Metrology

Metrology, the science of measurement, is a most important activity in the recording of sound: what we record is in fact a measurement of the sound pressure or electronic signal amplitude we wish to preserve. In order to manipulate and record audio signals, we measure characteristics of the signal, generally amplitude as a function of time, and in some manner preserve that measurement. Analog magnetic tape stores a continuous magnetic pattern representation of the signal. Digital audio systems measure the signal amplitude, producing a list of discrete measurements at regular times that can then be treated like any other computer data. Since sound measurements are a function of time, we must have a way of measuring time as well. Once we have a reliable measure of time and the signal, we can manipulate the signal electronically or numerically and make changes to suit our artistic desires: we can process and mix sounds into sonic landscapes that never before existed. Appreciating that the quality of a recording depends on how well we are able to make a measurement is an important first step in the study of sound recording.

In order to study the recording process in an organized way, we should first familiarize ourselves with the tools involved in the investigation. We are interested in the processes that allow the permanent storage of information contained in the transient air pressure variations we call sound. Because sound recording involves acoustics, mechanics, electronics, magnetism, and even physiology and psychology, we will need to employ the tools of science. Since we are concerned with measurements of the behavior of air pressures, electronic signals, and magnetism, physics is the branch of science that develops our understanding of these systems and it uses mathematics as the descriptive language. It turns out that similar mathematical descriptions apply to electronics, acoustics, and mechanics, so understanding an occasional equation will help to explain and understand the fundamentals of sound recording.

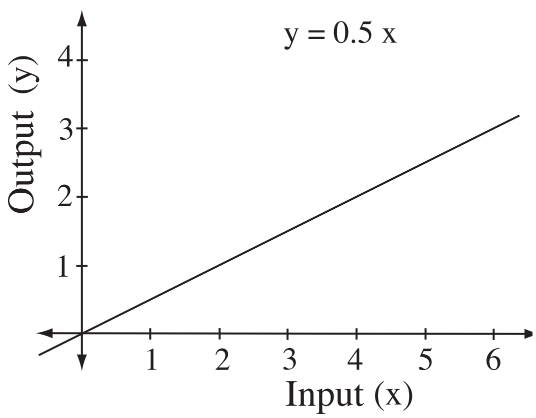

Figure 1 Linear

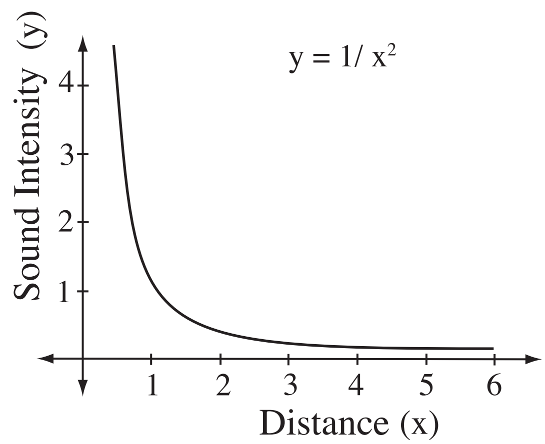

Figure 2 Non-linear

When we seek to establish how two phenomena are related, we do so by plotting a graph comparing two or more sets of measurements termed variables. The variable we control is called the independent variable while the one we measure is called the dependent variable because its value depends on the value of the independent variable we choose. For example, we might wonder how the amplitude of a sound system varies with frequency. We measure the amplitude at a number of frequencies and plot amplitude against frequency. The shape of the resulting curve shows us the function that describes the relationship between the two properties, in this case the frequency response of the system. Often such measurements fall along a straight line, in which case we call the relationship linear. Linearity is a requirement of most audio systems: for instance, we want the output of an audio amplifier to be a linearly amplified version of the input. In this case, the gain of the amplifier is the slope of the line plotting output voltage (y-axis) against input voltage (x-axis), where the slope is the output level divided by the input level (Fig. 1). For a linear system, this value is constant: in Fig. 1, the output is always half (x 0.5) the input level regardless of the input voltage. Other relationships, like the intensity of a sound as a function of distance from the source for example, are more complicated and deviate from linearity (Fig. 2). Nonlinear relationships are sometimes encountered that can lead to distortion in audio systems.

