both carts to one side, get the positions just right, and let go, the carts will move back and forth together as illustrated in Figure (5a). We will call this the ***sloshing mode*** of motion of the coupled air carts. In Figure (5b) we have recorded the velocity of one of the carts in the sloshing mode, and see that the curve is almost sinusoidal.

In general, the motion of the two carts is not sinusoidal. For example, if you pull one cart back and let go, you get a velocity curve like that shown in Figure (6). If you start the carts moving in slightly different ways you get differently shaped curves like the one seen in Figure (7). Only the vibrational and sloshing modes result in sinusoidal motion, all other motions are more complex. To study the complex motion of the carts, we will use the techniques of ***Fourier analysis***.

FOURIER ANALYSIS

As we mentioned in the introduction, Fourier analysis is essentially a mathematical prism that allows us to decompose a complex waveform into its constituent pure frequencies, much as a prism separates sunlight into beams of pure color or wavelength. We have just studied the motion of coupled air carts, which gave us an explicit example of a relatively complex waveform to analyze. While the two carts can oscillate with simple sinusoidal motion in the vibrational and sloshing modes of Figures (4) and (5), in general we get complex patterns like those in Figures (6) and (7). What we will see is that, by using Fourier analysis, the waveforms in Figures (6) and (7) are not so complex after all.

The MacScope program was designed to make it easy to perform Fourier analysis on experimental data. The MacScope tutorial gives you considerable practice using MacScope for Fourier analysis. What we will do here is discuss a few examples to see how the program, and how Fourier analysis works. We will then apply Fourier analysis to the curves of Figures (6) and (7) to see what we can learn. But first we will see how MacScope handles the analysis of more standard curves like a sine wave or square wave.

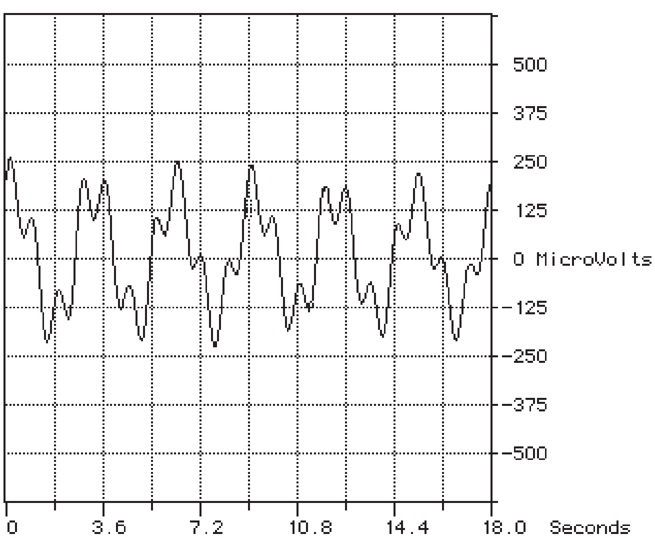


Figure 6
Complex motion of the coupled air carts.

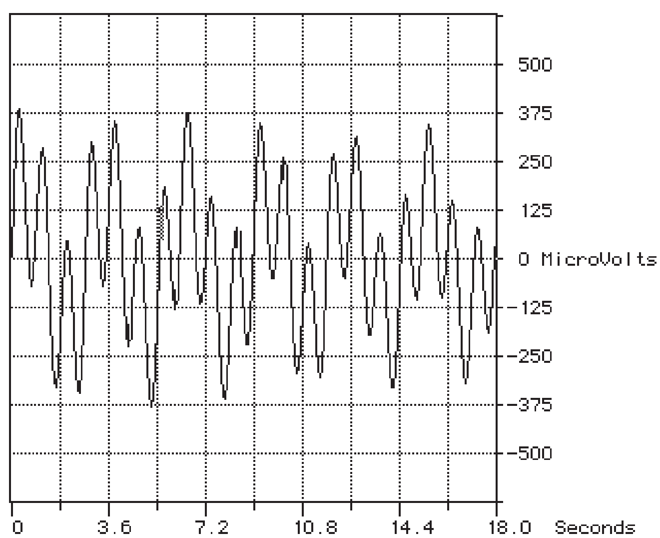


Figure 7
Another example of the complex motion of the coupled air carts.

Analysis of a Sine Wave

In Figure (8a) we attached MacScope to a sine wave generator and recorded the resulting waveform. The sine wave generator, which is usually called a ***signal generator***, is an electronic device that we will use extensively in laboratory work on the electricity part of the course. Typically the device has a dial that allows you to select the frequency of the wave, a knob that allows you to adjust the wave amplitude, and some buttons by which you can select the shape of the wave. Typically you can choose between a sine wave shape, , a square wave shape and a triangular wave .

In Figure (8a) we have selected one cycle of the wave and see that the frequency of the wave is 1515 Hz (1.515 KHz), which is about where we set the frequency dial at on the signal generator. To get Figure (8b) we pressed the ***Expand*** button that appears once a section of curve has been selected. This causes the selected section of the curve to fill the whole display rectangle.

The selection rectangle is obtained by holding down the mouse bottom and dragging across the desired section of the curve. The starting point of the selection rectangle can be moved by holding down the shift key while moving the mouse. The data box shows the period **T** *of the selected rectangle, and the corresponding frequency* **f**. *If you wish the data box to remain after the selection is made, hold down the option key when you release the mouse button. This immediately gives you the* **ImageGrabber™**, *which allows you to select any section of the screen to save as a PICT file for use in a report or publication.*

In Figure (9), we went up to the **Analyze** menu of MacScope and selected ***Fourier Analysis***. As a result we get the window shown in Figure (10). At the top we see the selected one cycle of a sine wave. Beneath, we see a rectangle with one vertical bar and a scale labeled ***Harmonics***. The vertical bar is in the first position, indicating that the section of the wave which we selected has only a first harmonic. Beneath the ex-

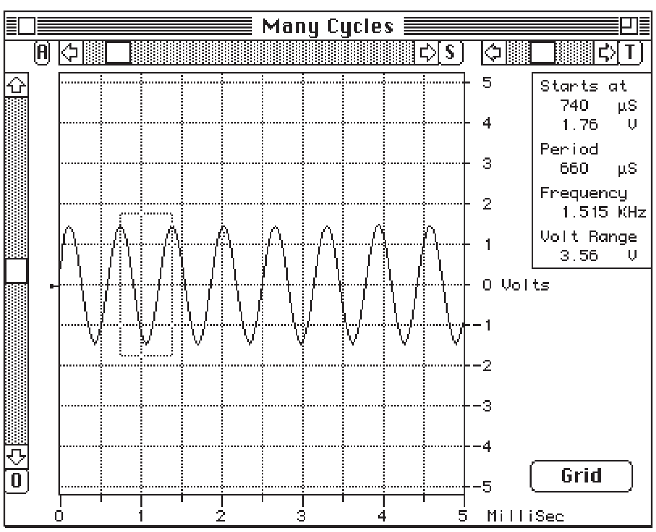


Figure 8a
Sine wave from a signal generator. Selecting one cycle of the wave, we see that the frequency of the wave is about 1.5 kilocycles (1.515 KHz). (Here we also see some of the controls that allow you to study the experimental data. The scrollbar labeled "S" lets you move the curve sideways. The "T" scrollbar changes the time scale, and the "O" scrollbar moves the curve up and down. In the text, we will usually not show controls unless they are important to the discussion.)

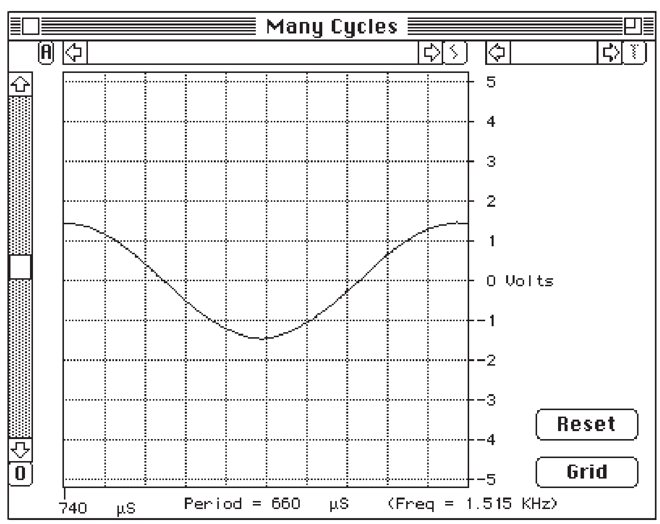
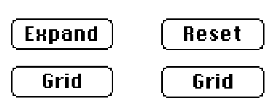


Figure 8b
Once we have selected a section of curve, a new control labeled* Expand *appears. When we press the* Expand *button, the selected section of the curve fills the entire display rectangle as seen above. (The control then becomes a* Reset *button which takes us back to the full curve.)



panded curve, there is a printout of the period and frequency of the selected section. Here we see again that the frequency is 1.515 KHz.

For Figure (11), we pressed **Reset** to see the full curve, selected 4 cycles of the sine wave, and expanded that. In the lower rectangle we now see one vertical bar over the 4th position, indicating that for the selected wave we have a pure fourth harmonic. In the MacScope tutorial we give you a MacScope data file for a section of a sine wave. Working with this data file, you should find that if you select one cycle of the wave, anywhere along the wave, you get an indication of a first harmonic. Select two cycles, and you get an indication of a pure second harmonic, etc.

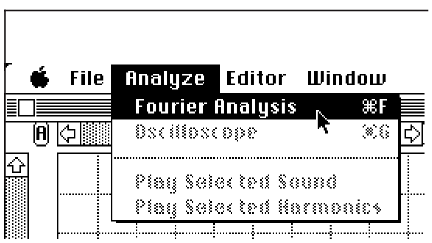


Figure 9
Choosing Fourier Analysis.

The basic function of the Fourier analysis program in MacScope is to determine how you can construct the selected section of the wave from ***harmonic sine waves***. The first harmonic has the frequency of the selected section. For example, suppose that you had a 10 cycle per second sine wave and selected one cycle. That selection would have a frequency of 10 Hz and a period of 0.1 seconds. The second harmonic would have a frequency of 20 Hz, and the third harmonic 30 Hz, etc. The nth harmonic frequency is n times greater than the first harmonic. Fourier's discovery was that *any continuous repeating wave form can be constructed from the harmonic sine waves*. So far we have considered only the obvious examples of sections of a pure sine wave. We will now go on to more complex examples to see how a waveform can be constructed by adding up the various harmonics.

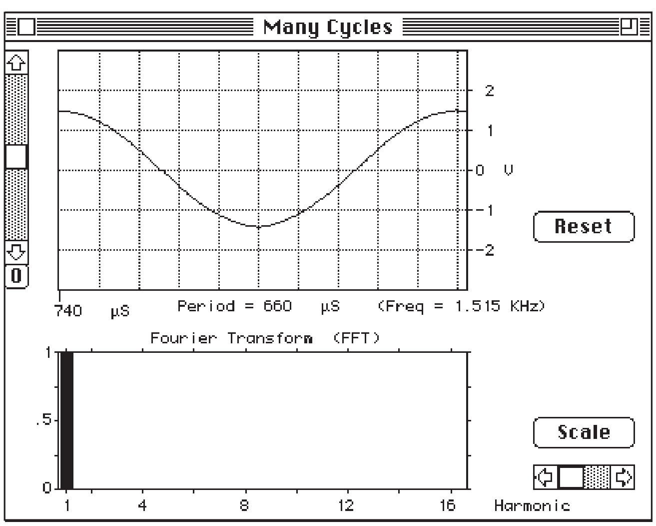


Figure 10
In the MacScope program, **Fourier Analysis** ***acts only on the selected section of the curve. Since we selected only one cycle, we see only a first harmonic.***

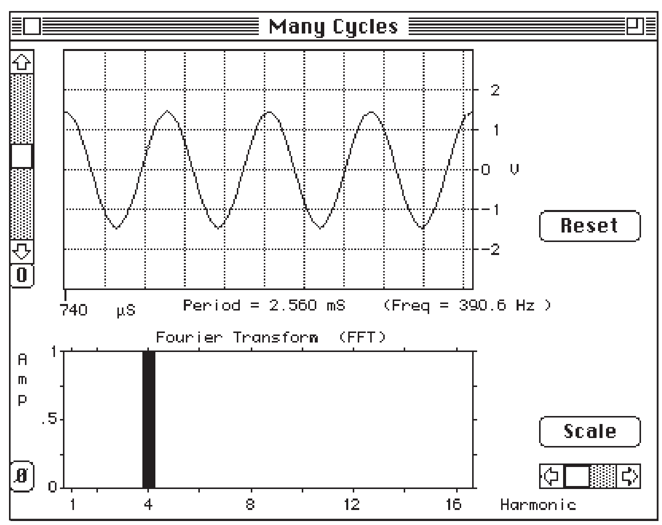


Figure 11
When 4 cycles are selected, the wave form consists of a pure 4th harmonic. We see the period of the selected section of wave, which is 4 times longer than one cycle. This makes the frequency 1/4 as high. (The difference between 1515 Hz/4 and 390.6 Hz indicates the accuracy of graphical selection.)

Analysis of a Square Wave

The so called square wave, whose shape is shown in Figure (12), is commonly used in electronics labs to study the response of various electronic circuits. The waveform regularly jumps back and forth between two levels, giving it a repeated rectangular shape. The ideal mathematical square wave jumps instantaneously from one level to another. The square waves we study in the lab are not ideal; some time is always required for the transition.

It is traditional to use the square wave as the first example to show students how a complex wave form can be constructed from harmonic sine waves. This is a bit ironic, because the ideal square wave has discontinuous jumps from one level to another, and therefore does not satisfy Fourier's theorem that any ***continuous wave shape*** can be made from harmonic sine waves. The result is that if you try to construct an ideal square wave from sine waves, you end up with a small blip at the discontinuity (called the Gibb's effect). Since our focus is experimental data where there is no true discontinuity, we will not encounter this problem.

In Figure (12) we have selected one cycle of the square wave. Selecting ***Fourier Analysis***, we get the result shown in Figure (13). We have clicked on the ***Expand*** button so that only the selected section of the wave shows in the upper rectangle. In the lower rectangle, which we will now call the FFT window, we see that this section of the square wave is made up from various harmonics. The MacScope program calculates 128 harmonics, but we have clicked three times on the ***Scale*** button to expand the harmonics scale so that we can study the first 16 harmonics in more detail.

In Figure (14) we have clicked on the first bar in the FFT window, the bar that represents the amplitude of the first harmonic. In the upper window you see one cycle of a sine wave superimposed upon the square wave. This is a picture of the first harmonic. It represents the best possible fit of the square wave by a single sine wave. If you want a better fit, you have to add in more sine waves.

In Figure (15) we clicked on the bar in the 3rd harmonic position, the bar representing the amplitude of the 3rd harmonic in the square wave. In the upper window you see a sine wave with a smaller amplitude and three times the frequency of the first harmonic. If you select a single harmonic, as we have just done in Figure (15), MacScope prints the frequency of both the first harmonic and the selected harmonic above the FFT window. Here you can see that the frequency of the first harmonic is 201.6 Hz and the selected harmonic frequency is 604.9 Hz as expected.

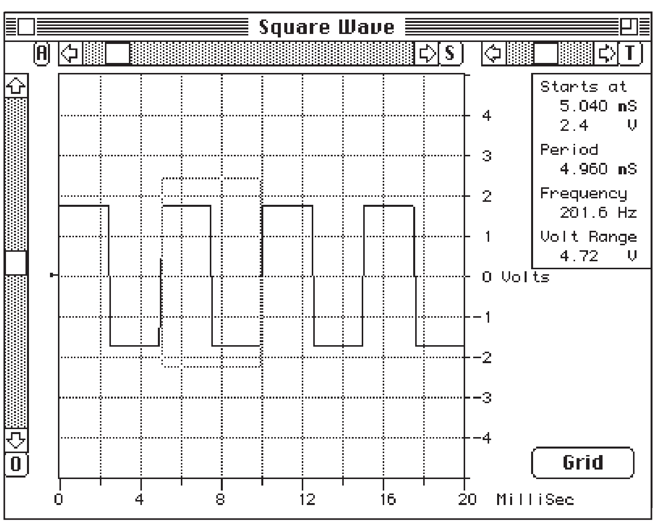


Figure 12
The square wave. The wave shape goes back and forth periodically between two levels. Here we have selected one cycle of the square wave.

