PY231: Notes on Linear and Nonlinear Oscillators, and Periodic Waves

B. Lee Roberts
Department of Physics
Boston University
DRAFT January 2011

1 The Simple Oscillator

In many places in music we encounter systems which can oscillate. If we understand such a system once, then we know all about any other situation where we encounter such a system. Our simple example system is a mass on a spring. It could also be a cork floating in water, coffee sloshing back and forth in a cup of coffee, or any number of other simple systems.

No system in the macroscopic world is a simple oscillator. Dissipative forces are always present. Sometimes they can be ignored, and for systems where the damping is light, we can understand much of the behavior by analogy to the simple oscillator. These notes are divided into four sections: the simple (undamped) oscillator; the damped oscillator, the driven damped oscillator, a discussion of nonlinear oscillators and systems, and finally a discussion of the related topic periodic waves.

Occasionally in this note, I have chosen to write a few equations. These equations will provide a mathematical statement of what is being described. If you don’t like equations, I also state in words what is going on.

Consider the mass on a spring as sketched in Fig. 1. We ignore the effect of gravity since

\[ y = 0 \]

Fig. 1: The Simple Oscillator
all it does is shift the equilibrium point. We have seen in class that when pulled down, the mass will oscillate about the point where it was originally at rest, which is often called the equilibrium position.

We say that the ideal spring provides a linear restoring force, that means that it pulls the mass back towards its equilibrium point with a force which is proportional to how far you stretch it. If we choose \( y = 0 \) to be the equilibrium point, then the force the spring exerts on the mass is given by

\[
F = -K y
\]

This is a linear restoring force (the \(-\) sign tells us the force pushes or pulls the mass back towards the equilibrium point). \( K \) is a constant which tells how stiff the spring is.

**aside:** If the force included a term like \( y^2 \) or \( y^3 \) then it would be a much more difficult problem to solve. We will briefly discuss such nonlinear forces at the end of these notes. Because all of the physical systems which appear in musical acoustics are nonlinear, we will have to address this issue.

---

**Fig. 2:** The motion of a simple oscillator. We show the vertical position as a function of time assuming that at the beginning the mass was released from the maximum position in the positive direction. The time when the mass has finished one cycle (period) is indicated by \( P \).

If we used the laws of physics to obtain the equation which described the motion of the mass on a spring, and its solution, we would find several simple facts about the motion of the mass:

1. The mass bounces up and down. We say that the motion is sinusoidal in time which is sketched in Fig. 2. This means that if we drew a graph of the position as a function of time, we would obtain a graph

\[
y = A \cos(2\pi f_0 t)
\]

where the frequency is given by

\[
f_0 = \frac{1}{2\pi} \sqrt{\frac{K}{m}}
\]

and the amplitude \( A \) is just the maximum vertical distance the mass travels.
2. There is only one frequency with which this system can oscillate, and that frequency is $f_0$. The period of the motion is just the inverse of the frequency, i.e. $P = 1/f_0$.

3. The frequency does not depend on the amplitude, or how the system is set into motion, but only on how stiff the spring is, and how large the mass is.

4. If we increase the mass (inertia) of the system, we lower the frequency it will vibrate with.

5. If we increase the stiffness of the spring, we will increase the frequency.

### 1.1 The Undamped Oscillator and Energy

If we stretch the spring, we have to do work against the spring force

$$F = -Ky$$

The amount of work depends on the distance squared, and is given by

$$W = \frac{1}{2}Ky^2$$

If we pull the spring out to a distance $y = A$, then the work which we do is $W = (1/2)KA^2$. This is the amount of energy that we put into the oscillator, and we conclude that the total energy of the oscillator is proportional to the square of the amplitude of the oscillation.

After we have pulled the spring back, we have given the mass-spring system some potential energy, which is energy which exists by virtue of the configuration (shape) of the system.

If we release the mass after pulling it aside, the spring will accelerate the mass, and it will return to the equilibrium position, but at that moment, where the potential energy is zero, the potential energy will have been transformed into kinetic energy (energy of motion) which is given by $(1/2)mv^2$ where $v$ is the velocity.

Since the mass has inertia associated with it, it will keep moving in the same direction past the equilibrium point. After the mass passes the equilibrium point, the spring starts to decelerate the mass and it will stop when it reaches a distance $A$ on that side of the equilibrium point. During this part of the cycle the spring is doing work on the mass, turning the kinetic energy into potential energy again.

The spring then accelerates the mass back in the direction it came from, changing the potential energy into kinetic energy.