Distortion is simply a change in the signal. It may even be used intentionally, as in the case of the audio compressor and guitar amplifier. Some forms of distortion are more noticeable than others. Understanding how linear and nonlinear systems are described mathematically will help clarify the difference.

As long as the variables under consideration are related by an equation not containing variables multiplied together or raised to powers, it is a linear relationship. Linear processes include mixing, in which signals are added together, and amplification, in which a signal is multiplied by a constant, the gain factor. Nonlinear relationships include modulations, in which one signal is multiplied by another (as in AM and FM radio), and exponential or logarithmic relationships. Nonlinear systems create distortion of the original signal by introducing harmonics (multiples) of signal frequencies that were not present in the original signal source.

Audio signals, whether measured as sound pressure levels or electronic voltages, are often represented as sine waves. Music, in fact all sounds, may be described using combinations of sine waves with varying amplitudes, frequencies, and phase relationships. Although real musical signals are quite complex, we often employ simple sinusoidal test signals in measuring audio systems. Sine waves are actually mathematic functions:

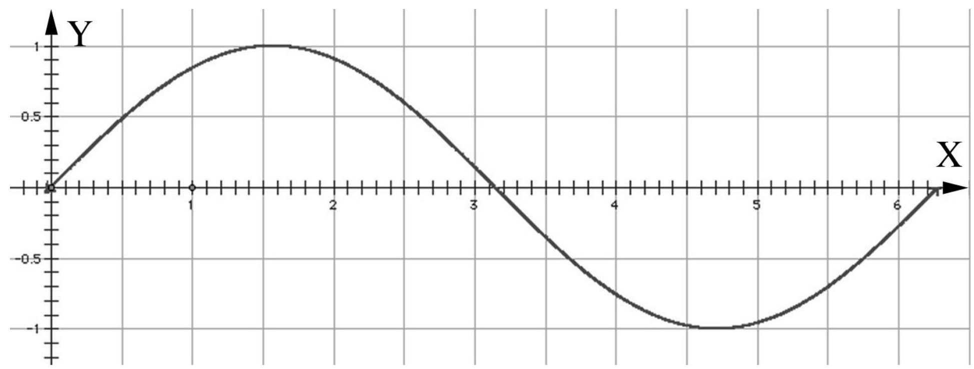

Figure 3 $y = A \sin(x)$,

where the amplitude y equals some maximum value A (A=1 in Fig. 3) multiplied by the sine of the variable x, with x ranging from 0 to 2π to generate one complete cycle. The sine function can be generated by the rotation of a point on a circle with a radius of 1, called the unit circle. If the point starts at the 3 o'clock position and rotates counterclockwise along the circle circumference, it generates a sine wave. (When the point starts at the maximum y=1 (12 o'clock) instead of from y=0, it generates the cosine function.) At any angle of rotation, the height of the sine wave along the y-axis will be the sine of the angle and the x-position will be the angle of rotation in radians. While we are familiar with degrees as the measure of angles where 360° describes a full circle, mathematics prefers the radian as the angular measure. There are 2π radians in a full circle and therefore in a complete sine wave cycle.