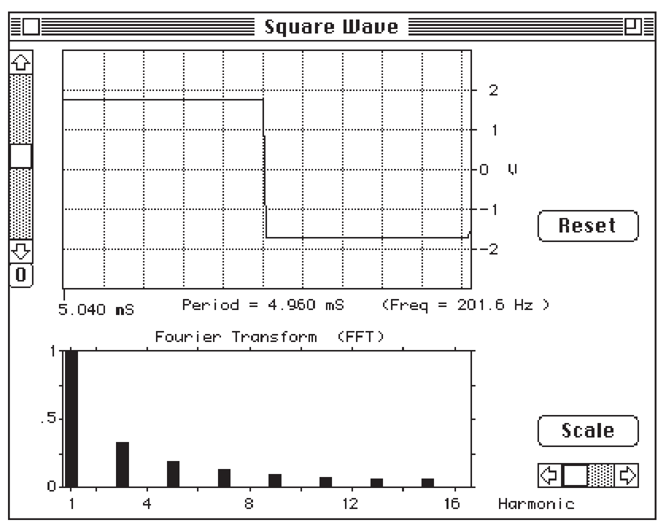


Figure 13
Expanding the one cycle selected, and choosing **Fourier Analysis*****, we see that this wave form has a number of harmonics. The computer program calculates the first 128 harmonics. We used the*** **Scale** ***button to display only the first 16.***

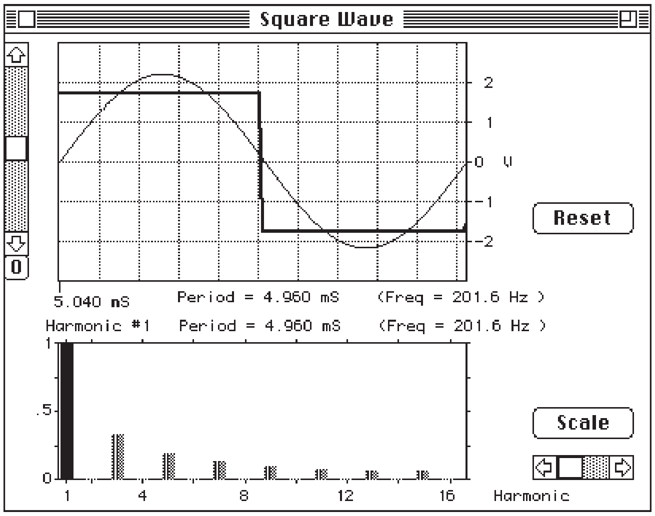


Figure 14
Select the first harmonic by clicking on the first bar.

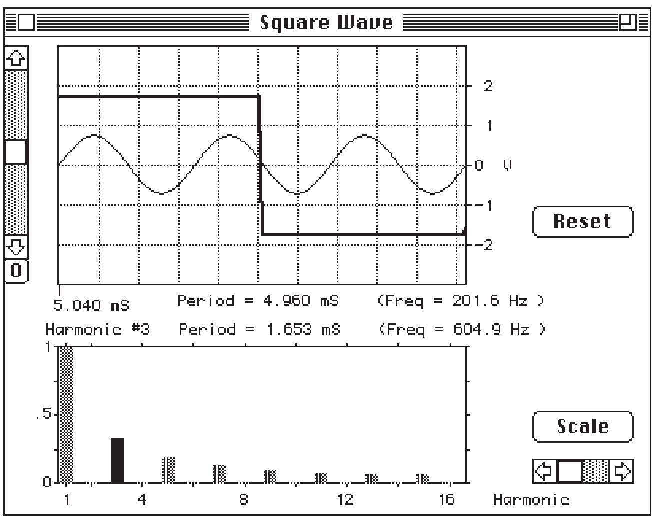


Figure 15
Select the second harmonic by clicking on the second bar.

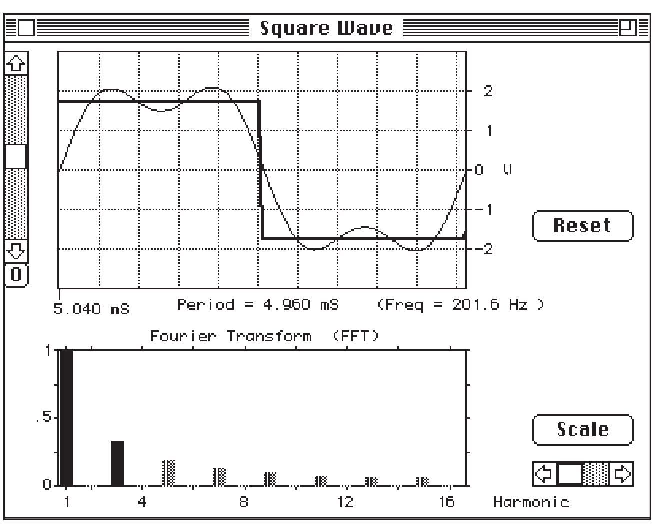


Figure 16
Sum of the first and second harmonic is obtained by selecting both.

If we select both the first and third harmonics together (either by dragging a rectangle over both bars, or by holding down the shift key while selecting them individually), the upper window displays the sum of the first and third harmonic, as shown in Figure (16). You can see that the sum of these two harmonics gives us a waveform that is closer to the shape of the square wave than either harmonic alone. We are beginning to build up the square wave from sine waves.

In Figures (17, 18 and 19), we have added in the 5th, 7th and 9th harmonics. You can see that the more harmonics we add, the closer we get to the square wave.

One of the special features of a square wave is that it contains only odd harmonics—all the even harmonics are absent. Another is that the amplitude of the nth harmonic is 1/n times as large as that of the first harmonic. For example, the third harmonic has an amplitude only 1/3 as great as the first. This is represented in the FFT window by drawing a bar only 1/3 as high as that of the first bar. In the MacScope program, the harmonic with the greatest amplitude is represented by a bar of height 1. All other harmonics are represented by proportionally shorter bars.

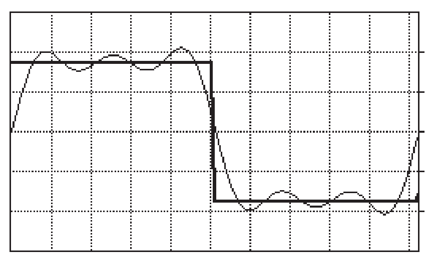


Figure 17
Sum of the harmonics 1, 3, and 5.

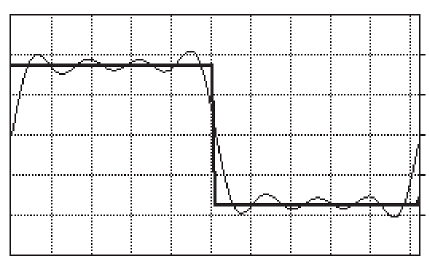


Figure 18
Sum of the harmonics 1, 3, 5, 7.

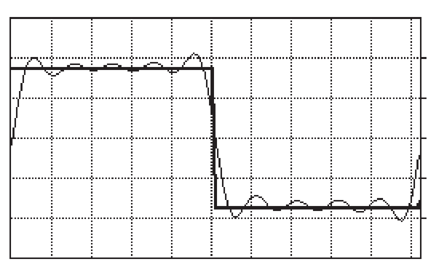


Figure 19
Sum of the harmonics 1, 3, 5, 7, 9.

Repeated Wave Forms

Before we apply the Fourier transform capability of MacScope to the analysis of experimental data, there is one more feature of the analysis we need to discuss. What we are doing with the program is reconstructing a selected section of a waveform from harmonic sine waves. Anything we build from harmonic sine waves ***exactly repeats at the period of the first harmonic***. Thus our reconstructed wave will always be a repeating wave, beginning again at the same height as the beginning of the previous cycle.

You can most easily see what we mean if you select a non repeating section of a wave. In Figure (20) we have gone back to a sine wave, but selected one and a half cycles. In the FFT window you see a slew of harmonics. To see why these extra harmonics are present, we have in Figure (21) selected the first 9 of them and in the upper window see what they add up to. It is immediately clear what has gone wrong. The selected harmonics are trying to reconstruct a repeating version of our 1.5 cycle of the sine wave. The extra spurious harmonics are there to force the reconstructed wave to start and stop at the same height as required by a repeating wave.

If you are analyzing a repeating wave form and select a section that repeats, then your harmonic reconstruction will be accurate, with no spurious harmonics. If your data is not repeating then you have to deal one way or another with this problem. One technique often used by engineers is to select a long section of data and smoothly force the ends of the data to zero so that the selected data can be treated as repeating data. Hopefully, forcing the ends of the data to zero does not destroy the information you are interested in. You can often accomplish the same thing by throwing away higher harmonics, assuming that the lower harmonics contain the interesting features of the data. You can see that neither of these techniques would work well for our one and a half cycles of a sine wave selected in Figure (21).

In this text, our use of Fourier analysis will essentially be limited to the analysis of repeating waveforms. As long as we select a section that repeats, we do not have to worry about the spurious harmonics.

(Most programs for the acquisition and analysis of experimental data have an option for doing Fourier analysis. Unfortunately, few of them allow you to select a precise section of the experimental data for analysis. As a result the analysis is usually done on a non repeating section of the data, which distorts the resulting plot of the harmonic amplitudes. For a careful analysis of data, the ability to precisely select the data to be analyzed is an essential capability.)

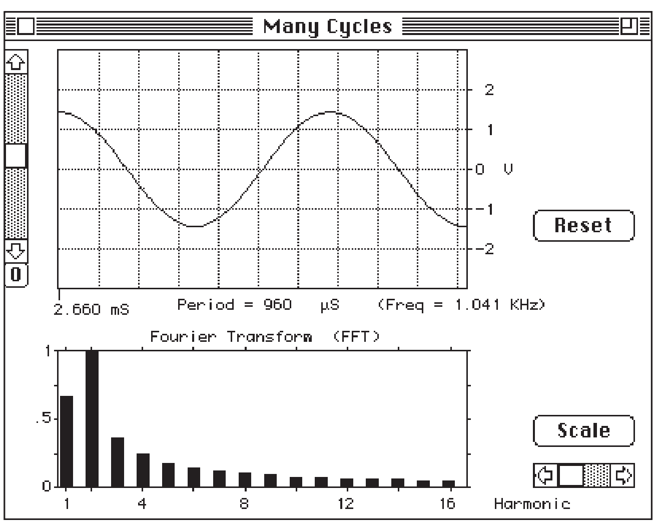


Figure 20
When we select one and a half cycles of a sine wave, we get a whole bunch of spurious harmonics.

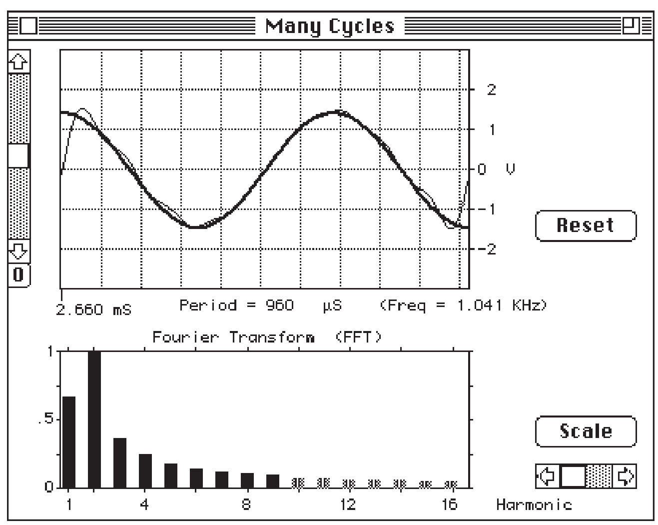


Figure 21
The spurious harmonics result from the fact that the reconstructed wave must be repeating—must start and stop at the same height.

ANALYSIS OF THE COUPLED AIR CART SYSTEM

We are now ready to apply Fourier analysis to our system of coupled air carts. Recall that there were two modes of motion that resulted in a sinusoidal oscillation of the carts, the vibrational motions shown in Figure (4) and the sloshing mode shown in Figure (5). In Figures (22) and (23) we expanded the time scales so that we could accurately measure the period and frequency of these oscillations. From the data rectangles, we see that the frequencies were 1.11Hz and 0.336 Hz for the vibrational and sloshing modes respectively.

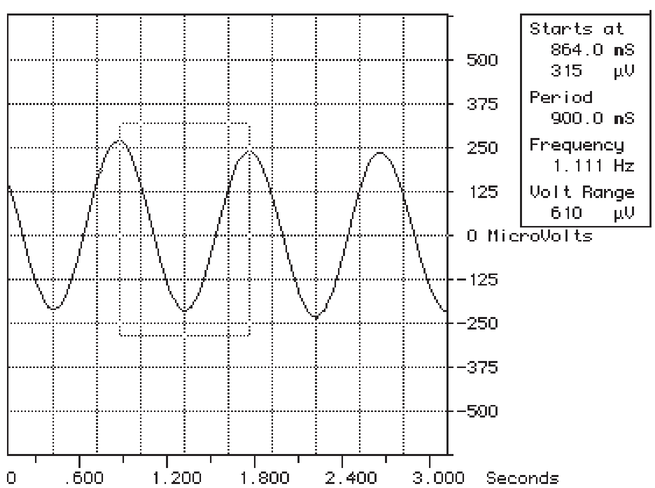


Figure 22
Vibrational mode of Figure (4). We have expanded the time scale so that we could accurately measure the period and frequency of the oscillation.

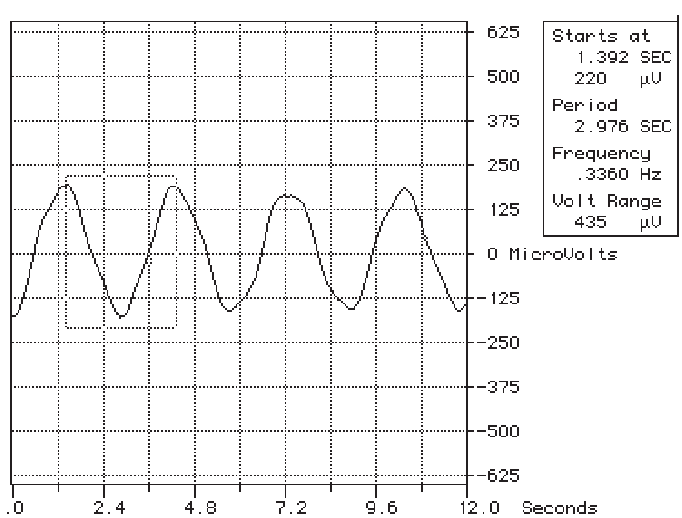


Figure 23
Sloshing mode of Figure (5). Less of an expansion of the time scale was needed to measure the period here.

When the carts were released in an arbitrary way, we generally get the complex motion seen in Figures (6) and (7). What we wish to do now is apply Fourier analysis to these waveforms to see if any simple features underlie this complex motion.

In Figure (24), which is the waveform of Figure (6), we see that there is a repeating pattern. The fact that the pattern repeats means that it can be reconstructed from harmonic sine waves, and we can use our Fourier analysis program to find out what the component sine waves are.

In Figure (24) we have selected precisely one cycle of the repeating pattern. This is the crucial step in this experiment—finding the repeating pattern and selecting one cycle of it. How far you have to look for the pattern to repeat depends upon the mass of the carts and the strength of the springs.

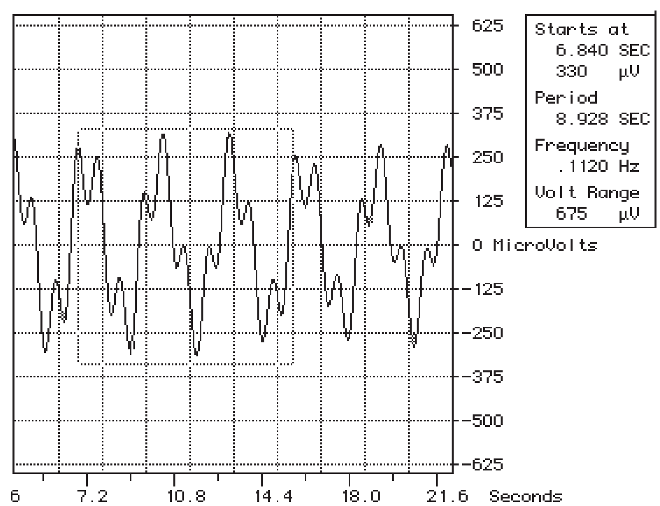


Figure 24
Complex mode of the coupled aircarts. We see that the waveform repeats, and have selected one cycle of the repeating wave.

Expanding the repeating section of the complex waveform, and choosing ***Fourier Analysis*** gives us the results shown in Figure (25). What we observe from the Fourier analysis is that the complex waveform is a mixture of two harmonics, in this case the third and tenth harmonic. If we click on the bar showing the amplitude of the third harmonic, we see that harmonic drawn in the display window of Figure (26), and we find that the frequency of this harmonic is 0.336 Hz. This is the frequency of the sloshing mode of Figure (23). Clicking on the bar above the tenth harmonic, we get the harmonic drawn in the display window of Figure (27), and see that the frequency of this mode is 1.12 Hz, within a fraction of a percent of the 1.11 Hz frequency of the vibrational mode of Figure (22).