Note that at any point except $y = 0$, which is the equilibrium position, or $y = \pm A$ which is the very top or bottom of the motion, the energy of the system is a mixture of kinetic and potential energy. Note also, that the total mechanical energy (potential plus kinetic) is a constant at any time during the motion. This is often called conservation of energy.

In Fig. 3 the total energy, the kinetic energy and the potential energy are shown as a function of time for the oscillator which starts its motion from rest at $y = A$.

The total mechanical energy is a constant which only depends on the maximum displacement and how stiff the spring is. It is a constant because there is no mechanism in the simple oscillator to dissipate the energy. The real oscillator will, of course, have damping and we will study this in the next section.
The Damped Oscillator

1.2 Summary of the Simple Oscillator

A linear restoring force leads to simple harmonic motion, which occurs at a frequency determined by the square root of the spring stiffness divided by the mass. The period of the motion is given by the inverse of the frequency. The amplitude of the oscillation is constant, since there is no way for energy to leave the system, since we said that for the simple oscillator there are no dissipative forces.

The example of the simple oscillator which we used was the mass on a spring. Another (almost simple) example which we have all seen is a swing, like those found on all playgrounds. Just as a clock pendulum, the rider swings back and forth with a frequency which, for small amplitudes, only depends on the acceleration of gravity and the length of the swing. We return to this example below when we discuss the driven oscillator.

Up to this point we have only discussed simple oscillators with one way to vibrate. i.e. they have only one frequency. However, there are many examples of oscillators which have more than one way to vibrate. All musical systems, e.g. bars, drum heads, strings, air columns in tubes, even the coffee in your cup have more than one way to vibrate. Nevertheless, the simple ideas developed above are still applicable.

2 The Damped Oscillator

The ideal oscillator discussed above does not really exist on the macroscopic scale. Real oscillators always experience a damping force, often one proportional to velocity. The prototypical system is a mass suspended on a spring but with a damper which is suspended in a viscous fluid.\footnote{The viscosity of a fluid tells us how much it resists something moving through it. Air is much less viscous than water. Water is much less viscous than molasses. If you have ever bought oil for your car engine, you may have seen the label \textit{10W40} or \textit{5W30} which indicate values of the viscosity both cold and hot.} Such a system is sketched in Fig. 4.

This sort of velocity-dependent force is familiar to us all. At 10 mph our hand sticking out of the car window does not feel too much resistance. At 60 or 70 mph the effect is
much more dramatic. The oscillator experiences an additional force, which in the simplest
approximation depends linearly on the velocity, $v$, which we write as $-bv$, where $b$ is called
the damping coefficient.\footnote{If the velocity is sufficient to cause turbulence, then a
damping force quadratic in velocity becomes important, i.e. $F_{damping} = -b_1v - b_2v^2$, and the
equation which describes the motion of the system becomes nonlinear. We will ignore this quadratic
term, which implies that the velocities under consideration are not too large and that the
coefficient $b_2$ is small.} The negative sign tells us that the damping force opposes the
motion.

The resulting motion of a system depends on how large the damping force is. Consider
what will happen when you pull the mass aside and let it go as we described above. You
can imagine that the damping force could be so large that shortly after you release the mass
the damping force just balances the spring force and the mass slowly moves back to its
equilibrium position. This situation is called overdamped.

On the other hand, the damping could be light enough to permit the mass to oscillate
a few cycles, or many cycles for that matter. Physicists call this underdamping or light
damping. However, each successive oscillation will have a reduced amplitude, since part of
the system’s energy will be lost to work done against the damping force.

A good example of this is an automobile. When a car hits a bump, it may bounce up
and down once or twice, but unless the shock absorbers are bad, it will very quickly stop
oscillating up and down.

We are all familiar with energy loss due to friction. While this viscous damping force
is not exactly like friction, some ideas do transfer over. For example, think about what
happens when you push a heavy box across the floor. We have to do work against friction
to get the box to slide across the floor. To get it started sliding we have to overcome the
(static) frictional force and the box’s inertia. If the box were on wheels, then once you got
it rolling, it would continue to roll. However, eventually it would stop because the wheels
also have some friction which will remove the kinetic energy which resulted from your doing
work on the box.

\begin{figure}
\centering
\includegraphics[width=0.2\textwidth]{damped_oscillator.png}
\caption{The Damped Oscillator}
\end{figure}
The motion is sketched in Fig. 5, where we see that the amplitude decreases with time. We can define the halving time, $T_{1/2}$ which is the time it takes for the amplitude to be reduced by one half. Note that this time ($T_{1/2}$), is independent of the time when you start to measure $T_{1/2}$. The important measure of the damping is the quantity $\gamma = b/m$ the ratio of the strength of the damping force to the mass.