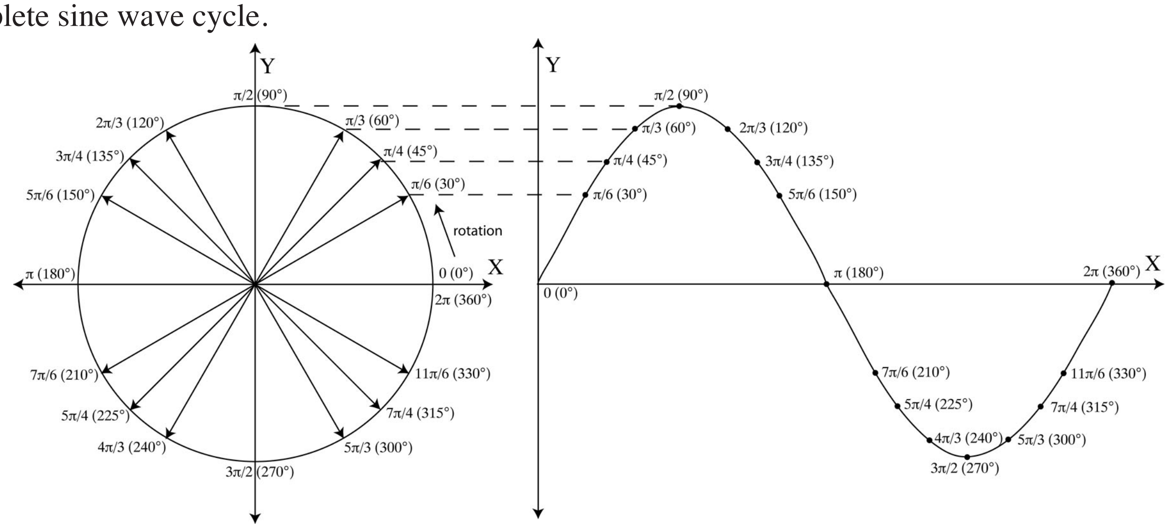

Figure 4 The unit circle and the sine wave

The phase of a sine wave is a measure of the displacement from the origin, along the x-axis, of its beginning, the positive-going 0 crossing. When beginning at the origin, the phase is taken to be zero. The amount of rotation of the unit circle required to turn one sine into the other is the relative phase angle. As we add sine waves together to make a complex signal, the relative starting points of the various sine components determine the phase relationship between them. If two sine waves are exactly in phase, they add constructively. If they are 180º (π radians) out of phase, they cancel completely provided they are of the same amplitude. We will see examples of such cancelations when we consider room acoustics and microphone polar patterns. Some electronic circuits alter the relative phases of the different frequency sine components of a signal in rather complex ways, notably filters.

Since sine waves are simple mathematical functions, they are easily described. We can measure the peak-to-peak swing of the signal, the absolute value of the signal, or even a time-average of the signal amplitude. Using sine waves as our test signals has its pitfalls though: real complex audio signals do not adhere to the same simple relationships that exists between these different measures when applied to simple sine waves. It will become apparent that how we choose to measure a signal makes a difference.

To keep track of our work, we are interested in displaying information about our audio signals graphically. By far the most common audio measurement is the signal amplitude. We need to create a meaningful display of the measurement that is easily interpreted and relatively inexpensive to duplicate, especially if we're dealing with dozens of separate signals, as is often the case with multitrack recordings. Early audio recording equipment solved this problem by using mechanical meters with a moving needle. More recently, we have seen the popularity of LED ladder displays. Each of these graphical indicators must conform to a set of rules about how we measure and display signal levels in order to be unambiguous. Since there are alternative methods of measuring the amplitude of signals, we should know which of these is employed in the graphical display we are using.