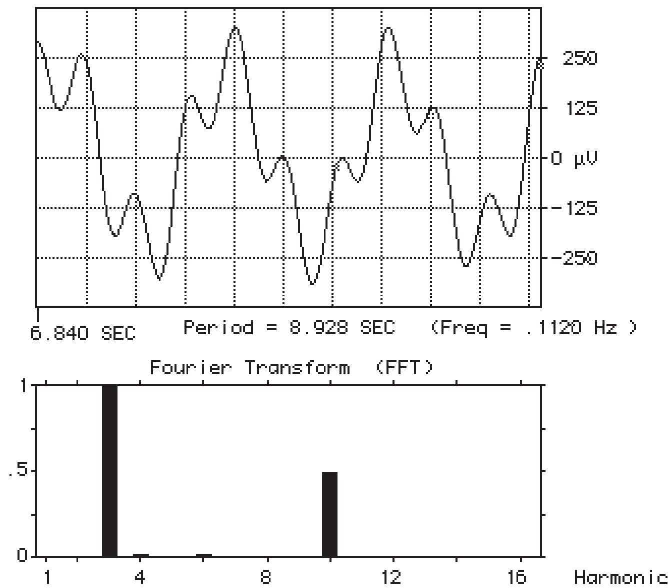


Figure 25
Fourier analysis of one cycle of the complex waveform. The FFT rectangle shows us that the wave consists of only two harmonics.

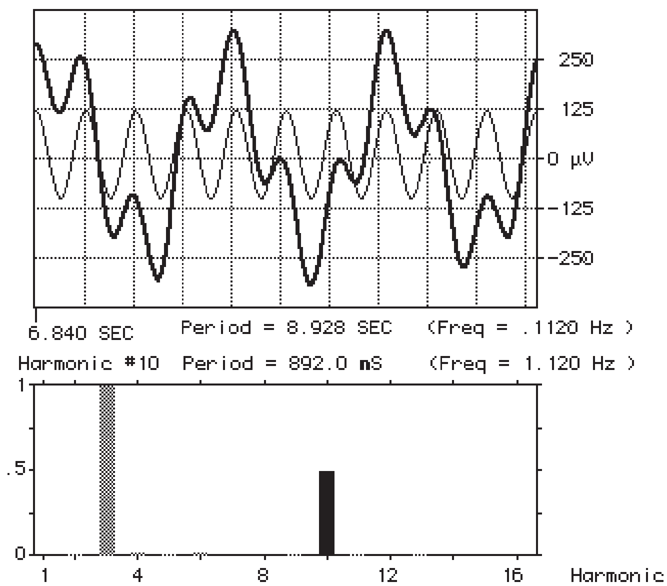


Figure 27
Selecting the tenth harmonic, we see that its frequency is essentially equal to the frequency of the vibrational mode of motion.

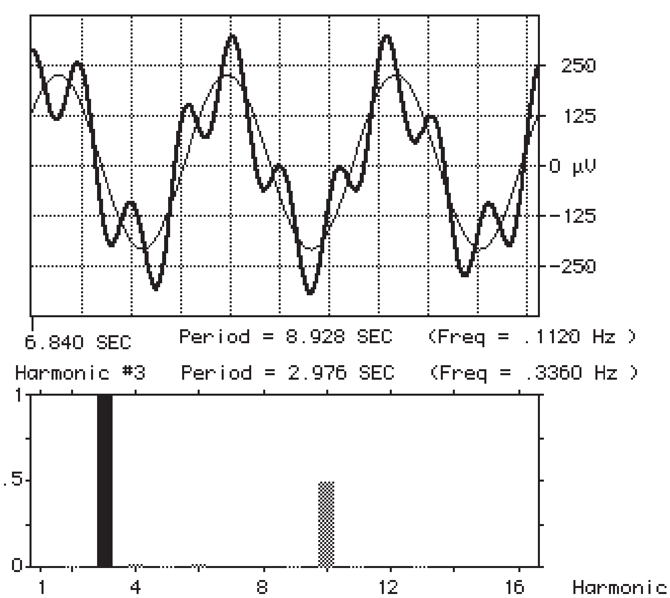


Figure 26
When we click on the third harmonic bar, we see that the frequency of the third harmonic is precisely the frequency of the sloshing mode of oscillation.

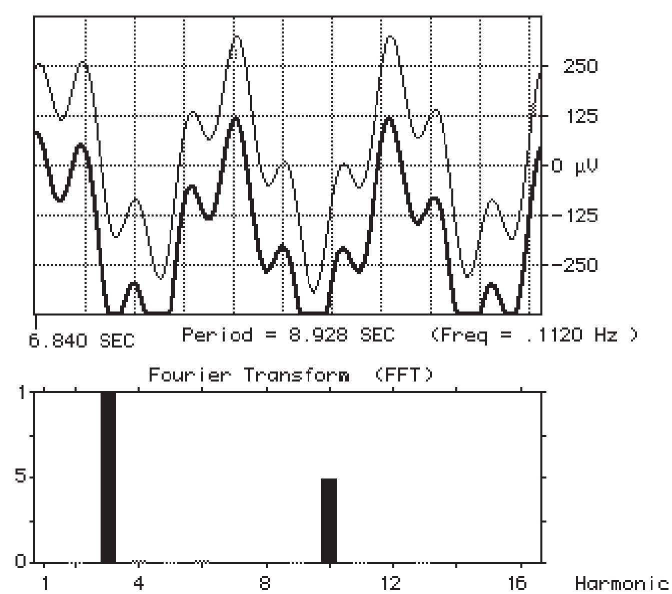


Figure 28
Selecting both modes shows us that the complex motion is simply the sum of the two sinusoidal modes of motion.

If we select both the third and tenth harmonics, the sum of these two harmonics is shown in Figure (28). These two harmonics together so closely match the experimental data that we had to move the experimental curve down in order to see both curves. What we have learned from this experiment is that ***the complex motion of Figure (6) is a mixture of the two simple, sinusoidal modes of motion.***

Back in Figures (6) and (7), we displayed two waveforms, representing different complex motions of the same two carts. Starting with the second waveform of Figure (7), we selected one cycle of the motion, expanded the selected section, and chose Fourier Analysis. The result is shown in Figure (29). What we see is that the second complex waveform is also a mixture of the third and tenth harmonics. The first waveform in Figure (25) had more of the third harmonic, more of the sloshing mode, while the second waveform of Figure (29) has more of the tenth harmonic, the vibrational mode. Both complex waveforms are simply mixtures of the vibrational and sloshing modes. ***They have different shapes because they are different mixtures.***

This experiment is beginning to demonstrate that for the two coupled aircarts, there is a strict limitation to the kind of motion the carts can have. They can either move in the vibrational mode, or in the sloshing mode, or in some combination of the two modes. ***No other kinds of motion are allowed! The various complex motions are just different combinations of the two modes.***

Adding another aircart so that we have three coupled aircarts, the motion becomes still more complex. However if we look carefully, we find that the waveform eventually repeats. Selecting one repeating cycle and choosing Fourier Analysis, we got the results shown in Figure (30). We observe that this complex motion is made up of three harmonics.

The sinusoidal modes of motion of the coupled air carts are called ***normal modes***. The general rule is that if you have n coupled objects, like n carts on an air track connected by springs, and they are confined to move in 1 dimension, there will be n normal modes. (With 2 carts, we saw 2 normal modes. With 3 carts, 3 normal modes, etc.) This result, which will play an important role in our discussion of the specific heat of molecules, can be extended to motion in 2 and 3 dimensions. For example, if you have n coupled particles that can move in 3 dimensions, as in the case of a molecule with n atoms, then the system should have 3n normal modes of motion. Such a molecule should have 3n independent ways to vibrate or move. We will have more to say about this subject in the next chapter.

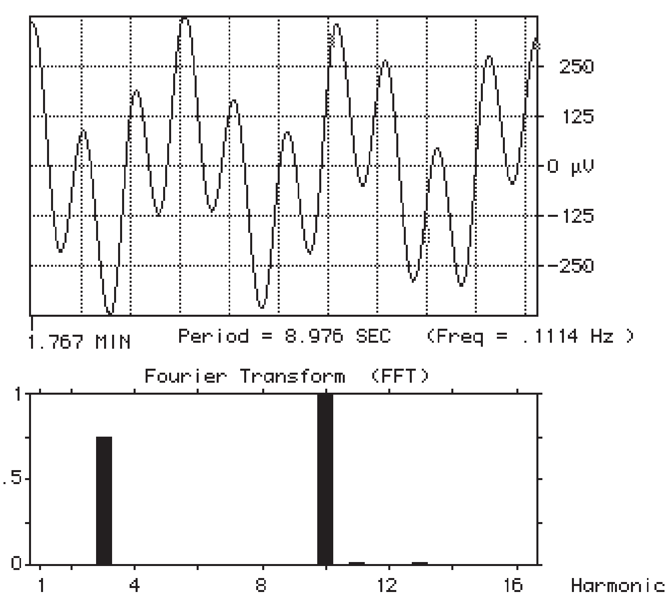


Figure 29
The second complex mode of motion, from Figure (7), is simply a different mixture of the same vibrational and sloshing modes of motion.

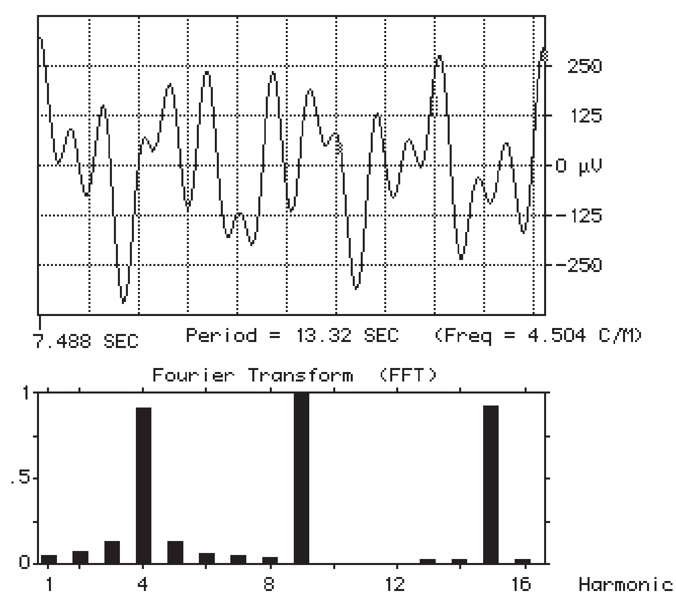


Figure 30
One cycle of the waveform for three coupled aircarts. With three carts, we get three normal modes of motion.

THE HUMAN EAR

The human ear performs a frequency analysis of sound waves that is not unlike the Fourier analysis of wave motion which we just studied. In the ear, the initial analysis is done mechanically, and then improved and sharpened by a sophisticated data analysis network of nerves. We will focus our attention on the mechanical aspects of the ear's frequency analysis.

Figure (31) is a sketch of the outer and inner parts of the human ear. Sound waves, which consist of pressure variations in the air, are funneled into the auditory canal by the external ear and impinge on the eardrum, a large membrane at the end of the auditory canal. The eardrum (tympanic) membrane vibrates in response to the pressure variations in the air. This vibrational motion is then transferred via a lever system of three bones (the malleus, incus, and stapes) to a small membrane covering the oval window of the snail shaped cochlea.

The cochlea, shown unwound in Figure (32), is a fluid filled cavity surrounded by bone, that contains two main channels separated by a membrane called the ***basilar membrane***. The upper channel (scala vestibuli) which starts at the oval window, is connected at the far end to the lower channel (scala tympani) through a hole called the helicotrema. The lower channel returns to the round window which is also covered by a membrane. If the stapes pushes in on the membrane at the oval window, fluid flows around the helicotrema and causes a bulge at the round window.

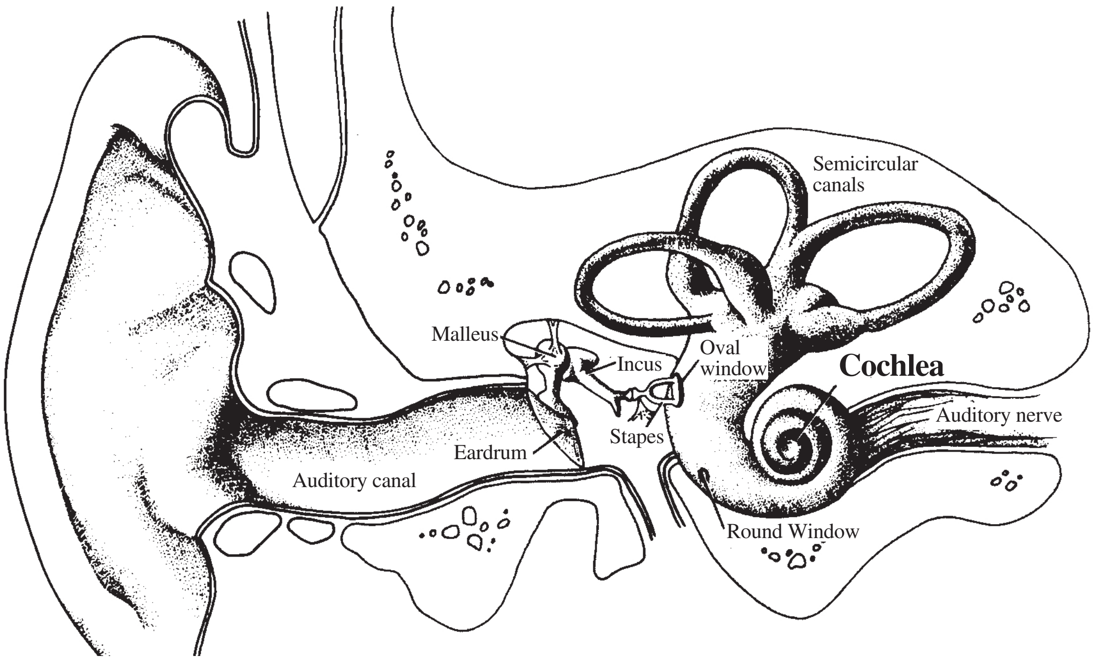


Figure 31
The human ear. Sound, entering the auditory canal, causes vibrations of the eardrum. The vibrations are transferred by a bone lever system to the membrane covering the oval window. Vibrations of the oval window membrane then cause wave motion in the fluid in the cochlea. *(Adapted from Lindsey and Norman,* **Human Information Processing.***)*

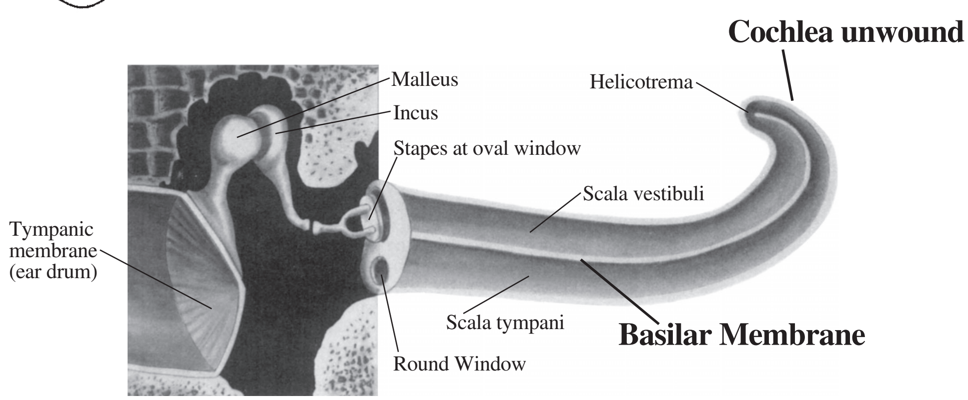


Figure 32
Lever system of the inner ear and an unwound view of the cochlea. The basilar membrane separates the two main fluid channels in the cochlea. Vibrations of the basilar membrane are detected by hair cells. *(Adapted from* **Principles of Neural Science** *Edited by E. R. Kandel and J. H. Schwartz, Elsevier/North-Holland, p260.)*

The purpose of the lever system between the eardrum and the cochlea is to efficiently transfer sound energy to the cochlea. The eardrum membrane is about 25 times larger in area than the membrane across the oval window. The lever system transfers the total force on the eardrum to an almost equal force on the oval window membrane. Since *force* equals *pressure* times *area*, a small pressure variation acting on the large area of the eardrum membrane results in a large pressure variation at the small area at the oval window. The higher pressures are needed to drive a sound wave through the fluid filled cochlea.