\textbf{aside:} For the curious, we give the the equation for the position as a function of time:

$$y(t) = Ae^{-\frac{1}{2}\gamma t} \cos (2\pi f_d t).$$  \hfill (6)

There are two important features to this solution which arise from damping. The amplitude of the oscillations dies away (the heavier the damping the faster the oscillations die out), and the frequency of oscillation is lowered from $f_0$ to some new value $f_d$. The size of the frequency change is small and we will not worry about it here.

The decrease in amplitude should come as no surprise since the damping force does work on the system and thus takes energy away from it. Each time the mass comes to rest, (let’s call the displacement where it comes to rest $Y_{max}$) the total energy remaining in the the system is given by $(1/2)KY_{max}^2$, however, since energy is lost, $Y_{max}$ is less each time.

Where does the energy go? The ultimate answer is into heat. The damper moving through the liquid stirs the molecules and heats them up. Even if you just have a block of wood on a spring in air, there will be some loss of energy due to the viscosity of the air.

In a real spring, there is a second source of energy loss. If you have ever bent a piece of metal back and forth until it breaks, you have noticed that it heats up. A real spring supporting the mass will dissipate some energy.

\section*{2.1 The Damped Oscillator in Music}

A number of musical instruments are damped non-simple oscillators, which have several frequencies with which they can vibrate. The simplest example is a tuned bar or tuning fork which is designed to give a single pitch for tuning. More complicated examples are the string of a guitar or violin, the air in a trumpet and the head of a drum. There are two ways in
which energy is put into these systems. For percussion instruments, the piano and the guitar we strike the instrument (or string) with a mallet or hammer or pluck it with a plectrum. Then the system is allowed to vibrate freely. The amplitude will decrease with time due to damping forces, and eventually the oscillations become so low in amplitude that the sound is no longer audible.

For the bowed violin, the trumpet, or a number of other instruments, the player provides a constant source of energy and a steady sound is produced. We discuss this further in the section below on driven oscillators.

### 2.2 Radiation Damping

The tuned bar used for demonstrations in this course has a resonator box below it which makes the oscillations of the bar at 440 Hz sound louder. In addition to the dissipative forces, which remove energy from the bar after it is struck, some of the energy goes into the sound waves which are radiated into the air.

What happens to the air in the box below the bar also leads us to the topic of resonance and the driven oscillator.

### 3 The Driven Oscillator

When you were small, chances are you went to a playground and rode in a swing. At first you had to be pushed by someone bigger, but soon you learned that you could pump the swing to get it going, and if you kept pumping, the amplitude of your swinging increased to a maximum value. What you didn’t realize was that you were providing a periodic driving force, and the period was that of the swing. If you pumped too fast, or too slow, it didn't work very well.

This was your introduction to the phenomenon of resonance, which is just the response of a system which can oscillate to a driving force with a frequency equal (or close) to the natural frequency of the oscillator.

Although the equations are not that complicated, our point here is to describe this behavior in a qualitative way. We can separate the response of an oscillatory system to a driving force as the sum of two parts:

- **A steady state** part with the frequency of the driving force. The amplitude is completely determined by the strength of the damping force, how far the driving frequency is from the natural frequency, and how strong the driving force is.

- **A transient part** which goes at the frequency $f_d$ (the frequency that system would oscillate with if we just impulsively put energy into it and then let it freely oscillate) and it will decay away with the same halving time as if it had been struck impulsively.

We first turn our attention to the steady-state part of the motion. It takes place at the frequency of the driving force, and its amplitude depends on the damping and how close you are to the natural frequency. In Fig. 6, we show how the steady state amplitude depends
Fig. 6: The steady state amplitude as a function of frequency for an oscillator with several values of damping. As the damping gets larger the maximum amplitude gets smaller.

on frequency for several values of the damping. Note that as the damping is increased, the maximum possible amplitude is decreased. However, the range of frequencies over which the response is half of the maximum value or greater is increased! The amount of damping desirable is different depending on the specific situation. On a brasswind instrument, you want the resonances of the air column to be broad enough that the player can excite one of them and have some flexibility to move the frequency up or down a little, but you don’t want them so broad that there is no center to the note or adjacent notes overlap. On the other hand, if you are talking about a soundboard, you want to “broadcast” a range of frequencies so you do not want very narrow resonances, since these would strongly color the sound, which in the worst case would make the sound of adjacent notes very different.