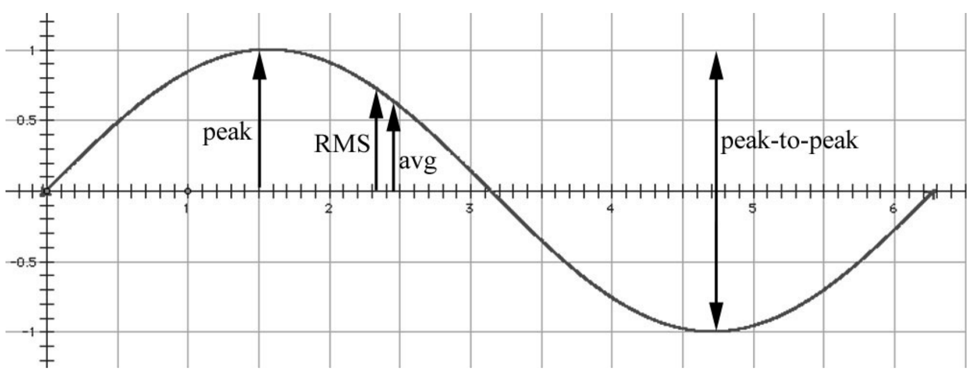

Figure 5 Sine amplitude measurements

The amplitude of a sine wave can be measured in different ways. The peak-to-peak measurement will tell us the swing between maximum and minimum values of our signal, but this may not directly correspond to the apparent loudness of that signal. The method that most closely approximates how we experience the loudness of a signal is the root-mean-squared (RMS) measure, in which we analyze the signal level over some time window and compute the square root of the mean of the squared value of the measured level. This is a rather complicated measurement to make, so a simpler method is often used: the average value. This value can be computed with simple analog circuitry, but it is a slightly less accurate measure of the perceived loudness of the signal. One problem with both of these approaches is that they average the signal, a process that will fail to track the maximum value of the signal. Since the maximum value determines how much signal the system must be able to accommodate, we are in danger of overloading the electronics if we simply measure the average level and ignore the peaks. Therefore very fast displays have evolved to augment the loudness-oriented displays. These peak-oriented displays may be as simple as an LED that flashes as the signal level approaches the limits of the electronic system.

Multiply below by x to convert to --->	Peak-to-Peak	Peak	Average	RMS
Peak-to-Peak	1	0.5	0.3185	0.3535
Peak	2	1	0.637	0.707
Average	0.785	1.57	1	1.11
RMS	2.828	1.414	0.9	1

Table 1 Sine amplitude measurement conversion factors

For sine waves, there is a simple conversion factor between these different measurements (Table 1). We notice that for a sine wave, the average level is 63.7% of the peak value while RMS is 70.7%. Unfortunately, for actual complex audio signals this simple relationship does not hold, so meters using different measures will not agree. In the studio, we often have to use equipment calibrated to different standards of measurement and we need to understand how they relate if we are to get the best performance from the system.

The range of values we encounter when measuring amplitude and frequency is enormous. We can hear frequencies from 20 Hz to 20 kHz (at least some people can) and hear sound levels from near silence to painfully loud, about six orders of magnitude in sound pressure amplitude. Working with numbers over this range is problematic, so mathematical shorthand is in order: the logarithm. Logarithms are exponents: the power to which a base number must be raised to equal the number in question. For example, the $\log_{10}$ (log base 10) of 100 (10^2) is 2. In the audio world, we generally use base 10 (unless otherwise designated, "log" implies base 10) however many physical processes are described by the so-called natural logarithm ln (base e = 2.718… don't ask) and digital signal processing is conducted ultimately in binary, or base 2. By using logarithmic measures, we avoid lots of 0's and keep the number of digits small.

$$\text{base}\rightarrow 10^{x \leftarrow \text{exponent}}$$

x	10^x
-2	0.01
-1	0.1
-0.5	0.316
0	1
0.303	2
0.5	3.16
0.603	4
0.903	8
1	10
2	100
3	1000

Table 2 Exponents in powers of 10

There are properties of logs that simplify their use. Any base to the zero power is one. Negative exponents yield numbers less that one, with a negative exponent giving the reciprocal of the positive value of the same exponent. Positive whole number exponents $base_{10}$ correspond to the number of zeros following the one. Logarithms only exist for numbers greater than zero.