If the oval window membrane is struck by a pulse, a pressure wave travels down the cochlea. The basilar membrane, which separates the two main fluid channels, moves in response to the pressure wave, and a series of hair cells along the basilar membrane detect the motion. It is the way in which the basilar membrane responds to the pressure wave that allows for the frequency analysis of the wave.

Figure (33) is an idealized sketch of a straightened out cochlea. (See Appendix B for more realistic sketches.) At the front end, by the oval window, the basilar membrane is narrow and stiff, while at the far end it is about 5 times as wide and much more floppy. To see why the basilar membrane has this structure, we have in Figure (34) sketched a mechanical model that has a similar function as the membrane. In this model we have a series of masses mounted on a flexible steel band and attached by springs to fixed rods as shown. The masses are small and the springs stiff at the front end. If we shake these small masses, they resonate at a high frequency $\omega = \sqrt{k/m}$. Down the membrane model, the masses get larger and the springs weaken with the result that the resonant frequency becomes lower.

If you gently shake the steel band at some frequency ω_0 a small amplitude wave will travel down the band and soon build up a standing wave of that frequency. If ω_0 is near the resonant frequency of one of the masses, that mass will oscillate with a greater amplitude than the others. Because the masses are connected by the steel band, the neighboring masses will be carried into a slightly larger amplitude of motion, and we end up with a peak in the amplitude of oscillation centered around the mass whose frequency ω is equal to ω_0. (In the sketch, we are shaking the band at the resonant frequency ω_7 of the seventh mass.)

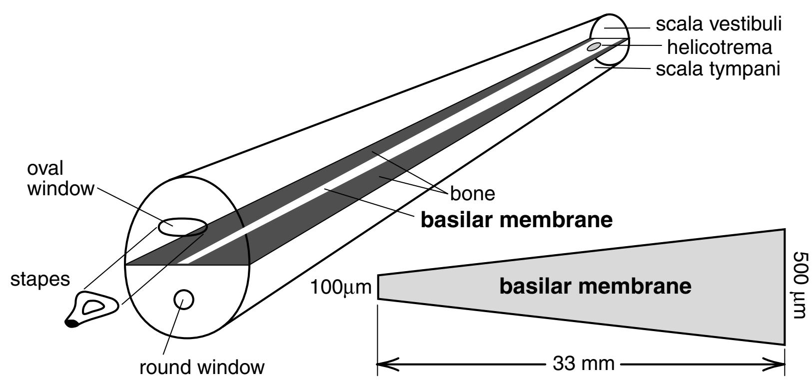


Figure 33
The basilar membrane in the cochlea. (Adapted from Green,* An Introduction to Hearing, *John Wiley & Sons, p66.)

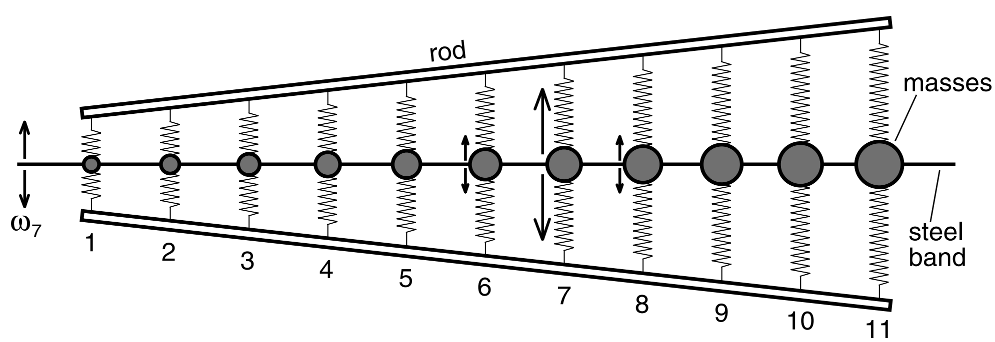


Figure 34
Spring model of the basilar membrane. As we go down the steel band, the masses become larger, the springs weaker, and the resonant frequency drops. If we vibrate the end of the band at some frequency, the mass which resonates at that frequency will have the biggest amplitude of oscillation.

Because the masses m get larger and the spring constants k get smaller toward the far end, there is a continuous decrease in the resonant frequency as we go down the band. If we start shaking at a high frequency, the resonant peak will occur up near the front end. As we lower the frequency, the peak of the oscillation will move down the band, until it finally gets down to the lowest resonant frequency mass at the far end. We can thus measure the frequency of the wave by observing where along the band the maximum amplitude of oscillation occurs.

The basilar membrane functions similarly. The stiff narrow membrane at the front end resonates at a high frequency around 20,000 Hz, while the wide floppy back end has a resonant frequency in the range of 20 to 30 Hz. Figure (35) shows the amplitude of the oscillation of the membrane in response to driving the fluid in the cochlea at different frequencies. We can see that as the frequency increases, the location of the maximum amplitude moves toward the front of the membrane, near the stapes and oval window. Figure (36) depicts the shape of the membrane at an instant of maximum amplitude when driven at a frequency of a few hundred Hz. The amplitude is greatly exaggerated; the basilar membrane is about 33 millimeters long and the amplitude of oscillation is less than .003 mm.

Although the amplitude of oscillation is small, it is accurately detected by a system of about 30,000 hair cells. How the hair cells transform the oscillation of the membrane into nerve impulse signals is discussed in Appendix B at the end of this chapter.

The human ear is capable of detecting tiny changes in frequency and very subtle mixtures of harmonics in a sound. Looking at the curves in Figure (35) (which were determined from a cadaver and may not be quite as sharp as the response curves from a live membrane), it is clear that it would not be possible to make the ear's fine frequency measurements simply by looking for the peak in the amplitude of the oscillation of the membrane. But the ear does not do that. Instead, measurements are continuously made all along the membrane, and these results are fed into a sophisticated data analysis network before the results are sent to the brain. The active area of current research is to figure out how this data analysis network operates.

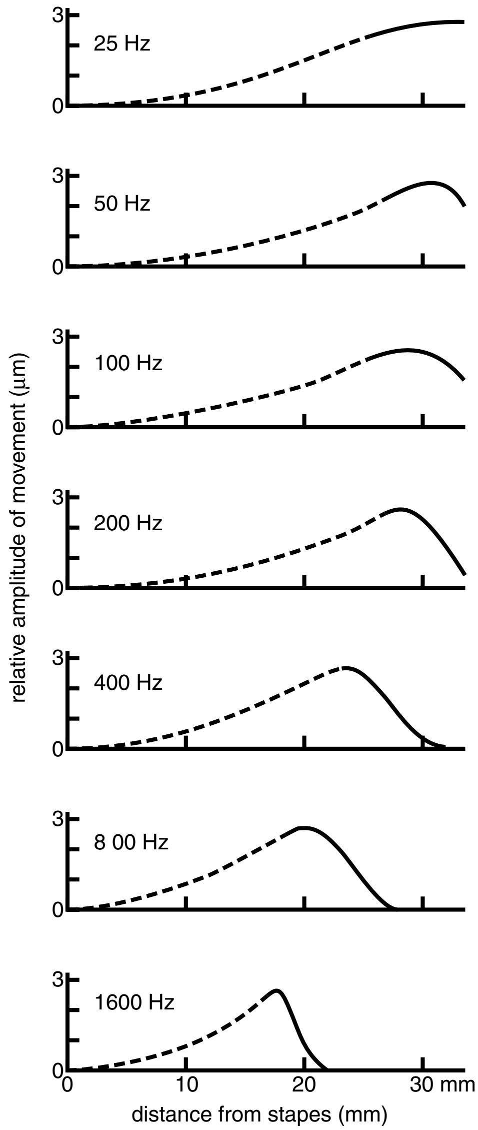


Figure 35
Amplitude of the motion of the basilar membrane at different frequencies. (Adapted from* Principles of Neural Science *Edited by E. R. Kandel and J. H. Schwartz, Elsevier/North-Holland, p 263.)

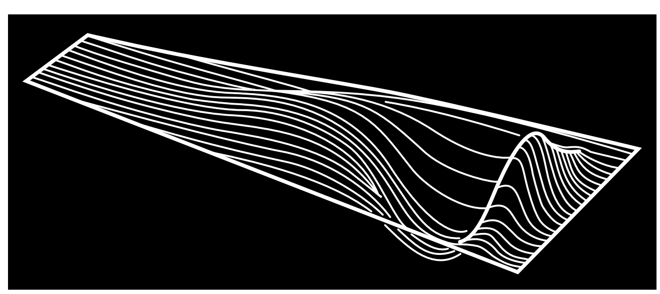


Figure 36
Response of the basilar membrane to a moderately low frequency driving force. (From Vander,A; Sherman,J; and Luciano,D.* Human Physiology*, 4th edition, 1985, P662. McGraw Hill Publishing Co., NY.)

STRINGED INSTRUMENTS

The stringed instruments provide the clearest example of how musical instruments function. The only possible modes of oscillation of the string are those with nodes at the ends, and we have seen that the frequencies of these modes form a harmonic series $f_n = n f_1$. This suggests that if we record the sound produced by a stringed instrument and take a Fourier transform to see what harmonics are present in the sound, we can tell from that what modes of oscillation were present in the vibrating string.

This is essentially correct for the electric stringed instruments like the electric guitar and electric violin. Both of these instruments have a magnetic pickup that detects the velocity of the string at the pickup, using the same principle as the velocity detector we used in the air cart experiments discussed earlier in this chapter. The voltage signal from the magnetic pickup is then amplified electronically and sent to a loudspeaker. Thus the sound we hear is a fairly accurate representation of the motion of the string, and an analysis of that sound should give us a good idea of which modes of oscillation of the strings were excited.

The situation is different for the acoustic stringed instruments, like the acoustic guitar used by folk singers, and the violin, viola, cello and base, found in symphony orchestras. In these instruments the vibration of the string does not produce that much sound itself. Instead, the vibrating string excites resonances in the sound box of the instrument, and it is the sound produced by the resonating sound box that we hear. As a result the quality of the sound from an acoustic string instrument depends upon how the sound box was constructed. Subtle differences in the shape of the sound box and the stiffness of the wood used in its construction can lead to subtle differences in the harmonics excited by the vibrating string. The human ear is so sensitive to these subtle differences that it can easily tell the difference between a great instrument like a 280 year old Stadivarius violin, and even the best of the good instruments being made today. (It may be that it takes a couple of hundred years of aging for a very good violin to become a great one.)

To demonstrate the difference between electric and acoustic stringed instruments, and to illustrate how Fourier analysis can be used to study these differences, my daughter played the same note, using the same bowing technique, on the open E string of both her electric and her acoustic violins. Using the same microphone in the same setting to record both, we obtained the results shown in Figure (37).

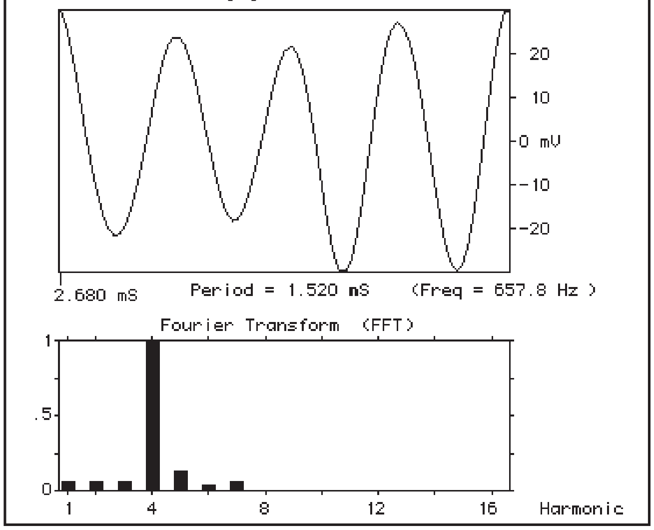
From Figure (37) we immediately see why acoustic stringed instruments sound differently from their electric counterparts. With the acoustic instruments you get a far richer mixture of harmonics. In the first trial, labeled *E(1) Electric Violin*, the string was bowed so that it produced a nearly pure 4th harmonic. The sound you hear corresponds to a pure tone of frequency 657.8/4 = 264 Hz. The corresponding sound produced by the acoustic violin has predominately the same 4th harmonic, but a lot of the sound is spread through the first 8 harmonics.

A number of recordings were made, so that we could see how the sounds varied from one playing to the next. The examples shown in Figure (37) are typical. It is clear in all cases that for this careful bowing a single harmonic predominated in the electric violin while the acoustic violin produced a mixture of the first eight. It is rather surprising that the ear hears all of these sounds as representing the same note, but with a different quality of sound.

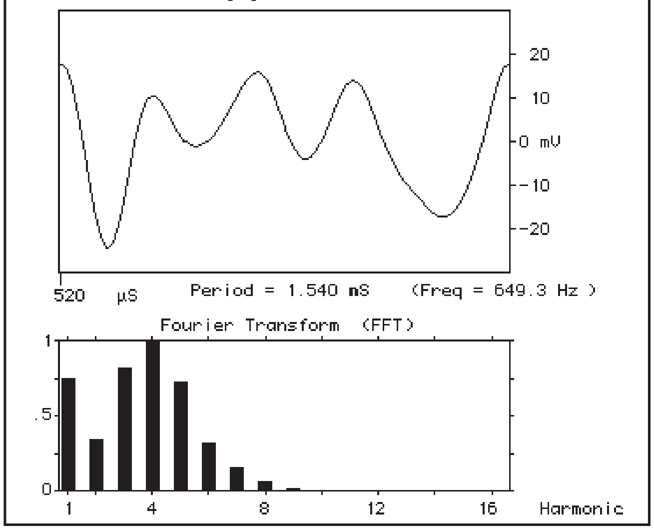
We chose the violin for this comparison because by using a bow, one can come much closer to exciting a pure mode of vibration of the string. We see this explicitly in the *E(1) Electric Violin* example. When you pluck or strum a guitar, even if you pluck only one string, you get a far more complex sound than you do for the violin. If you pluck a chord on an acoustic guitar, you get a very complex sound. It is the complexity of the sound that gives the acoustic guitar a richness that makes it so effective for accompanying the human voice.

Figure 37
Comparison of the sound of an electric and an acoustic violin. In each case the open E string was bowed as similarly as possible. The electric violin produces relatively pure tones. The interaction between the string and the sound box of the acoustic violin gives a much richer mixture of harmonics.