We now ask: “What happens when we start driving an oscillator?” In musical terms, this is equivalent to asking what happens when a brass player begins a note.

There are two possibilities: The player can have the lips vibrating at exactly the same frequency as the air column wants to vibrate with, or the lips can be slightly off. The air column wants to vibrate with its own frequency, \( f_d \) as we called it. This part of the motion is called the transient part of the motion. Eventually the air column will only vibrate with the frequency that the lips are driving it with. This part of the motion is called the steady-state part of the motion.

The time just after the driving force is turned on can be very interesting, if the frequency of the lips is not that of the air column. We can get beating between the driving frequency and the air column’s natural frequency, which sounds like a little fuzz (or big splat if the player really misses the note).

If we wait long enough, then the transient part will die out and we will be left with only the steady state part. This is shown for several examples of damping and frequency in Fig. 7(b) and Fig. 7(a). If the damping is not too heavy, there can be beating between the driving frequency and the natural frequency, i.e. between the steady state response and the transient response. This is clearly visible in the figures.
Fig. 7: (a) The complete motion, along with the transient part and the steady-state part when the driving frequency is exactly on resonance. (b) The response of a driven oscillator, with the driving frequency different from the natural frequency. Notice that the transient part of the response dies out more quickly with heavier damping.

Fig. 8: The response of a driven oscillator with modest damping and the driving force off resonance. Note the regular change in the amplitude of the full motion during the time that the transient term dies out. This is called beating between the driving frequency and the natural frequency.
4 Nonlinear Oscillators

A linear oscillator can oscillate with only one frequency, its motion is sinusoidal and periodic. If the return force in the spring shown in Fig. 1 is not linear, the motion will still repeat itself, but it will no longer have only a single frequency in its motion. The oscillations will repeat over and over, always with the same period, but the position as a function of time will not be given by $y = A \cos(2\pi f_1 t)$, where $f_1 = 1/P$. The reason for giving $(1/P)$ the name $f_1$ will soon be apparent.

So what are the building blocks of this complex motion? What property do we need for these building blocks? We need building blocks that depend on time, since we are trying to describe the motion of the nonlinear oscillator. The simplest building block would be a sine wave, but with a different frequency, but the immediate question is “What frequency?” If we want to describe motion that exactly repeats itself cycle after cycle, we need building blocks that oscillate with exactly the same period $P$, or at least integer fractions of the period, namely $P$, $P/2$, $P/3$, etc., which we can write symbolically as $P_n = P/n$ where $n$ is an integer, $n = 1, 2, 3, \ldots$. Since the frequency $f_n$ is the inverse of the period,

$$f_n = \frac{n}{P} = \frac{n}{P} = nf_1. \quad (7)$$

Fig. 9: The first four harmonics. The fourth harmonic is shown below the others for clarity. The dashed vertical line shows half the period, so $t = P/2$. Note the difference between the behavior of the odd (1 and 3) harmonics and the even harmonics (2 and 4) just after $t = P/2$.

These building blocks, called harmonics, are simple sine waves with frequencies that are integer multiples of the lowest frequency $f_1 = 1/P$. They go through exactly 1, 2, etc. complete oscillations in the period $P$. The frequency $f_1$ is called the fundamental of the harmonic series. The first four harmonics are shown in Fig. 9, which are given by

$$y_n = A \cos \left\{n(2\pi f_1 t)\right\} \quad n = 1, 2, 3, 4. \quad (8)$$

Notice that at the midpoint, all of the harmonics are zero, but the even harmonics have gone through an integer number of cycles, and are going positive again while the odd harmonics
have gone through 1/2, 3/2, 5/2, etc cycles, and are going negative. This will become important when we talk about periodic waves, in the last section of this document.

If the period of the oscillation is $P$, then the frequencies present in the motion are

$$f_1 = \frac{1}{P}, \ f_2 = 2f_1 = \frac{2}{P}, \ f_3 = 3f_1, \ f_4 = 4f_14f, \ etc.$$  \hspace{1cm} (9)

To summarize, the motion contains the frequency $f_1$ which is the inverse of the period, plus harmonics (integer multiples) of this frequency. **This is very different from the simple oscillator.** In the simple oscillator we had one frequency which only depended on the stiffness and inertia of the system. Now, with a nonlinear return force, we get something quite new. The motion of the nonlinear oscillator consists of a complex motion made up of harmonics of $f_1$. The participation of each harmonic in a complex oscillation depends on the details of the nonlinearity.