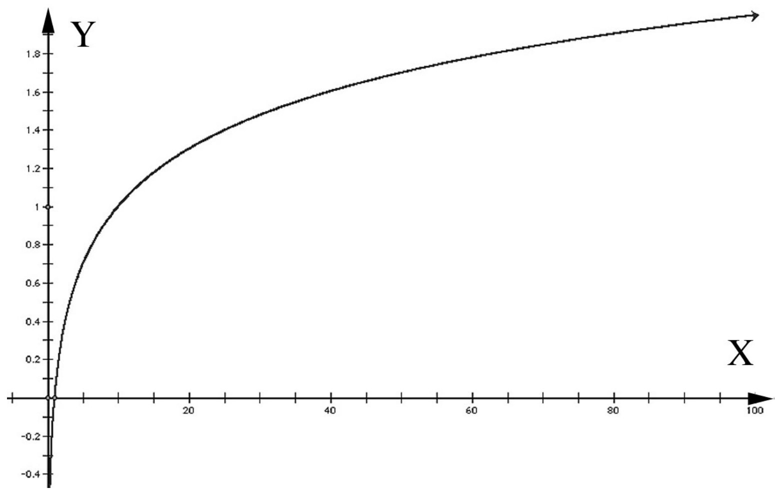

Figure 6 $y = \log(x)$

The primary audio application of the logarithm is the decibel. Since signal amplitude measurements must be made over many orders of magnitude, the bel $(\log_{10}(\text{power}_2/\text{power}_1))$ was adopted as a unit of measure by the early telecommunications industry. While appropriate in telegraphy where signal power transmission was important, the unit was too large to be convenient for electronic audio circuits and the decibel (1/10th bel) is now used. An important attribute of the decibel is that it isn't an absolute measure, but rather a ratio. It is used to describe how much larger or smaller a sound level or signal amplitude is than a standard reference level. Reference levels are chosen according to the application: they may represent the quietest sound we can perceive (dB SPL), the maximum signal a system can produce (dB FS), the recommended input level a device is designed to see (dBv), or whatever we choose (VU). Each of these assumes a different reference quantity. We will later discuss the various reference levels and why some decibels are $20\log_{10}(x)$ while others are $10\log_{10}(x)$.

$$\text{decibel} = 10\log_{10}\frac{\text{power}_2}{\text{power}_1} = 20\log_{10}\frac{\text{voltage}_2}{\text{voltage}_1}$$

Measurement Quantity	**Reference Level**
Sound Pressure	0 dB SPL = 0.0002 dyne/cm^2
Voltage	0 dBV = 1 volt
Voltage	0 dBu (or dBv) = 0.775 volt
Power	0 dBm = 0.001 watt

Table 3 Amplitude conversion factors

In dealing with physical systems, we are often describing quantities like forces. A force has two components: a magnitude and a direction or phase. In order to mathematically describe a force, we employ vector math. An example of a vector we commonly encounter without trepidation is wind: it has a velocity and a direction. Similarly, a force has a magnitude and a direction. (Measures of magnitude without a direction are called scalars: temperature, for example.) We can draw a vector as an arrow of length proportional to magnitude and pointing in the direction of the action. If we wish to add or combine two forces, we must use vector math in order to combine both the magnitude and direction information.

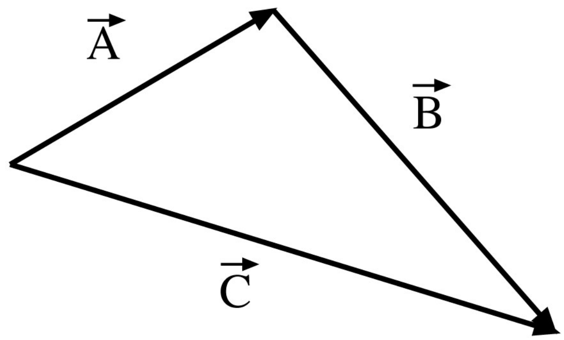

Figure 7 Vector addition

The above diagram shows how vectors add. Vectors $\vec{B}$ and $\vec{B}$ combine to produce vector $\vec{C}$: in order to get the magnitude and phase or direction correct we must use vector addition. The magnitude and direction of the resulting vector are found by separately adding the x and y components of the individual vectors. While using vector math simplifies computations for mechanical, acoustic, and electronic systems, we will only need to understand that vector math is different from scalar math for the purposes of our investigation.

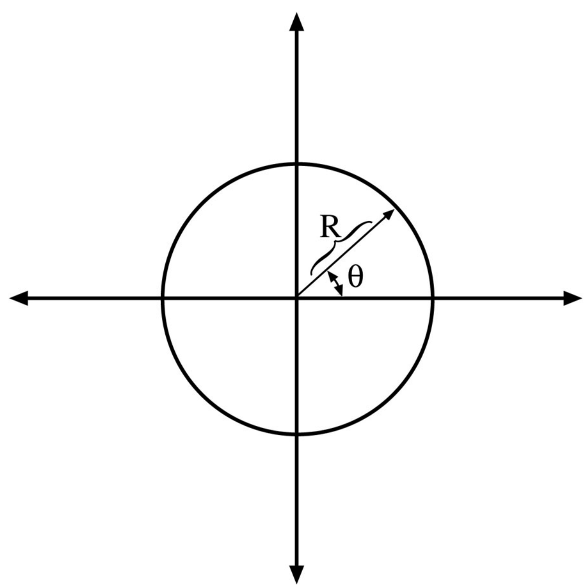

Figure 8 Polar coordinates

Another technique often used in describing magnitude and direction is polar coordinates. Instead of plotting vectors by giving pairs of Cartesian coordinates (x,y), polar coordinates consist of a magnitude and an angle (r,θ). This technique makes tasks like describing microphone directional sensitivity patterns and sound radiation patterns simpler, since the length of the line r indicates the microphone's sensitivity (output voltage for a given sound level input) and the angle θ indicates the direction from which the sound originates.