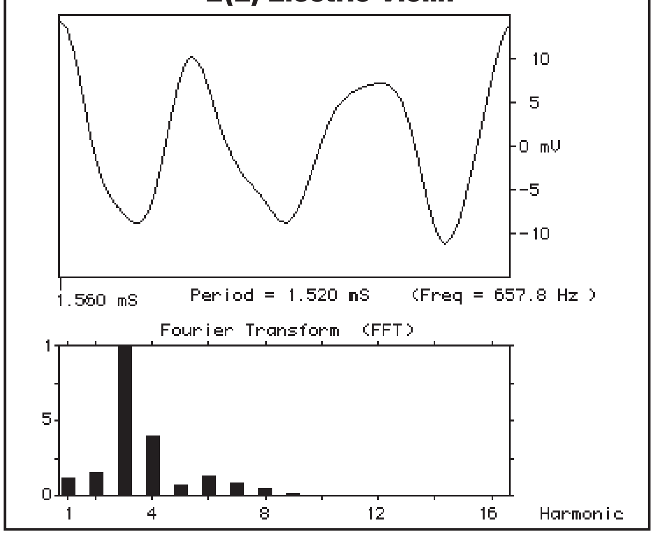
E(1) Electric Violin



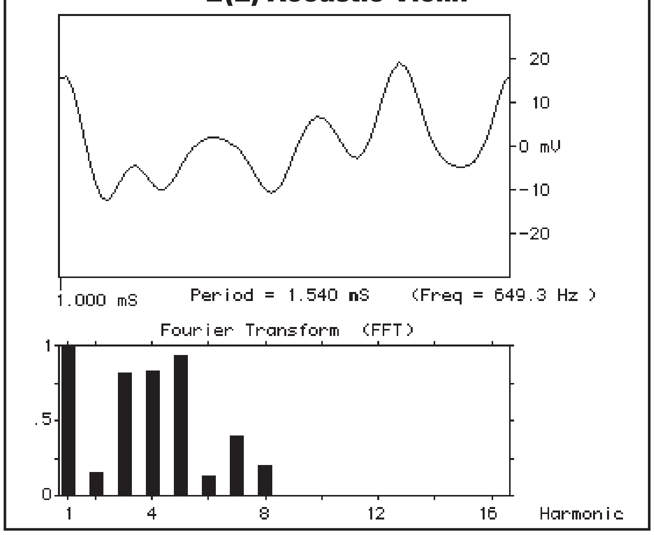
E(1) Acoustic Violin



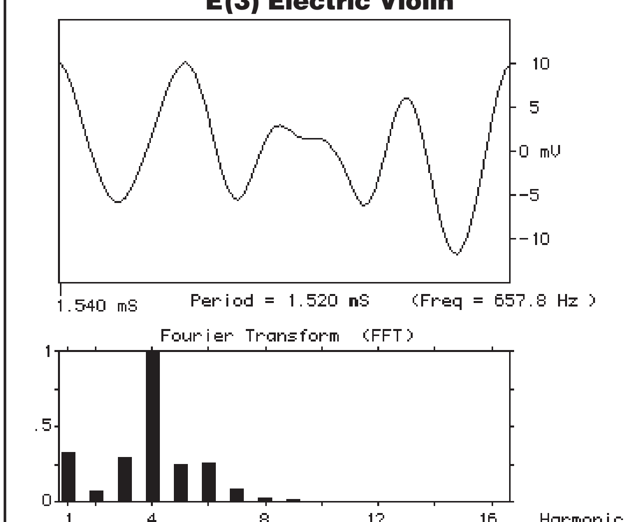
E(2) Electric Violin



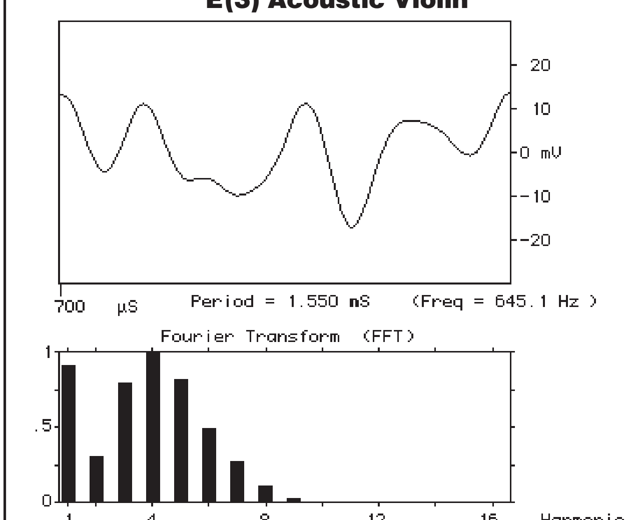
E(2) Acoustic Violin



E(3) Electric Violin



E(3) Acoustic Violin



Recording the sound of an instrument and using Fourier analysis can be an effective tool for studying musical instruments, but care is required. For example, in comparing two instruments, start by choosing a single note, and try to play the note the same way on both instruments. Make several recordings so that you can tell whether any differences seen are due to the way the note was played or due to the differences in the instruments themselves. With careful work, you can learn a lot about the nature of the instrument.

In the 1970s, before we had personal computers, students doing project work would analyze the sound produced by instruments or the voice, using a time sharing mainframe computer system to do the Fourier analysis. Hours of work were required to analyze a single sound, but the results were so interesting that they served as the incentive to develop the MacScope™ program when the Macintosh computer became available

I particularly remember an early project in which two students compared the spinet piano, the upright piano and the grand piano. They recorded middle C played on each of these pianos. Middle C on the spinet consisted of a wide band of harmonics. From the upright there were still a lot of harmonics, but the first, third, and fifth began to predominate. The grand piano was very clean with essentially only the first, third, and fifth present. You could clearly see the effect of the increase in the size of the musical instrument.

The same year, another student, Kelly White, took a whale sound from the Judy Collins record *Sound of the Humpback Whales*. Listening to the record, the whale sounds are kind of squeaky. But when the sound was analyzed, the results were strikingly similar to those of the grand piano. The analysis suggested that the whale sounds were by an instrument as large as, or larger, than a grand piano. (The whale's blowhole acts as an organ pipe when the whale makes the sound.)

WIND INSTRUMENTS

While the string instruments are all based on the oscillation of a string, the wind instruments, like the organ, flute, trumpet, clarinet, saxophone, and glass bottle, are all based on the oscillations of an air column. Of these, the bottle is the most available for studying the nature of the oscillations of an air column.

When you blow carefully across the top of a bottle, you hear a sound with a very definite frequency. Add a little water to the bottle and the frequency of the note rises. When you shortened the length of a string, the pitch went up, thus it is not surprising that the pitch also goes up when you shorten the length of the air column.

You might guess that the mode of vibration you set up by blowing across the top of the bottle has a node at the bottom of the bottle and an anti node at the top where you are blowing. For a sine wave the distance from a node to the next anti node is 1/4 of a wavelength, thus you might predict that the sound has a wavelength 4 times the height of the air column, and a frequency

$$f = \frac{v_{sound}}{\lambda} = \frac{v_{sound}}{4d} \qquad \left(\text{for } \lambda = 4d\right) \tag{4}$$

This prediction is not quite right as you can quickly find out by experimenting with various shaped bottles. Add water to different shaped bottles, adjusting the levels of the water so that all the bottles have the same height air columns. When you blow across the top of the different bottles, you will hear distinctly different notes, the fatter bottles generally having lower frequencies than the skinny ones. Unlike the vibrating string, it is not just the length of the column and the speed of the wave that determine the frequency of oscillation, the shape of the container also has a noticeable effect.

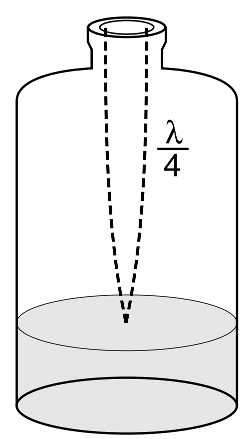


Figure 38
Blowing across the top of a bottle produces a note whose wavelength is approximately 4 times the height of the air column.

Despite this additional complexity, there is one common feature to the air columns encountered in musical instruments. They all have unique frequencies of oscillation. Even an organ pipe that is open at both ends—a situation where you might think that the length of the column is not well defined—the air column has a precise set of frequencies. The fussiness in defining the length of the column does not fuzz out the sharpness of the resonance of the column.

Organ pipes, with their straight sides, come closest to the simple standing waves we have seen for a stretched instrument string. In all cases, the wave is excited at one end by air passing over a sharp edge creating a turbulent flow behind the edge. This turbulence excites the air column in much the same way that dragging a sticky bow across a violin string excites the oscillation of the string.

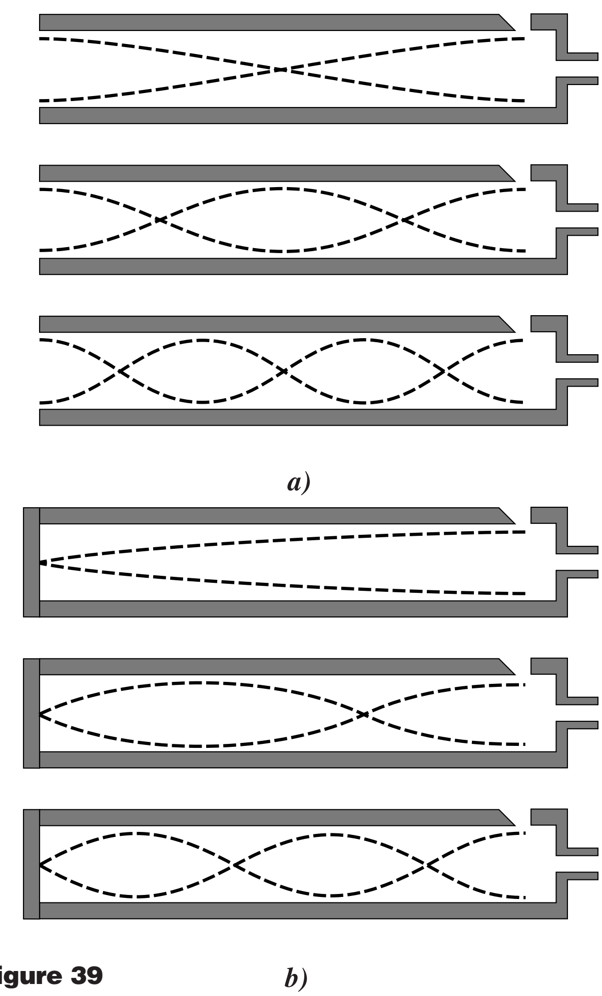


Figure 39
Modes of oscillation of an air column in open and closed organ pipes.

The various modes of oscillation of an air column in an organ pipe are shown in Figure (39). All have an anti node at the end with sharp edges where the turbulence excites the oscillation. The open ended pipes shown in Figure (39a) also have anti nodes at the open end, while the closed pipes of Figure (39b) have a node at the far end. The pictures in Figure (39a) are a bit idealized, but give a reasonably accurate picture of the shape of the standing wave. We can compensate for a lack of accuracy of these pictures by saying, for example, that the anti node of the open ended pipes lies somewhat beyond the end of the pipe.

Exercise 1

(a) Find the formula for the wavelength λ_n of the nth harmonic of the open ended pipes of Figure (39).

(b) Assuming that the frequencies of vibration are given by $f(cycle/sec) = v(meter/sec) / \lambda(meter/cycle)$ where v is the speed of sound, what is the formula for the allowed frequencies of the open organ pipe of length L?

(c) What should be the length L of an open ended organ pipe to produce a fundamental frequency of 440 cycles/second, middle A?

(d) Repeat the calculations of parts a, b, and c for the closed organ pipes of Figure (39).

(e) If you have the opportunity, find a real organ and check the predictions you have made (or try the experiment with bottles).

If you start with an organ pipe, and drill holes in the side, essentially converting it into a flute or one of the other wind instruments like a clarinet, you considerably alter the shape and frequency of the modes of oscillation of the air column. As a first approximation you could say that you create an anti node at the first open hole. But then when you play these instruments you can make more subtle changes in the pitch by opening some holes and closing others. The actual patterns of oscillation can become quite complex when there are open holes, but the simple fact remains that, no matter how complex the wave pattern, there is a precise set of resonant frequencies of oscillation. It is up to the maker of the instrument to locate the holes in such a way that the resonances have the desired frequencies.

PERCUSSION INSTRUMENTS

We all know that the string and wind instruments produce sound whose frequency we adjust to produce melodies and chords. But what about drums? They seem to just make noise. Surprisingly, drumheads have specific modes of oscillation with definite frequencies, just as do vibrating strings and air columns. But one does not usually adjust the fundamental frequency of oscillation of the drumhead, and the frequencies of the higher modes of oscillation do not follow the harmonic patterns of string and wind instruments.

To observe the standing wave patterns corresponding to modes of vibration of a drumhead, we can drive the drumhead at the resonant frequency of the oscillation we wish to study. It turns out to be a lot easier to drive a drumhead at a precise frequency than it is to find the normal modes of the coupled air cart system.

The experiment is illustrated in Figure (40). The apparatus consists of a hollow cardboard cylinder with a rubber sheet stretched across one end to act as a drumhead. At the other end is a loudspeaker attached to a signal generator. When the frequency of the signal generator is adjusted to the resonant frequency of one of the normal modes of the drumhead, the drumhead will start to vibrate in that mode of oscillation.

To observe the shape and motion of the drumhead in one of its vibrational modes, we place a strobe light to one side of the drumhead as shown. If you adjust the strobe to the same frequency as the normal mode vibration, you can stop the motion and see the pattern.

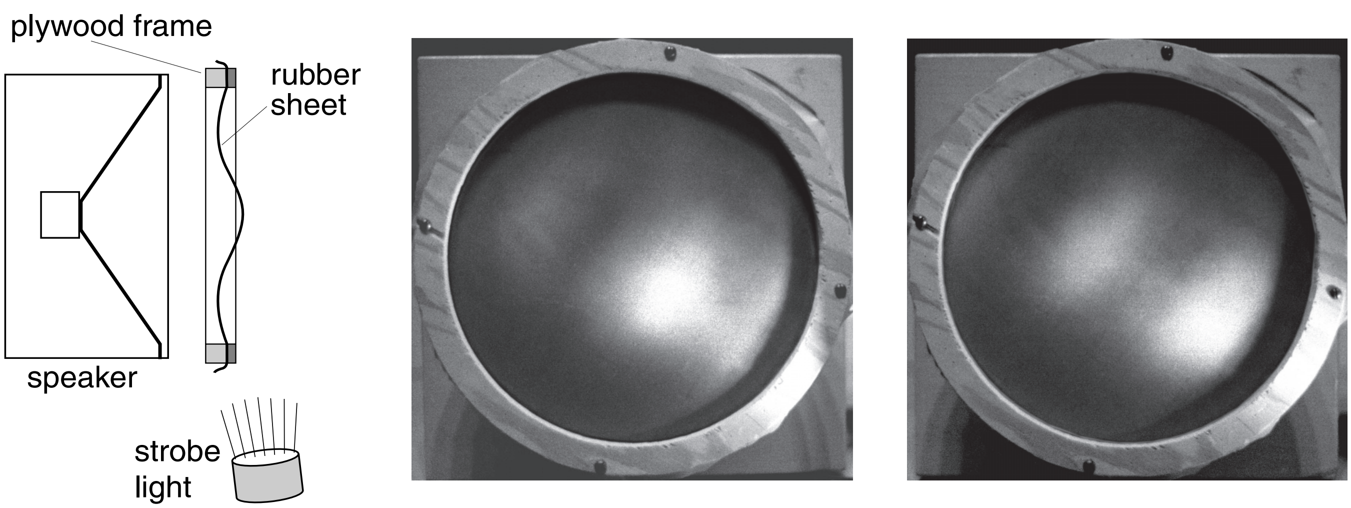


Figure 40
Studying the modes of oscillation of a drumhead.

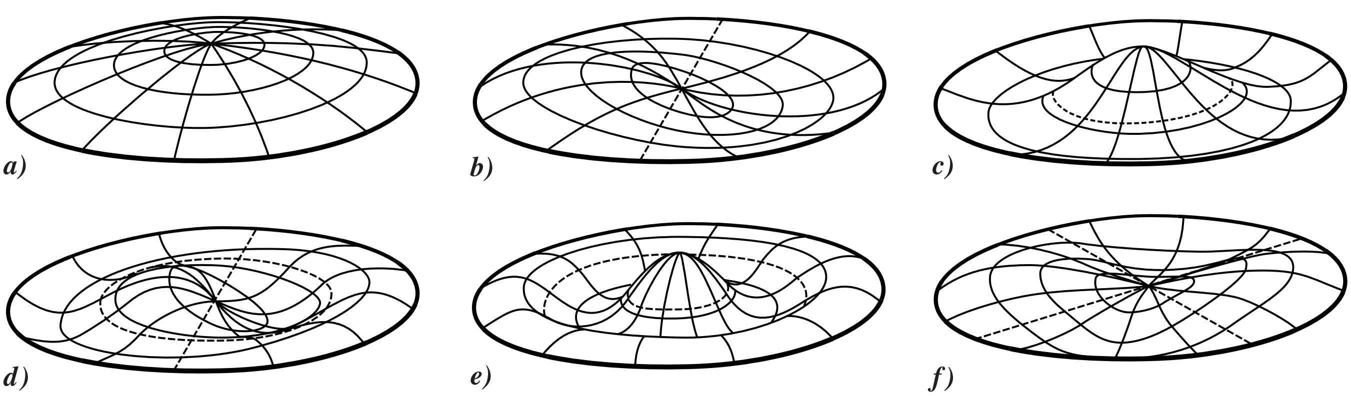


Figure 41
Modes of oscillation of a drumhead. (Adapted from* Vibration and Sound *by Phillip M. Morse, 2nd ed., McGraw-Hill, New York, 1948.)

Turn the frequency a bit off resonance, and you get a slow motion moving picture of the motion of that mode.

Some of the low frequency normal mode or standing wave patterns of the drumhead are illustrated in Figure (41). In the lowest frequency mode, Figure (41a), the entire center part of the drumhead moves up and down, much like a guitar string in its lowest frequency mode. In this pattern there are no nodes except at the rim of the drumhead.

In the next lowest frequency mode, shown in Figure (41b), one half the drumhead goes up while the other half goes down, again much like the second harmonic mode of the guitar string. The full two dimensional nature of the drumhead standing waves begins to appear in the next mode of Figure (41c) where the center goes up while the outside goes down. Now we have a circular node about half way out on the radius of the drumhead.

As we go up in frequency, we observe more complex patterns for the higher modes. In Figure (41d) we see a pattern that has a straight node like (41b) and a circular node like (41c). This divides the drumhead into 4 separate regions which oscillate opposite to each other. Finer division of the drumhead into smaller regions can be seen in Figures (41e) and (41f). The frequencies of the various modes are listed with each diagram. You can see that there is no obvious progression of frequencies like the harmonic progression for the modes of a stretched string.