There are two important characteristics of the nonlinear oscillator.

1. The effects of the nonlinearity become much more important as the amplitude is increased.

2. For some types of nonlinearity, the frequency of the oscillator will change with amplitude.

Thus when we drive a nonlinear system, the larger the amplitude the more important the higher harmonics are. The second property can make for very interesting response when the system is driven at different amplitudes producing a very curious shape to the resonance curve, but we will not discuss it here. We note that the pendulum is not a simple oscillator, but rather a nonlinear one, with the frequency decreasing with increased amplitude.

### 4.1 Harmonic Distortion (THD)

So, the next question we have to answer, is “What happens when we apply a sinusoidal (single frequency) driving force to our nonlinear oscillator?” Recall that when we drove the simple oscillator with a sinusoidal force, we got a response at the driving frequency. So we might expect that the nonlinear oscillator will respond at the driving frequency as well. However, we might suspect that because of the nonlinearity, the system will respond in a more complicated way. In fact, the resulting motion will have the same period as the driving force, but harmonics of the driving force will also be present in the motion. In Fig. 10 we represent this schematically, where we represent the nonlinear system as a box, and don’t worry what is inside. The important point is that we start out with a single frequency and end up with a set of frequencies that contains the driving frequency and its harmonics. These new frequencies which were not in the original driving force, are called **harmonic distortion.** Incidentally, these new frequencies which appear because of the nonlinearities are sometimes called by their engineering name, **heterodyne components.**

The set (collection) of frequencies which are related to each other by integer ratios, which we could define as

$$f_n = nf_1$$  \hspace{1cm} (10)
is called the harmonic series. These frequencies have musical significance. In Fig. 11, the harmonic series based on $C_2$ is shown. The whole notes are more or less in tune with the notes you would expect from the piano. The black notes are not “in tune” with what we expect when we see those notes written on the staff, and many musically trained listeners would find them difficult to listen to.

Suppose that we drive a nonlinear system with a frequency $f$. The output consists of the original frequency and its harmonics, i.e. $f$, $2f$, $3f$, $4f$ etc. The details of how much of each harmonic is present depends on the details of the nonlinearity.

To illustrate this point, we will consider what an amplifier does. We are all familiar with the volume control on our radio, i-pod, etc. that increases the loudness of the music. This is possible because the music is put through an amplifier which has an adjustment so we can increase or decrease the output, depending on how loud we want the music to be. An ideal amplifier would produce an exact copy of the input but with a greater, or smaller amplitude. Thus if we listened to a sine wave, we would expect to get an exact copy at the same frequency. This would be a linear amplifier. Now suppose we had an amplifier that was non linear. An example is given in Fig. 12, which has a linear response for the negative going part of the sine wave, but a quadratic response for the positive going part. The effect is to stretch the positive part of the input sine wave, so it’s no longer a simple sine wave. It’s still has the same period, but it’s no longer a simple sine wave. We understand from our discussion above, that we started with a frequency $f$, and ended up with $f$ plus its harmonics, $f$, $2f$, $3f$ · · ·.

Harmonic distortion is a key problem to be controlled in any high-fidelity sound reproduction system. While the total harmonic distortion (called THD and defined in the appendix) is quite small in almost all commercially available amplifiers, the mechanical transducers
4.2 Intermodulation Distortion (IMD)

If two sinusoidal driving forces with frequencies $f_a$, $f_b$ act on a nonlinear system we get an even more surprising result. In addition to the harmonics which you would expect from the discussion above, you also get sums and differences of the two driving frequencies, i.e. $f_a + f_b$, $f_a - f_b$, $2f_a + f_b$ etc. You might say that the whole is greater than the sum of its parts. These additional frequencies are called intermodulation or IM distortion.

We can classify these components according to their order, which is done in Table 1. The simplest ones which are not present in the original frequencies are those listed in the second order row. It is these simplest sum and difference tones which are musically significant, and useful. The ear is a nonlinear system, and can generate sum and difference tones when presented with two pitches. Since the frequencies in the harmonic series are related
Tab. 1: The Heterodyne Components Generated by Two Frequencies. We assume that $f_a < f_b$. The first, second (quadratic), third (cubic) and fourth (quartic) order frequencies are given.