$$\mathbf{z} = x + yi$$

Vectors can also be represented as complex numbers: numbers with real and imaginary components. While this is not our usual way of thinking about numbers, the use of imaginary numbers makes analysis of acoustic and electronic systems simpler. We are familiar with real numbers, but imaginary numbers may not be so obvious. The square root of –1 is an imaginary number, since there is no real number which, when squared, generates a negative number. So we make some up: *i* and *j*. These are simply defined as the square root of –1 even though we don't come across such numbers in our daily activities (unless we're physicists or engineers.) If we make the real part of the number represent the magnitude of the vector and assign the imaginary part the role of describing direction or phase, we can use a single number to represent both components of the vector. By applying algebraic manipulation to the complex numbers, vector computations become more convenient.

We will not need to use vector mathematics much in our quest for a scientific understanding of audio recording, but we need not be intimidated by its presence when encountered in technical papers. It is merely a descriptive language to simplify the complex task of accurately and precisely analyzing complicated behavior of electronic and acoustic systems.

Units of Measure

In order to apply mathematical analysis to our meaurements, we need a standardized set of units of measure. Unfortunately, the need to standardize our system of measurement has led to several competing systems in different parts of the world. Most of the world uses the metric system while the US continues to use the older British system. The scientific community has adopted the metric system and that will be used here, although we will sometimes include British measurements as well since they are more commonly understood in the US.

The metric system uses the meter as a unit of length, the gram as the unit of mass, and the second as the unit of time. For large measurements, the MKS (meter-kilogram-second) system is convenient, while for smaller measurements the CGS (centimeter-gram-second) system is preferred. Scientists have adopted a system known as the International System of Units (SI) as the preferred system of measure. Some frequently used SI units of measure are shown in Table 4.

Quantity	SI Unit of Measure	CGS Unit of Measure
distance	meter	centimeter
time	second	second
mass	kilogram	gram
force	newton (N)	dyne (= 10^{-5} N)
pressure	pascal (P)	barye (= 0.1 P)
work, energy	joule (J)	erg (= 10^{-7} J)
electric current	ampere (A)	ampere
magnetic flux	weber (Wb)	maxwell (= 10^{-8} Wb)
magnetic field strength	ampere-turns/meter ($A \cdot m^{-1}$)	oersted (= 79.6 $A \cdot m^{-1}$)

Table 4 SI and cgs units of measure

Physics

When we think of sound recording, we usually think of microphones, wires, mixing boards and tape recorders. Or perhaps we think of computers and plug-ins now, but the basic foundation of the process of recording sound lies in the physics that describes the behavior of molecules and electromagnetic fields. The most critical elements of the recording process, the transduction of air pressure variations into voltage changes and the storage of the converted information as magnetic or optical records, involve the science of physics. Physics is a branch of science that seeks to describe and quantify the behavior of matter and energy in the precise language of mathematics. Since the interaction of air molecules with a microphone transducer element determines how accurately we are able to record the different sound frequencies and how the microphone responds to sounds from different directions, we benefit from an understanding of the physics involved. While the final decision about how well a microphone performs rests with our ears, understanding the physics involved helps us decide which microphone type and placement is likely to produce the desired result without listening to every microphone in every position possible.

We begin our discussion of physics by considering the basic principles of Newtonian mechanics. Much of our modern understanding of the physical world begins with the 17th century work of Sir Isaac Newton. While everyone who has thrown a baseball innately knows how it will behave, the mathematical equations that predict the path it takes are less well known. Newton's three laws describe the basic behavior of matter and energy in simple mathematical terms.