When you strike a drumhead you excite a number of modes at once and get a complex mixture of frequencies. However, you do have some control over the modes you excite. Bongo drum players, for example, get different sounds depending upon where the drum is struck. Hitting the drum in the center tends to excite the lowest mode of vibration and produces a lower frequency sound. Striking the drum near the edges excites the higher harmonics, giving the drum a higher frequency sound.

Even more complex than the modes of vibration of a drumhead are those of the components of a violin. To construct a successful violin, the front and back plates of a violin must be tuned before assembly. Figure (42) shows a violin backplate under construction, while Figure (43) shows the first 6 modes of oscillation of a completed backplate. Note again that the resonant frequencies do not form a harmonic series.

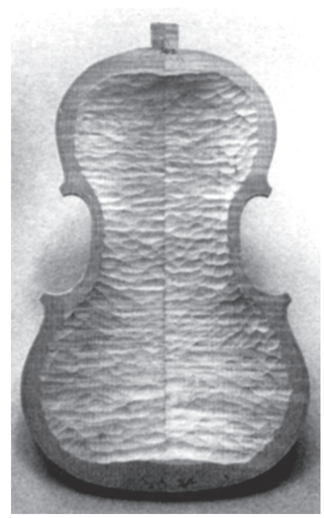
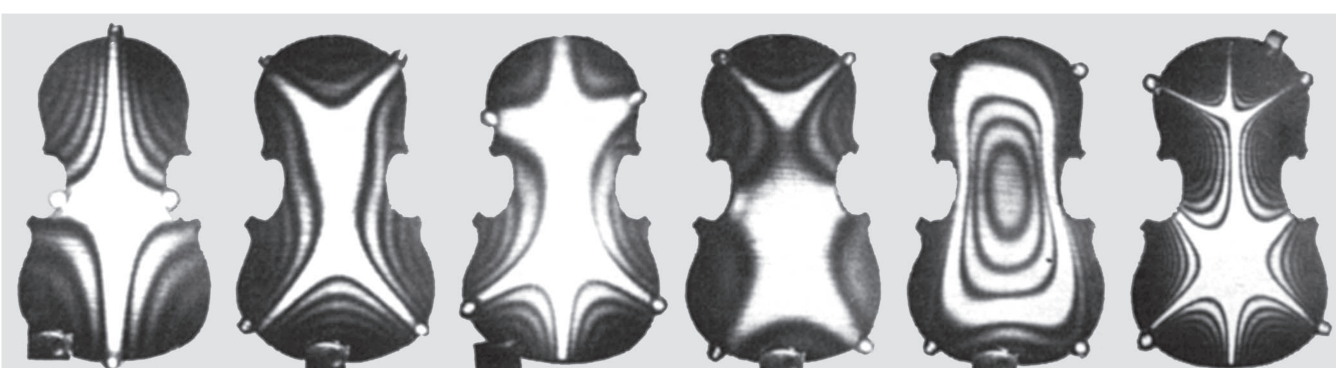


Figure 42
Back plate of a violin under construction. The resonant frequencies are tuned by carving away wood from different sections of the plate.



116Hz *167Hz* *222Hz* *230Hz* *349Hz* *403Hz*

Figure 43
Modes of oscillation of the backplate made visible by holographic techniques. Quality violins are made by tuning the frequencies of the various modes. (Figures 42 and 43 from "The Acoustics of Violin Plates," by Carleen Maley Hutchins, **Scientific American*****, October 1981.)***

SOUND INTENSITY

One of the amazing features of the human eardrum is its ability to handle an extreme range of intensities of sound waves. We define the ***intensity*** of a sound wave as the ***amount of energy per second being carried by a sound wave through a unit area***. In the MKS system of units, this would be the number of joules per second passing through an area of one square meter. Since one joule per second is a unit of power called a ***watt***, the MKS unit for sound intensity is ***watts per square meter***. The human ear is capable of detecting sound intensities as faint as 10^{-12} watts/m^2, but can also handle intensities as great as 1 watt/m^2 for a short time. This is an astounding range, a factor of 10^{12} in relative intensity.

The ear and brain handle this large range of intensities by essentially using a logarithmic scale. Imagine, for example, you are to sit in front of a hi fi set playing a pure tone, and you are told to mark off the volume control in equal steps of loudness. The first mark is where you just barely hear the sound, and the final mark is where the sound just begins to get painful. Suppose you are asked to divide this range of loudness into what you perceive as 12 equal steps. If you then measured the intensity of the sound you would find that the intensity of the sound increased by approximately a factor of 10 after each step. Using the faintest sound you can hear as a standard, you would measure that the sound was 10 times as intense at the end of the first step, 100 times as intense after the second step, 1000 times at the third, and 10^{12} times as intense at the final step.

The idea that the intensity increases by a factor of 10 for each equal step in loudness is what we mean by the statement that the loudness is based on a logarithmic scale. We take the faintest sound we can hear, an intensity $I_0 = 10^{-12}$watts/m^2 as a basis. At the first setting $I = I_0$, at the second setting $I_1 = 10\, I_0$, at the 3rd setting $I_2 = 100\, I_0$, etc. The factor I/I_0 by which the intensity has increased is thus

$$
\begin{aligned}
I_0/I_0 &= 1 \qquad\qquad (I_0 = 10^{-12}\text{ watts/m}^2)\\
I_1/I_0 &= 10\\
I_2/I_0 &= 100\\
&\cdots\cdots\\
I_{12}/I_0 &= 10^{12}
\end{aligned}
\tag{5}
$$

Taking the logarithm to the base 10 of these ratios gives

$$
\begin{aligned}
\mathrm{Log}_{10}(I_0/I_0) &= \mathrm{Log}_{10}(1) = 0\\
\mathrm{Log}_{10}(I_1/I_0) &= \mathrm{Log}_{10}(10) = 1\\
\mathrm{Log}_{10}(I_2/I_0) &= \mathrm{Log}_{10}(100) = 2\\
&\cdots\cdots\\
\mathrm{Log}_{10}(I_{12}/I_0) &= \mathrm{Log}_{10}(10^{12}) = 12
\end{aligned}
\tag{6}
$$

Bells and Decibels

The scale of loudness defined as $\mathrm{Log}_{10}(I/I_0)$ with $I_0 = 10^{-12}$watts/m^2 is measured in bells, named after Alexander Graham Bell, the inventor of the telephone.

$$
\boxed{\left.\begin{array}{l}\text{loudness of a sound}\\ \text{measured in bells}\end{array}\right\} \equiv \mathrm{Log}_{10}\left(\frac{I}{I_0}\right)}
$$
$$
I_0 = 10^{-12}\text{watts/m}^2 \tag{7}
$$

From this equation, we see that the faintest sound we can hear, at $I = I_0$, has a loudness of zero bells. The most intense one we can stand for a short while has a loudness of 12 bells. All other audible sounds fall in the range from 0 to 12 bells.

It turns out that the bell is too large a unit to be convenient for engineering applications. Instead one usually uses a unit called the ***decibel (db)*** which is 1/10 of a bell. Since there are 10 decibels in a bell, the formula for the ***loudness*** $\boldsymbol{\beta}$, in decibels, is

$$
\boxed{\beta(\text{decibels}) = (10\text{ db})\mathrm{Log}_{10}\left(\frac{I}{I_0}\right)} \tag{8}
$$

On this scale, the loudness of sounds range from 0 decibels for the faintest sound we can hear, up to 120 decibels (10×12 bells) for the loudest sounds we can tolerate.

The average loudness of some of the common or well known sounds is given in Table 1.

Table 1 Various Sound Levels in db

threshold of hearing	0
rustling leaves	10
whisper at 1 meter	20
city street, no traffic	30
quiet office	40
office, classroom	50
normal conversation at 1 meter	60
busy traffic	70
average factory	80
jack hammer at 1 meter	90
old subway train	100
rock band	120
jet engine at 50 meters	130
Saturn rocket at 50 meters	200

An increase in loudness of 10 db corresponds to an increase of 1 bell, or an increase of intensity by a factor of 10. A rock band at 110 db is some 100 times as intense (2 factors of 10) as a jack hammer at 90 db. A Saturn rocket is about 10^{20} times as intense as the faintest sound we can hear.

Our sensitivity to sound depends not only to the intensity of the sound, but also to the frequency. About the lowest frequency note one can hear, and still perceive as being sound, is about 20 cycles/second. As you get older, the highest frequencies you can hear decreases from around 20,000 cycles/sec for children, to 15,000 Hz for young adults to under 10,000 Hz for older people. If you listen to too much, too loud rock music, you can also decrease your ability to hear high frequency sounds.

Figure (44) is a graph of the average range of sound levels for the human ear. The faintest sounds we can detect are in the vicinity of 4000 Hz, while any sound over 120 db is almost uniformly painful. The frequency ranges and sound levels usually encountered in music are also shown.

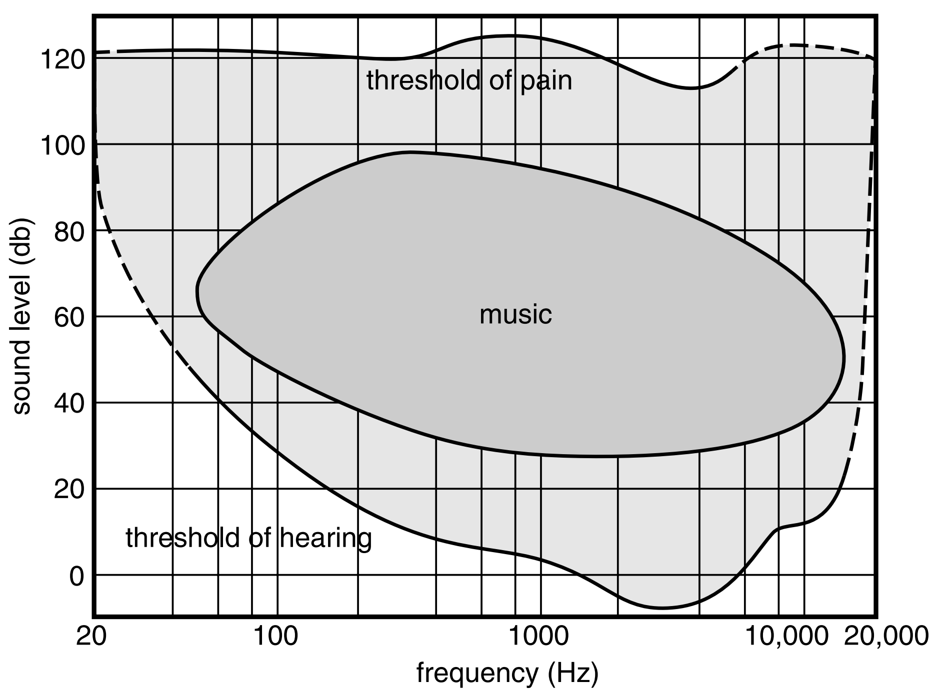


Figure 44
Average range of sound levels for the human ear. Only the very young can hear sound frequencies up to 20,000 Hz. (Adapted from* Fundamental Physics *by Halliday and Resnick, John Wiley & Sons.)

Sound Meters

Laboratory experiments involving the intensity or loudness of sound are far more difficult to carry out than those involving frequencies like the Fourier analysis experiments already discussed. From the output of any reasonably good microphone, you can obtain a relatively good picture of the frequencies involved in a sound wave. But how would you go about determining the intensity of a sound from the microphone output? (There are commercial sound meters which have a scale that shows the ambient sound intensity in decibels. Such devices are often owned by zoning boards for checking that some factory or other noise source does not exceed the level set by the local zoning ordinance, often around 45 db. The point of our question is, how would you calibrate such a device if you were to build one?)

The energy in a wave is generally proportional to the square of the amplitude of the wave. A sound wave can be viewed as oscillating pressure variations in the air, and the energy in a sound wave turns out to be proportional to the square of the amplitude of the pressure variations. The output of a microphone is more or less proportional to the amplitude of the pressure variations, thus we expect that the intensity of a sound wave should be more or less proportional to the square of the voltage output of the microphone. However, there is a great variation in the sensitivity of different microphones, and in the amplifier circuits used to produce reasonable signals. Thus any microphone that you wish to use for measuring sound intensities has to be calibrated in some way.

Perhaps the easiest way to begin to calibrate a microphone for measuring sound intensities is to use the fact that very little sound energy is lost as sound travels out through space. Suppose you had a speaker radiating 100 watts of sound energy, and for simplicity let us assume that the speaker radiates uniformly in all directions and that there are no nearby walls.

If we are 1 meter from this speaker, all the sound energy is passing out through a 1 meter radius sphere centered on the speaker. Since the area of a sphere is $4\pi r^2$, this 1 meter radius sphere has an area of $4\pi \text{meters}^2$, and the average intensity of sound at this 1 meter distance must be

$$\left.\begin{array}{l}\text{average intensity of}\\ \text{sound 1 meter from}\\ \text{a 100 watt speaker}\end{array}\right\} = \frac{100\text{ watts}}{4\pi\text{ meters}^2} = 8.0\,\frac{\text{watts}}{\text{meters}^2} \tag{9}$$

If we wish to convert this number to decibels, we get

$$\begin{aligned}\beta\begin{pmatrix}\text{sound intensity 1 meter}\\ \text{from a 100 watt speaker}\end{pmatrix} &= (10\text{ db})\operatorname{Log}\frac{I}{I_0}\\ &= (10\text{ db})\operatorname{Log}_{10}\left(\frac{8.0\text{ watts / m}^2}{10^{-12}\text{ watts / m}^2}\right)\\ &= (10\text{ db})\times\operatorname{Log}_{10}\left(8\times 10^{12}\right)\\ &= (10\text{ db})\times 12.9\\ &= 129\text{ db}\end{aligned} \tag{10}$$

From our earlier discussion we see that this exceeds the threshold of pain. One meter from a 100 watt speaker is too close for our ears. But we could place a microphone there and measure the amplitude of the signal output for our first calibration point.

Move the speaker back to a distance of 10 meters and the area that the sound energy has to pass through increases by a factor of 100 since the area of a sphere is proportional to r^2. Thus as the same 100 watts passes through this 100 times larger area, the intensity drops to 1/100 of its value at 1 meter. At a distance of 10 meters the intensity is thus $8/100 = .08$ watts/m^2 and the loudness level is

$$\beta\begin{pmatrix}10 \text{ meters from a} \\ 100 \text{ watt speaker}\end{pmatrix}$$

$$= (10\text{ db})\text{Log}_{10}\left(\frac{.08\text{ watts / m}^2}{10^{-12}\text{ watts / m}^2}\right)$$

$$= 109\text{ db} \qquad (11)$$

We see that when the intensity drops by a factor of 100, it drops by 20 db or 2 bells.

To calibrate your sound meter, record the amplitude of the signal on your microphone at this 10 meter distance, then set the microphone back to a distance of 1 meter, and cut the power to the speaker until the microphone reads the same value as it did when you recorded 100 watts at 10 meters. Now you know that the speaker is emitting only 1/100th as much power, or 1 watt. Repeating this process, you should be able to calibrate a fair range of intensities for the microphone signal. If you get down to the point where you can just hear the sound, you could take that as your value of I_0, which should presumably be close to $I_0 = 10^{-12}$watts/m^2. Then calibrate everything in db and you have built a loudness meter. (The zoning board, however, might not accept your meter as a standard for legal purposes.)

Exercise 2

What is the loudness, in db, 5 meters from a 20 watt speaker? (Assume that the sound is radiated uniformly in all directions).

Exercise 3

You are playing a monophonic record on your stereo system when one of your speakers cuts out. How many db did the loudness drop? (Assume that the intensity dropped in half when the speaker died. Surprisingly you can answer this question without knowing how loud the stereo was in the first place. The answer is that the loudness dropped by 3 db).

Speaker Curves

When you buy a hi fi loudspeaker, you may be given a frequency response curve like that in Figure (45), for your new speaker. What the curve measures is the intensity of sound, at a standard distance, for a standard amount of power input at different frequencies. It is a fairly common industry standard to say that the frequency response is "flat" over the frequency range where the intensity does not fall more than 3 db from its average high value. In Figure (45), the response of that speaker, with the woofer turned on, is more or less "flat" from 62 Hz up to 30,000 Hz.

Why the 3 db cutoff was chosen, can be seen in the result of Exercise 3. There you saw that ***if you reduce the intensity of the sound by half, the loudness drops by 3 db.*** This is only 3/120 (or 1/36) of our total hearing range, not too disturbing a variation in what is supposed to be a flat response of the speaker.