<table>
<thead>
<tr>
<th>Order 1:</th>
<th>(the original frequencies) $f_a$, $f_b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Order 2:</td>
<td>$2f_a$, $2f_b$, $f_a + f_b$, $f_b - f_a$</td>
</tr>
<tr>
<td>Order 3:</td>
<td>$3f_a$, $3f_b$, $2f_a + f_b$, $f_a + 2f_b$, $2f_b - f_a$, $2f_a - f_b$</td>
</tr>
<tr>
<td>Order 4:</td>
<td>$4f_a$, $4f_b$, $3f_a + f_b$, $f_a + 3f_b$, $2f_a + 2f_b$, $2f_b - 2f_a$, $3f_a - f_b$, $3f_b - f_a$</td>
</tr>
</tbody>
</table>

by integers, with the frequency of the $n^{th}$ harmonic being $nf_a$, we can find the sum and difference tones by adding and subtracting the harmonic numbers. For example, consider the musical interval $C_5 - E_5$, which is a major third (M3), and is the interval between the $8^{th}$ and $10^{th}$ harmonics. The difference tone is the $(10 - 8)^{th}$ or $2^{nd}$ harmonic $C_3$, which is two octaves below $C_5$.

This can be used to musical advantage. For example, Robert Osmun (of Osmun Brass Instruments) told the author that at a recital of music written for the baroque trumpet, Edward Tarr tuned to the organ which accompanied him by playing a M3, and changing his trumpet until the difference tone two octaves lower was in tune with the root played by the organ. This produced a beat-free just tempered M3.

Richard Merewether, who was the principal designer of the Paxman horns made in London, once complained bitterly to the author that some conductors preferred equally tempered M3s rather than the just ones where the difference tone is in tune. He told the author of a performance of Judas Maccabaeus in London where he (RM) and the second horn player were almost fired for playing a just M3 rather than playing the equal tempered M3 which the pianist and organist conductor had demanded.

If you have ever bought a high fidelity amplifier, you may have noticed that the specifications include a statement of THD (total harmonic distortion) and IMD (intermodulation distortion). Every amplifier exhibits distortion, it is a quantitative question of how much distortion is present, compared to how sensitive you are to it. To put it differently, designers try to minimize the distortion, and if you are willing to pay more, you will probably get lower distortion specifications. In the end, you have to listen to the system and see if it matters.

The perfect amplifier would produce a perfect copy of its input signal, with absolutely no change except a scale factor, i.e. everything is just larger by the same linear factor. If we put a sine wave of a single frequency in, we get one out. Since no amplifier is perfect, instead we get out a waveform which repeats itself, but it is no longer simple. Said differently, it contains the original frequency and its harmonics.

### 4.3 Nonlinearities in Musical Acoustics

Two examples of the musical relevance of difference tones were given above, to emphasize that the seemingly abstract idea of difference tones appears naturally in musical settings.

The topic of harmonic generation by nonlinear systems is at the core of musical acoustics. The production of harmonics when a nonlinear system is driven by a single frequency is
central to speech and to the production of musically useful sounds from all instruments which produce a steady sound. In every instance where a steady sound is produced there is a nonlinearity which produces a harmonic spectrum of a single driving frequency. For example, even though a speaker’s vocal folds, or the lips of a brass player, or the reed of a clarinet may open and close sinusoidally at small amplitudes, the volume of air which flows through an opening does not vary linearly with the dimension of the opening. Thus the pressure variations downstream are periodic but not sinusoidal so they contain harmonics of the fundamental frequency. They are periodic, and have a fundamental frequency which is the inverse of the period. We will learn later that the Fourier theorem will tell us that these puffs of air, (pressure waves) contain a number of harmonics of the fundamental frequency.

In the violin family it is the nonlinear dependence of the slip stick force between the bow and string which causes the harmonic spectrum. Thus the motion of the string is not simple harmonic at one frequency, but represents a combination of modes of the string.

5 Periodic Waves

The motion of a nonlinear oscillator is periodic, with a period $P$, but it is not a simple sine function with one frequency. Instead we have argued that it is built up of harmonics $f_1 = 1/P$. We can generalize this idea to waves that repeat themselves, which are called periodic waves. For sound waves, it is the air pressure that varies periodically with time. For light, it is an electric and magnetic field. For water waves it is a disturbance that travels through the water. Just as for the motion of a nonlinear oscillator, we want to know which harmonics are present, as well as their intensity.

The mathematical process that tells us how to calculate the frequencies in a periodic wave is called Fourier analysis, in honor of Jean Baptiste Joseph Fourier who invented the procedure. Our statement above is commonly called Fourier’s Theorem: Any periodic waveform of period $P$ may be built up out of a set of sine waves whose frequencies form a harmonic series with $f_1 = 1/P$. Each sine wave must have the right amplitude and phase (which can be determined using calculus).