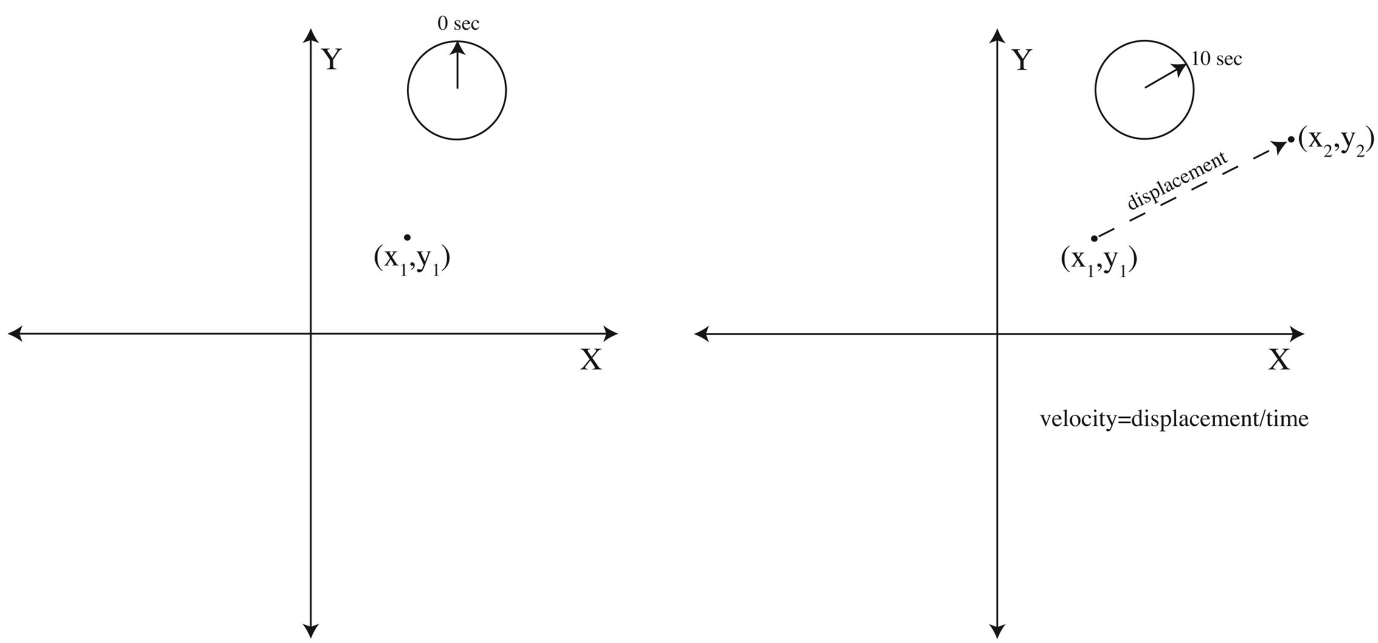

Figure 9 Displacement

To understand the physics of mechanical systems, we need to consider the basic properties of objects. Mass is the property we associate with the weight of an object, however strictly speaking that is only due to our earthly frame of reference and the pull of its gravity. Nonetheless, our everyday experience gives an intuitive understanding of what mass is. Displacement is the distance and direction an object moves. Velocity is the rate and direction of movement, its magnitude is the first derivative of position with respect to time (dx/dt). Since displacement and velocity have both magnitudes and directions, they are vector quantities. (The magnitude of velocity is the scalar we call speed.) Changing velocity is acceleration, the first derivative of velocity (dv/dt) with respect to time and the second derivative of position (d^2x/dt^2) with respect to time. Force describes a push or pull on an object, like the pull of gravity or the electric force of voltage on electrons. The more mass an object possesses, the more work a force must do to move it. That work is measured in terms of power or energy transferred. While we intuitively know these terms, formalizing their definitions is not so obvious.

Newton's First Law states that if no force is exerted on an object, its velocity cannot change. This is often simplified as "a body at rest tends to remain at rest." While we understand intuitively what a force is, it is less clear how to formally describe it. Loosely, force is a push or pull on an object. We deal with forces when we consider how air movement propagates and pushes and pulls on the diaphragm of a microphone. Forces are also involved in electric circuits and magnetic attraction and repulsion.