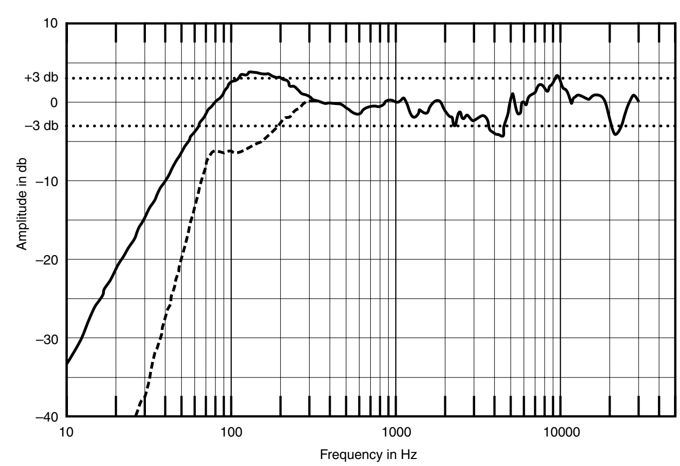


Figure 45
Speaker response curve from a recent audio magazine. The dashed line shows the response when the woofer is turned off. (We added the dotted lines at + and – 3 db.)

APPENDIX A
FOURIER ANALYSIS LECTURE

In our discussion of Fourier Analysis, we saw that any wave form can be constructed by adding together a series of sine and cosine waves. You can think of the Fourier transform as a mathematical prism which breaks up a sound wave into its various wavelengths or frequencies, just as a light prism breaks up a beam of white light into its various colors or wavelengths. In MacScope, the computer does the calculations for us, figuring out how much of each component sine wave is contained in the sound wave. The point of this lecture is to give you a feeling for how these calculations are done. The basic ideas are easy, only the detailed calculations that the computer does would be hard for us to do.

Square Wave

In Figure A-1 we show a MacScope window for a square wave produced by a Hewlett Packard oscillator.

We have selected precisely one cycle of the wave, and see that the even harmonics are missing. A careful investigation shows that the amplitude of the Nth odd harmonic is 1/N as big as the first (e.g., the 3rd harmonic is 1/3 as big as the first, etc.). Thus the mathematical formula for a square wave F(t) can be written:

$$F(t) = (1)\sin(t) + (1/3)\sin(3t) + (1/5)\sin(5t) + (1/7)\sin(7t) \; ...$$

where, for now, we are assuming that the period of the wave is precisely 2π seconds. The coefficients (1), (1/3), (1/5), (1/7), which tell us how much of each sine wave is present, are called the ***Fourier coefficients***. Our goal is to calculate these coefficients.

Calculating Fourier Coefficients

In general we cannot construct an arbitrary wave out of just sine waves, because sine waves, sin(t), sin(2t), etc., all have a value 0 at $t = 0$ and at $t = 2\pi$. If our wave is not zero at the beginning ($t = 0$) of our selected period, or not zero at the end ($t = 2\pi$), then we must also include cosine waves which have a value 1 at those points. Thus the general formula for breaking an arbitrary repetitive wave into sine and cosine waves is:

$$\begin{aligned} F(t) = A_0 &+ A_1\cos(1t) + A_2\cos(2t) \\ &+ A_3\cos(3t) + \; ... \\ &+ B_1\sin(1t) + B_2\sin(2t) \\ &+ B_3\sin(3t) + \; ... \end{aligned} \qquad \text{(A-1)}$$

The question is: How do we find the coefficients A_0, A_1, A_2, B_1, B_2 etc. in Equation (A1)? (These are the ***Fourier coefficients.***)

To see how we can determine the Fourier coefficients, let us take an explicit example. Suppose we wish to find the coefficient B_3, representing the amount of sin(3t) present in the wave. We can find B_3 by first multiplying Equation (1) through by sin(3t) to get:

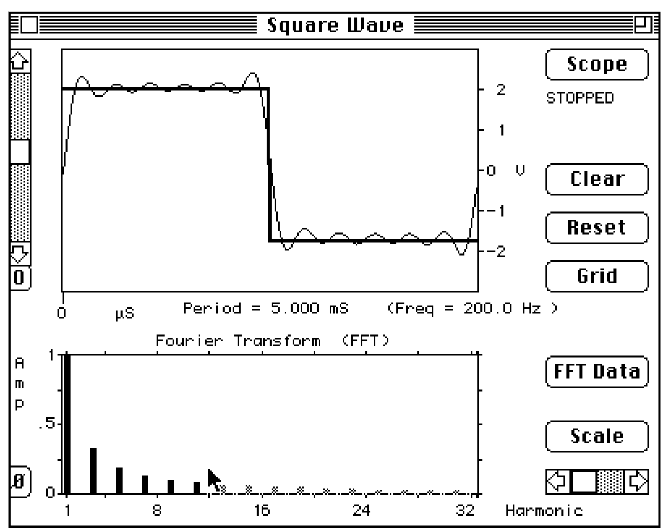


Figure A-1
The square wave has only odd harmonics.

$$
\begin{aligned}
F(t)\sin(3t) = \; & A_0 \sin(3t) \\
& + A_1 \cos(1t)\sin(3t) \\
& + A_2 \cos(2t)\sin(3t) \\
& + A_3 \cos(3t)\sin(3t) \\
& + \ldots \\
& + B_1 \sin(1t)\sin(3t) \\
& + B_2 \sin(2t)\sin(3t) \\
& + B_3 \sin(3t)\sin(3t) \\
& + \ldots
\end{aligned}
\qquad \text{(A-2)}
$$

At first it looks like we have created a real mess. We have a lot of products like cos(t)sin(3t), sin(t)sin(3t), sin(3t)sin(3t), etc. To see what these products look like, we plotted them using **True BASIC™** and obtained the results shown in Figure (A2) (on the next page).

Notice that in all of the plots involving sin(3t), the product $\sin(3t)\sin(3t) = \sin^2(3t)$ is special; it is the only one that is always positive. (It has to be since it is a square.) A careful investigation shows that, in all the other non square terms, there is as much negative area as positive area, as indicated by the two different shadings in Figure (A3). If we define the "net area" under a curve as *the positive area minus the negative area*, then only the sin(3t)sin(3t) term on the right side of the Equation A-2 has a net area.

The mathematical symbol for finding the net area under a curve (in the interval $t = 0$ to $t = 2\pi$) is:

$$
\left.\begin{array}{l}\text{Area under} \\ \sin(3t)\sin(3t) \\ \text{from } t = 0 \\ \text{to } t = 2\pi\end{array}\right\} = \int_0^{2\pi} \sin(3t)\sin(3t)\,dt \qquad \text{(A-3)}
$$

(Those who have had calculus say we are taking the **integral** of the term sin(3t)sin(3t).)

A basic rule learned in algebra is that if we do the same thing to both sides of an equation, the sides will still be equal. This is also true if we do something as peculiar as evaluating the net area under the curves on both sides of an equation.

If we take the net area under the curves on the right side of Equation A-2, only the $\sin^2(3t)$ term survives and we get:

$$
\int_0^{2\pi} F(t)\sin(3t)\,dt = B_3 \int_0^{2\pi} \sin^2(3t)\,dt \qquad \text{(A-4)}
$$

In Figure (A4), we have replotted the curve $\sin^2(3t)$, and drawn a line at height $y = .5$. We see that the peaks above the $y = .5$ line could be flipped over to fill in the valleys below the $y = .5$ line. Thus $\sin^2(3t)$ in the interval $t = 0$ to $t = 2\pi$ has a net area equal to that of a rectangle of height .5 and length 2π. I.e., the net area is π:

$$
\int_0^{2\pi} \sin^2(3t)\,dt = \pi \qquad \text{(A-5)}
$$

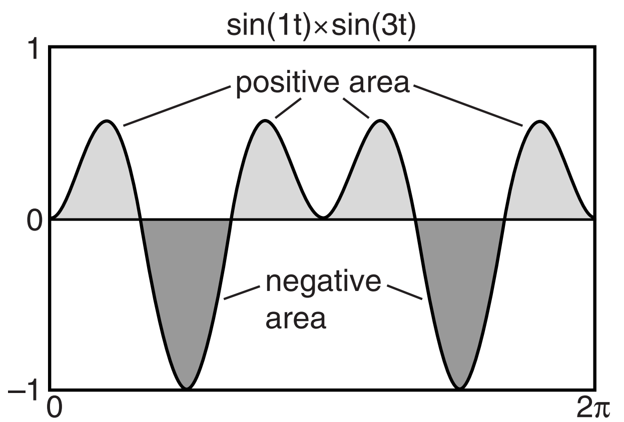


Figure A-3
Only the square terms such as sin(3t)sin(3t) have a net area. This curve has no net area.

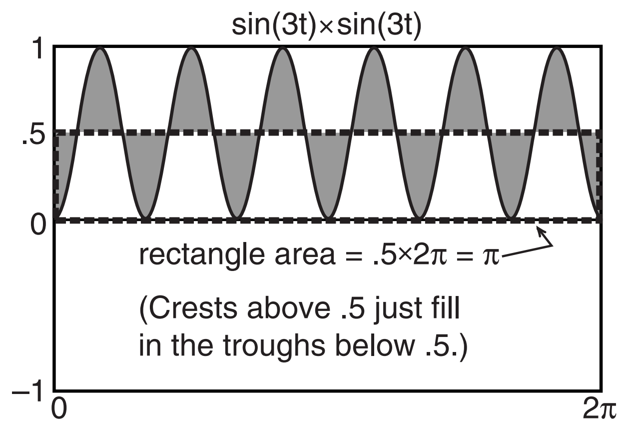


Figure A-4
The area under the curve $\sin^2(3t)$ is equal to the area of a rectangle .5 high by 2π long.

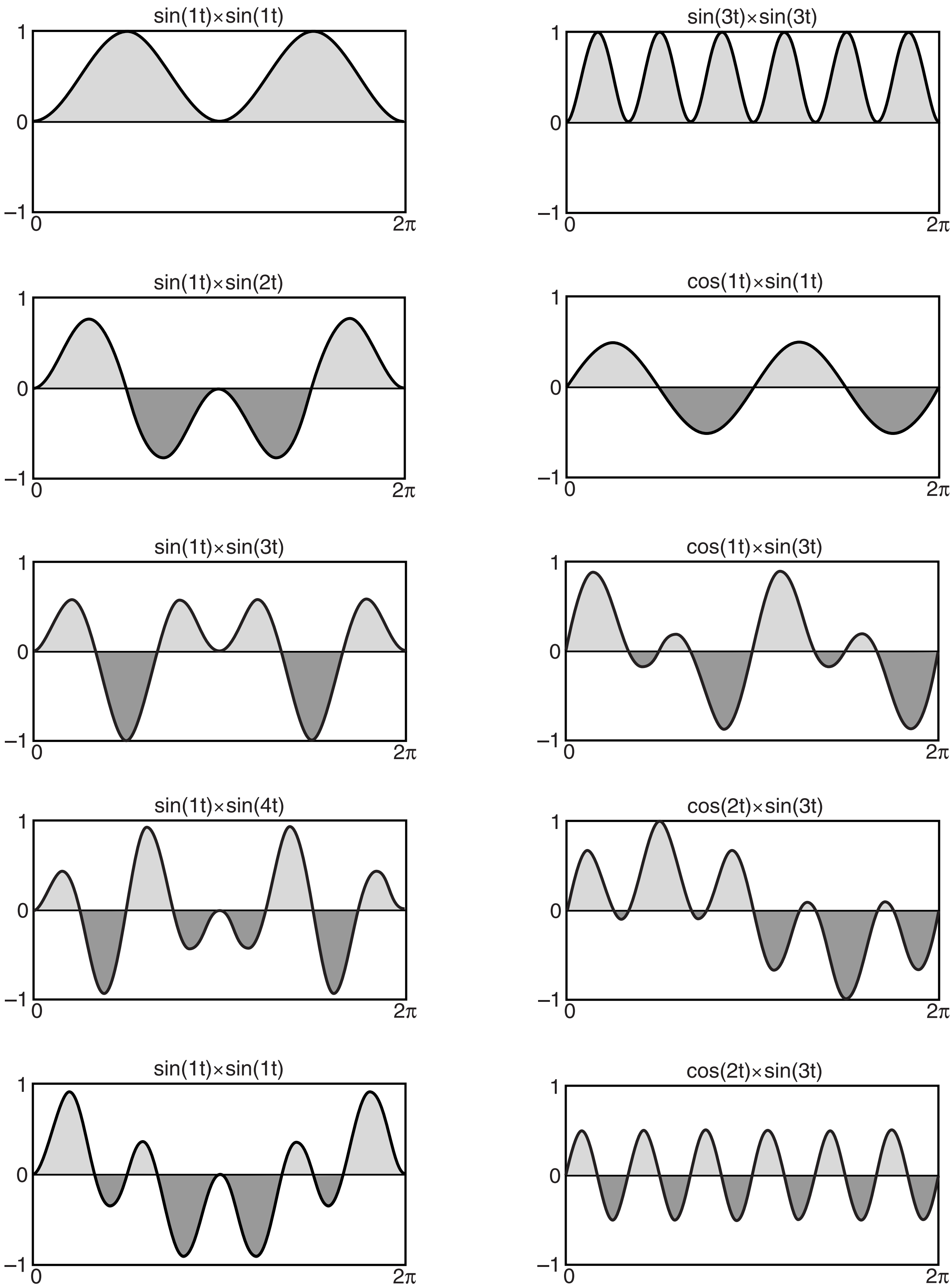


Figure A-2
Product wave patterns.

Substituting Equation A-5 in Equation A-4 and solving for the Fourier coefficient B_3 gives:

$$B_3 = \frac{1}{\pi}\int_0^{2\pi} F(t)\sin(3t)\,dt \qquad \text{(A-6)}$$

Similar arguments show that the general formulas for the Fourier coefficients A_n and B_n are:

$$\boxed{A_n = \frac{1}{\pi}\int_0^{2\pi} F(t)\cos(nt)\,dt} \qquad \text{(A-7)}$$

$$\boxed{B_n = \frac{1}{\pi}\int_0^{2\pi} F(t)\sin(nt)\,dt} \qquad \text{(A-8)}$$

These integrals, which were nearly impossible to do before computers, are now easily performed even on small personal computers. Thus the computer has made Fourier analysis a practical experimental tool.

Amplitude and Phase

Instead of writing the Fourier series as a sum of separate sine and cosine waves, it is often more convenient to use amplitudes and phases. The basic formula we use is **

$$A\cos(t) + B\sin(t) = C\cos(t-\phi) \qquad \text{(A-9)}$$

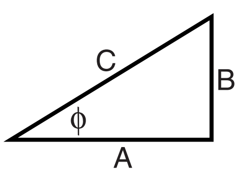
where

C = amplitude

$C^2 = A^2 + B^2$

ϕ = phase

$\tan(\phi) = B/A$



** Start with

$\cos(x-y) = \cos(x)\times\cos(y) + \sin(x)\times\sin(y)$

Let $x = t$, $y = \phi$, and multiply through by C to get

$C\cos(t-\phi) = \{C\cos(\phi)\}\cos(t) + \{C\sin(\phi)\}\sin(t)$

This is Equation A–9 if we set

$A = \{C\cos(\phi)\}$; $B = \{C\sin(\phi)\}$

Thus $\tan\phi = \sin\phi/\cos\phi = B/A$

The function $\cos(t-\phi)$ is illustrated in Figure (A-5). We see that C is the amplitude of the wave, and the phase angle ϕ is the amount the wave has been moved to the right. (When $t=0$, $\cos(t-\phi) = \cos(-\phi)$.) With Equation A-9, we can rewrite equation A-1 in the form:

$$\begin{aligned} F(t) = C_0 &+ C_1\cos(\ t-\phi_1) \\ &+ C_2\cos(2t-\phi_2) \\ &+ C_3\cos(3t-\phi_3) \\ &+ \ldots \end{aligned} \qquad \text{(A-10)}$$

The advantage of Equation A-10 is that the coefficients C represent how much of each wave is present, and sometimes we do not care about the phase angle ϕ. For example our ears are not particularly sensitive to the phase of the harmonics in a musical note, thus the tonal quality of a musical instrument is determined almost entirely by the amplitudes C of the harmonics the instrument produces.

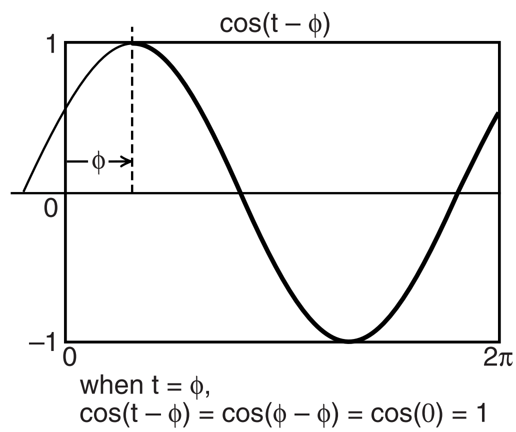


Figure A-5
The function* $\cos(t-\phi)$*. When* $\phi = 90°$ *you get a sine wave.