Two examples of periodic waves are shown in Fig. 13. The sum of harmonics 1-3 is shown, with their correct amplitudes. Notice the difference between the two waveforms. The sawtooth rises slowly in the first half of the period, then drops abruptly to zero. In the second half, it first drops immediately to the maximum negative amplitude, then rises slowly. Notice that the square wave goes negative in the second half of the period exactly as it went positive in the first half period. Recall that in Fig. 9 we observed that the odd harmonics are going negative after half a cycle, and that the even harmonics are going positive. We conclude that the even harmonics cannot contribute to a wave that goes negative in the second half of the cycle exactly like it went positive. Thus such a wave contains only odd harmonics. On the other hand, the sawtooth contains all harmonics. In both cases the amplitude of the $n$th harmonic has an amplitude of $A_1/n$ where $A_1$ is the amplitude of the first harmonic. The difference is that the even harmonics are missing in the square wave, which means that the sawtooth wave contains more energy in the high frequencies.
Fig. 13: (a) A “sawtooth” wave, and the sum of 3 harmonics. (b) A square wave and the sum of the harmonics 1 and 3.

In Fig. 14(a) a “triangle” wave is shown along with the first harmonic, and in Fig. 14(b) the sawtooth is shown along with the sum of many harmonics. Because of the sharp edge on the sawtooth (and also for the square wave) the rapid wiggling you see there never goes away, no matter how many harmonics you add in. (aside: this overshoot is called the Gibbs phenomenon.)

Fig. 14: (a) A triangle wave, which like the square wave only contains odd harmonics. Only the first harmonic is shown. (b) A sawtooth wave showing the sum of many harmonics.

While we will not introduce the mathematics necessary for Fourier analysis, there are some simple rules that we can write down which give a qualitative description of the frequencies present in a waveform.

- If the waveform is smooth and does not have any parts which vary rapidly, then only the first few harmonics are important.
- If there are sharp kinks, or rapid wiggles, there are many high harmonics.
• Any wave whose second half merely repeats the first half exactly except negative (upside down) has only odd harmonics.

• If the waveform does not exactly repeat itself negatively on the second half of the period, then all harmonics are present.

• For a pulse train, as shown in Fig. 15 that is positive for a fraction of the period (called the “duty factor”), then harmonics that are integer multiples the inverse of the duty factor are missing.

![Fig. 15: A pulse train. The wave is positive for a time $\Delta t$ and then zero for the rest of the period $P$. In this picture the duty factor $\Delta t/P$ is 1/4 so every 4th harmonic will be absent.](image)

The square wave is an example of the last point, since the duty factor is 1/2, and harmonics 2, 4, 6, (all even ones) are missing. In the figure, the duty factor is 1/4, so harmonics 4, 8, 12, 16 · · · are missing.

Another example of Fourier’s theorem is clipping of a sine wave by a diode, or by an amplifier. A diode is a one-way flow valve for current in an electrical circuit. If a sine wave is sent through a diode, it only lets the positive (or negative) part of the sine wave through (depending on which way the diode is inserted in the circuit). This is shown schematically in Fig. 16. (aside: It is possible to add a constant offset to the sine wave, so that the negative-going part is only partially clipped. In either case, such a signal will contain all harmonics.

![Fig. 16: A sine wave before and after it goes through a diode.](image)

A related topic is clipping by an amplifier. If the input to an amplifier is too large, or the volume control is turned up too high, the amplifier will be asked to put out a signal
that is larger than it is capable of doing. The result is to output a signal that is chopped off at the maximum amplitude that the amplifier is capable of. This is shown schematically in Fig. 17. As it is shown, this waveform will only contain odd harmonics. (Why?) The details of exactly what happens when the amplifier clips depends on the amplifier circuit, especially on whether it is a tube or solid-state amplifier.

![Fig. 17: The output of an amplifier after clipping.](image)

The possibility of clipping is enhanced if the amplifier is not matched to the speakers in a sound system. Suppose the speakers require significantly more power to play loud than the amplifier can produce. The listener will turn up the volume, and the amplifier will begin clipping, introducing high frequencies into the signal driving the speakers. Depending on the details, the energy in these extra high frequencies could destroy, or damage the tweeters. Thus care should be given when purchasing a sound system that the speakers are well matched to the amplifier power.