Newton's Second Law states that the net force on an object is equal to the product of the mass and its acceleration, or in equation form:

$$\vec{F} = m\vec{a}$$

where $\vec{F}$ is the force vector, m is the mass and $\vec{a}$ is the acceleration or change in velocity. Note that force and acceleration both have a direction of action as well as a magnitude and are therefore represented by vectors. This simple relationship describes a surprising percentage of physical interactions.

Newton's Third Law states that when two objects interact, the forces exerted by the objects on each other are equal in magnitude and opposite in direction or, "for every action there is an equal and opposite reaction.") This can be illustrated by a drunk leaning on a light pole: each object exerts an equal force on the other but in opposite directions, hence neither accelerates. It also applies to billiard balls bouncing off one another, much like molecules behave on a smaller scale. Using only Newton's laws, we are able to describe most of the physics involved in recording sound.

Since we use electronic and magnetic representations of sound, we also need to understand electromagnetism. Electronic circuits are analogous to mechanical systems. Voltage is the equivalent of mechanical force, pushing charges through a circuit. An inductor acts as a mass, storing energy in a magnetic field, while capacitors behave like springs, storing energy in an electric field. Resistors act like friction, simply dissipating energy as heat and opposing the flow of charge. Magnetism exerts a force on susceptible particles much like a voltage does on a charged particle. Magnetic and electric forces have a direction as well as a magnitude and are therefore represented by vectors. The distribution in space of electric and magnetic forces is described by a collection of vectors with a magnitude and direction for every point in space. The representation of these fields is usually an arrangement of lines of force eminating from one pole (north or south pole for magnetism and positive or negative charge for electricity) and terminating at the opposite pole. The greater the path length, the weaker the force at points along that path. (This makes sense if you consider the overall strength of the force between the poles is constant and the distance over which the force is exerted grows as the path lengthens. Therefore the difference in force per unit of distance decreases as the path length increases.)

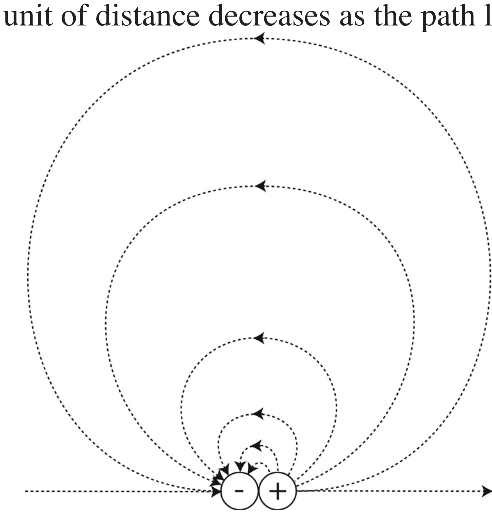

Figure 10 Dipole lines of force

When a force causes the movement of an object, energy is transfered to the object and work is done. Energy may take one of two forms: potential and kinetic. An object at rest may possess potential energy but its kinetic energy is zero. A moving object has kinetic energy and may also have potential energy. The amount of kinetic energy in a physical system is:

$$K = \frac{1}{2}mv^2$$

where K is the kinetic energy (measured in joules), m is the mass and v is the velocity. Potential energy represents the ability to do work even if no work is currently being done: a ball at the top of a hill, for example. Mechanical systems allow for the back and forth interchange of potential and kinetic energy, however the total amount of energy in a closed system remains constant. In the case of gravitational potential energy, the relationship is represented by the equation:

$$U(y) = mgy$$

where U(y) is the potential energy at a height of y (meters), m is mass and g is the force of gravity. There are other forms of potential energy, tension in a spring for example. Voltage, also called potential, is a form of potential energy.

It will help in the explanation of how microphones operate and how electronic circuits function if we keep in mind the simple physical interactions that underly the more complex processes. Our physical models of systems are correct only to some level of approximation. In most cases we can describe a system well enough to understand the basic operations involved in recording sound with simple mathematical descriptions, realizing there are often more subtle factors at play if we carefully scrutinize the processes. A complete picture of the magnetic field generated by a record head and its interaction with tape takes an entire book, still we can understand how bias current adjustments shape the recovered sound recording by using a quite simplified model. Circuit designers sometimes find they cannot explain what they hear even by detailed analysis, but we are at first concerned with understanding the everyday audio world.