In the Fourier transform plots we have shown so far, the graph of the harmonics has been representing the amplitudes C. If you wish to see a plot of the phases ϕ, then press the button labeled ø as shown, and you get the result seen in Figure(A-6). In that figure we are looking at the phases of the odd harmonic sine waves that make up a square wave. Since

$$\sin(t) \;=\; \cos(t - 90°)$$

all the sine waves should have a phase shift of 90°.

If for any reason, you need accurate values of the Fourier coefficients, they become available if you press the **FFT Data** button to get the results shown in Figure (A-7). When you do this, the **Editor** window is filled with a text file containing the A, B coefficients accurate to 3 or 4 significant figures.

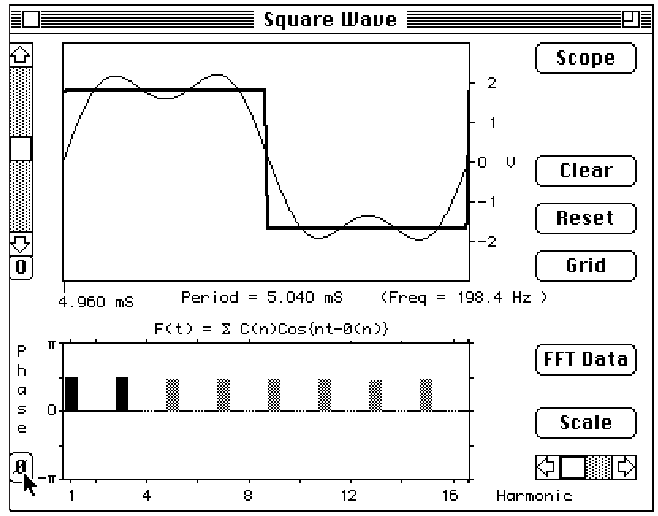


Figure A-6
When you press the ø button, the Fourier Transform display shifts from amplitudes to phases. Since the square wave is made up of pure odd harmonic sine waves, each odd harmonic should have a 90 degree or $\pi/2$ phase shift.

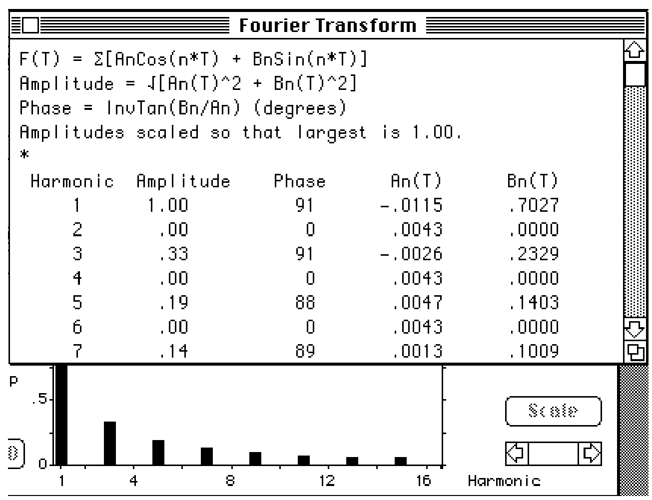


Figure A-7
For greater numerical accuracy, you can press the FFT Data button. This gives you a text file with the A, B coefficients given to four places.

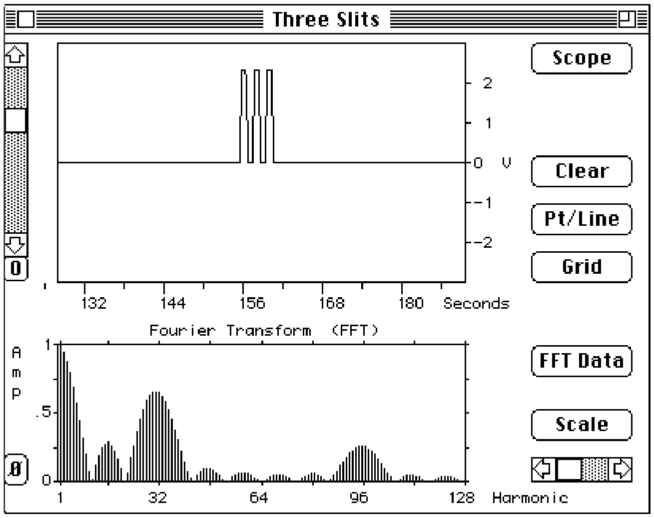


Figure A-8
Amplitude. The Fourier transform of a 3-slit pattern gives the amplitude of the diffraction pattern that would be produced by a laser beam passing through these slits. (Selecting the data to give wider slits would correspond to using a different wavelength laser beam.)

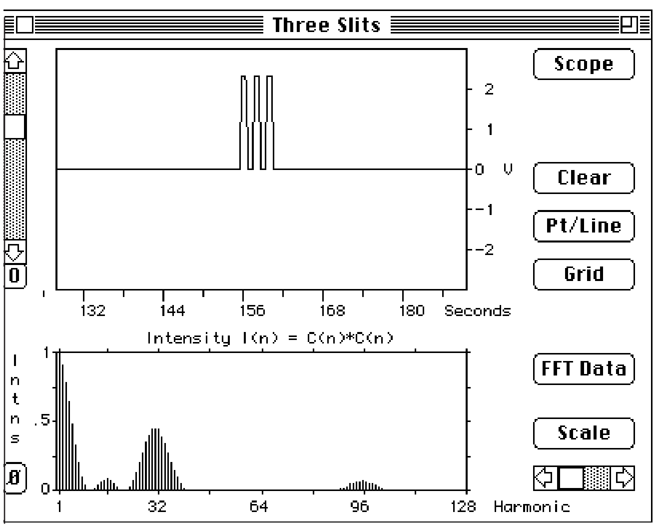


Figure A-9
Intensity. In the lab, you see the intensity of the diffraction pattern. MacScope will display the intensity of the Fourier transform if you click one more time on the ø button.

Amplitude and Intensity

An experiment that has become possible with MacScope, is to have students compare the Fourier transform of a multiple slit grating with the diffraction pattern produced by a laser beam passing through that grating. For example, in Figure (A-8) we have taken the transform of a 3-slit grating. In this case, the 3-slit "pattern" was made simply by turning a 2 volt power supply on and off. We are now working on ways for students to record the slit pattern directly.

The problem with Figure (A-8) is that the Fourier transform of the slits gives the amplitude of the diffraction pattern, while in the lab one measures the intensity of the diffraction pattern. The intensity of a light wave is proportional to the square of its amplitude.

In order that students can compare a slit Fourier transform with an experimental diffraction pattern, we have designed MacScope so that one more press on the ø button takes us from a display of phases to intensities. Explicitly, the ø button cycles from *amplitudes* to *phases* to *intensities*. In Figure (A-9) we have clicked over to intensities, and this pattern may be directly compared with the intensity of the 3-slit diffraction pattern seen in lab.

(The Fourier transform of the diffraction pattern amplitude should give the slit pattern. Unfortunately, if you take the Fourier transform of the experimental diffraction pattern, you are taking the transform of the intensity, or square, of the amplitude. What you get, as Chris Levey of our department demonstrated, is the *convolution* of the slit pattern with itself.)

APPENDIX B
INSIDE THE COCHLEA

In Figure (32), our simplified unwound view of the cochlea, we show only the basilar membrane separated by two fluid channels (the scala vestibuli which starts at the oval window, and the scala tympani which ends at the round window). That there is much more structure in the cochlea is seen in the cochlea cross section of Figure (B-1). The purpose of this additional structure is to detect the motion of the basilar membrane in a way that is sensitive to the harmonic content of the incoming sound wave.

Recall that when the basilar membrane is excited by a sinusoidal oscillation, the maximum amplitude of the response of the basilar membrane is located at a position that depends upon the frequency of the oscillation.

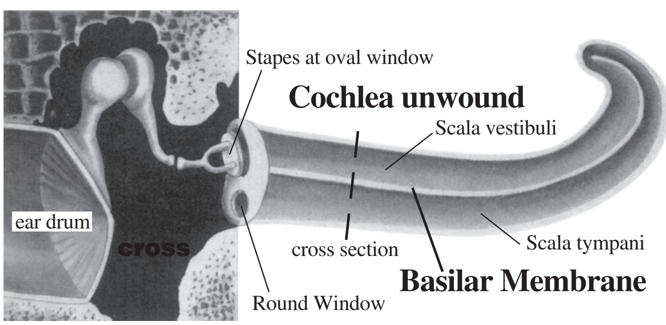


Figure 32 (repeated)
The cochlea unwound.

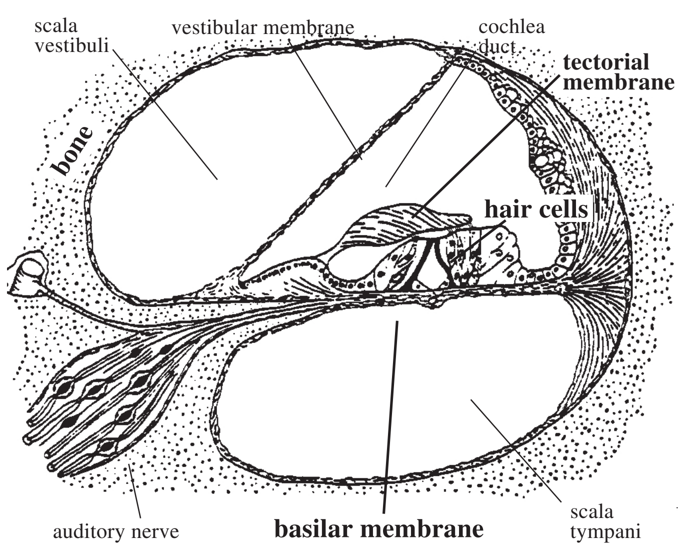


Figure B-1
Cross section of the cochlea. *(From Vander,A; Sherman,J; and Luciano,D.* **Human Physiology**, *4th edition, 1985, P662. McGraw Hill Publishing Co., NY.)*

As seen in Figure (35), the lower the frequency, the farther down the membrane the maximum amplitude occurs. Along the top of the basilar membrane is a system of hair cells that detects the motion of the membrane and sends the needed information to the brain.

Figure (B-2a) is a close up view of the hair cells that sit atop of the basilar membrane. (There are about 30,000 hair cells in the human ear.) Above the hair cells is another membrane called the tectorial membrane which is hinged on the left hand side of that figure. Fine hairs go from the top of each hair cell up to the tectorial membrane as shown. When the basilar membrane is deflected by an incoming sound wave the hairs are bent as shown in Figure (B-2b). It is the bending of the hairs that triggers an electrical impulse in the hair cell.

Figure (B-3) is a mechanical model of how the bending of the hairs creates the electrical impulse. The fluid in the cochlea duct surrounding the hair cells has a high concentration of positive potassium ions (k^+). The

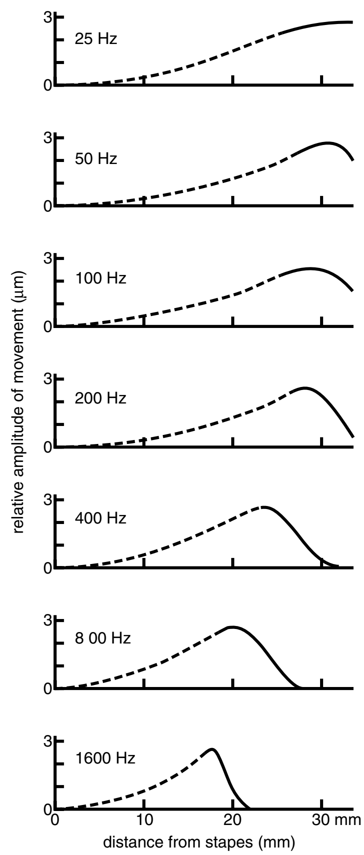


Figure 35 (repeated)
Amplitude of the motion of the basilar membrane at different frequencies, as we go down the basilar membrane.

bending of the hair cell opens small channels allowing potassium ions to flow into the hair cell. This flow of positive charge into the cell changes the electrical potential of the cell, triggering reactions that will eventually result in an electrical impulse in the nerve fiber that is connected to the hair cell.

After the channel at the top of the hair cell closes, the excess potassium is pumped out of the hair cell, and the cell returns to its normal resting voltage, ready to fire again.

There are various ways that a hair cell can transmit frequency information to the nervous system. One is by its location down the basilar membrane. The lower the frequency of the sound wave, the farther down the membrane an oscillation of the membrane takes place. Thus high frequency waves excite cells at the front of the basilar membrane, while low frequency oscillations excite cells at the back end.

Secondly, hair calls in a given area show special sensitivities to different frequencies. Figure (B-4) shows the amplitude, in db, of the sound wave required to excite a nerve fiber connected to that particular region of hair cells. You can see that the nerve is most sensitive to a 2 killocycle (2kHz) frequency. At 2 kHz, that nerve fires when excited by a 15 db sound wave. It is not excited by a 4 kHz wave until the sound intensity rises to 80 db.

Ultimately the exquisite sensitivity of the human ear to different frequency components in a sound wave results from the fact that there are about 30,000 hair cells continuously monitoring the motion of the basilar membrane. Effective processing of this vast amount of information leads to the needed sensitivity. Much of this processing of information occurs in the nervous system in the ear, before the information is sent to the brain.

Figure B-3
Model of the valves at the base of the hair cells.
(From Shepard,G.M., Neurobiology, 3rd Edition, 1994, P316. Oxford Univ. Press.)

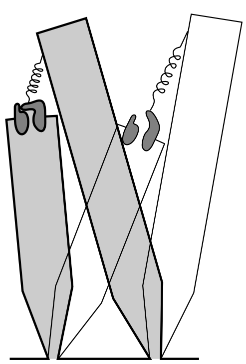
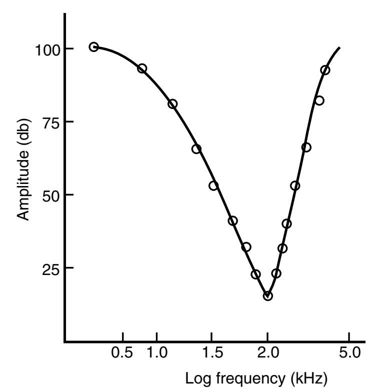


Figure B-4
Frequency dependence. A much lower amplitude sound will excite this nerve fiber at 2kHz than any other frequency. Different nerve fibers connected to the hair cells have different frequency dependence.



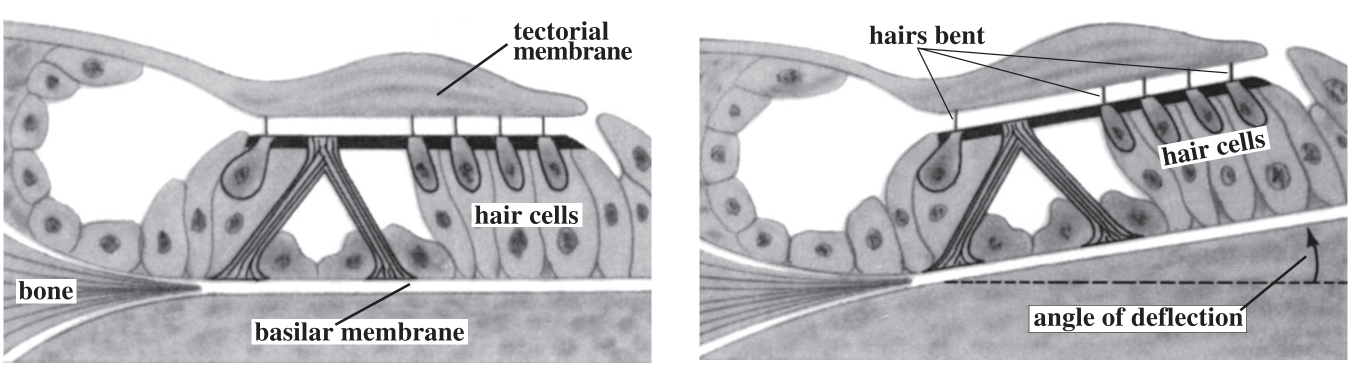


Figure B-2 a,b
When the basilar membrane is deflected by a sound wave, the hair cells are bent. This opens a valve at the base of the hair call. ***(Figures 2 & 4 adapted from Kandel, E; Schwartz, J; and Jessell, T;*** **Principles of Neural Science*****, 3rd Edition, 1991; pages 486 and 489.)***

Index