What if a function is not periodic? Then it can contain all frequencies. If they are random, then the resulting pattern is called noise. There are two types of noise that are quite useful in acoustics. White noise contains equal energy per frequency interval. This means that there is the same energy between 10 to 20 Hz as there is from 440 to 460 Hz. Pink noise contains equal energy per octave. That means that there is the same amount of energy between 40 and 80 Hz, as there is between 80 and 160 Hz. You can approximate white noise by holding your tongue close to the roof of your mouth and just behind your front teeth and forcing an airstream through. The noise your FM receiver makes when tuned between stations is also close to white noise. If you relax your tongue as if you were saying “sh” as in should, the sound will be close to pink noise. Thus white noise contains much more energy in the high frequencies. Pink noise is sometimes used to balance a sound system in a performance or listening space, because the energy distribution in pink noise is a close approximation to the average energy in musical performance.

6 References

There are four classic references on musical acoustics. Each has its strong points and its weaknesses. The books by Hall, Benade, Backus and Rossing are briefly annotated below.

1. Donald E. Hall, *Musical Acoustics*, Brooks/Cole Publishing Company, Second Edition 1991, contains a discussion of many of these issues. However the discussion of simple and damped oscillators and resonance is divided between chapters 2 (§2.4–2.5), chapter

---

3 A loudspeaker has several different components, one for the high frequencies called a tweeter, and one for low frequencies called a woofer. Expensive speakers might have a separate mid-range component as well.
9 (§9.7) and chapter 11 (§11.3). Nonlinear behavior is discussed in chapter 17 (§17.5), but not in a comprehensive way.

2. Arthur H. Benade, *Fundamentals of Musical Acoustics*, Oxford University Press, 1976, which has been reprinted by Dover Press in its musical texts series (not the physics series). This book contains a much more systematic and thorough discussion of these topics beginning with chapter 4, and going through chapter 10. Along the way he discusses systems which can vibrate with more than one frequency in an integrated way with simple oscillators. Damping is developed in a natural way by discussing impulsively excited systems which decay with some characteristic time. In Chapter 14 an extensive discussion of nonlinear systems is given.


4. Thomas D. Rossing, *The Science of Sound*, Addison-Wesley Publishing Company, Second Edition 1990. This book is slightly more technical than the above references, but should be accessible if you read it carefully. Chapters 1-4 form an introduction to all of the physical principles used later, chapters 2 and 4 cover oscillators and resonance.

7 Appendix: Definitions of THD and IMD

The production of harmonics of the driving frequency is commonly called harmonic distortion. That is because these new frequencies (the harmonics) were not in the original driving force, and thus their introduction represents a distortion of the original signal. Often devices are characterized by their total harmonic distortion, which is abbreviated as THD and is defined as

\[ THD = \frac{\sqrt{A_2^2 + A_3^2 + A_4^2 + \ldots}}{A_1} \times 100\% \]  

where \( A_j \) is the amplitude of the \( j \)th harmonic.

Intermodulation Distortion (IMD) is defined order by order. The second order IMD is characterized by

\[ IMD_2 = \frac{A_{f_a + f_b} + A_{f_b - f_a}}{A_{f_b}} \times 100\% \]  

and similar definitions exist for higher orders.

8 Exercises

1. You are surrounded by physical systems which can oscillate. All that is needed is a mechanical system which has an equilibrium position, and a restoring force if it is
moved away from the equilibrium position. Many of them are driven and you observe 
resonance. Spend a day looking for examples, and list four of them. Please do not 
list all examples from acoustics. Look beyond your music experience. Tell how each 
is an example of resonance. You should clearly identify what it is that oscillates, 
what provides the driving force, and how the frequency dependence comes into to play. 
Remember if the response is not dependent on frequency, it is not resonance.

2. The note $A_4$ is assigned a frequency of 440 Hz by agreement. What is the frequency 
of $A_3$? What is the frequency of $A_4$? Suppose you play $A_4$ on an instrument which 
produces a sinusoidal waveform (no harmonics). What frequencies will be present if 
this acts on a nonlinear device. (give at least three) Now suppose that an $A_4$ and an $E_5$ 
(which you can take to have a frequency of 660 Hz) act on a nonlinear device? Using the 
back cover of your book, figure out which notes the order 2 (Table 4.1) components are. 
(Hint: Because you have rounded off the frequency of $E_5$, you will not get exactly the 
frequencies listed on the back cover.) Suppose you heard this collection of frequencies. 
What chord would you hear?

3. If you are regularly in musical performance situations, listen for difference tones. What 
are the circumstances where you hear them the most? If you do not have the opportu-
nity to listen for these tones in musical settings, then discuss where you might expect 
them.