A Scientific Exploration of How Sound Becomes Music

Luis A. Anchordoqui,[1,2,3] Gabriela V. Anchordoqui,[4] and Luca Girolano Anchordoqui[5]

[1] *Department of Physics & Astronomy, Lehman College, City University of New York, NY 10468, USA*
[2] *Department of Physics, Graduate Center, City University of New York, NY 10016, USA*
[3] *Department of Astrophysics, American Museum of Natural History, NY 10024, USA*
[4] *Sociedad Odontológica La Plata, Calle 13 No. 680, La Plata 1900, Argentina*
[5] *Facultad de Ingeniería, Universidad Nacional de La Plata, La Plata 1900, Argentina*

(Dated: Spring 2026)

I. GENERAL PRINCIPLES OF EVERYDAY PHYSICAL PHENOMENA

A. Newtonian Mechanics

Newtonian mechanics is a branch of classical physics that describes the motion of macroscopic objects using concepts of force, mass, and acceleration, based on principles formulated by Newton in 1687 [1]. It acts as a foundational, predictive model for everyday objects moving at speeds much slower than the speed of light. Here is a summary of the basic concepts:

1. *Position, path length, and displacement*

Position can be understood as the location of an object with respect to a reference point. It is expressed as a position vector. Shown in Fig. 1 is the position of a cyclist with respect to a lamp post. The cyclist starts at the lamp post and goes to positions one (P1 = +3 units), two (P2 = +5 units), three (P3 = −2 units), and four (P4 =−6 units), respectively. The path length (or distance) is the total length of the path covered by the moving object. In Fig. 2 we show the distance (path lengh = 16 units) covered by the same cyclist. The displacement is the shortest distance between the initial and final positions

$$\Delta\vec{x} = \text{final position} - \text{intial position}\,. \qquad (1)$$

The displacement is a vector.[1] The magnitude of the displacement may or may not be equal to the distance traveled. If the object does not change direction during motion, then the distance is equal to the magnitude of the displacement. If the object changes direction during motion, then the distance is always greater than the displacement. The total displacement of the cyclist is shown in Fig. 3, whereas the displacement between position P2 and position P4 is shown in Fig. 4.

We will regularly use the metric system; in this system the prefix *deci* means 1/10, *centi* means 1/100, *milli* means 1/1,000, *micro* means 1/1,000,000, *kilo* means 1,000, and *mega* means 1,000,000. Units of length or can be meters (m), centimeters (cm), millimiters (mm), kilometers (km), etc.[2]

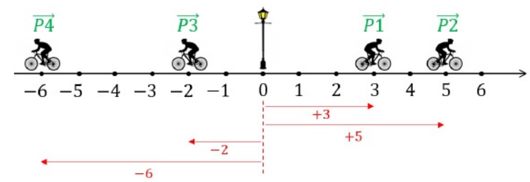

FIG. 1: Position of a cyclist with respect to a lamp post. Cyclist starts at the lamp post and goes to positions P1, P2, P3, and P4 respectively.

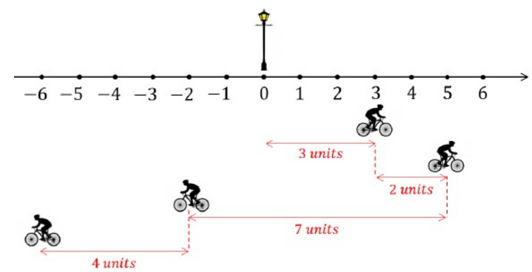

FIG. 2: Distance (total path length) covered by the same cyclist.

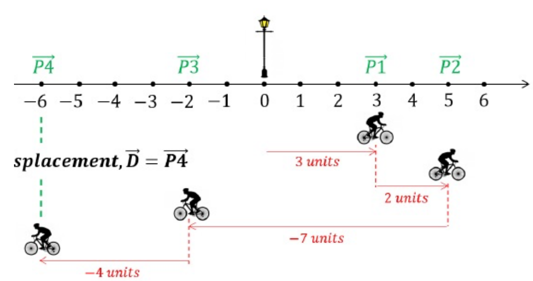

FIG. 3: Total displacement of the cyclist.

[1] A vector is a quantity that has both a magnitude (a size or length) and a direction. It can be represented by a direct line segment, with the arrow's length showing the magnitude and the arrow's point showing the direction.

[2] One meter is approximately 39.37 inches; 1 inch (in) is 2.54 cm.

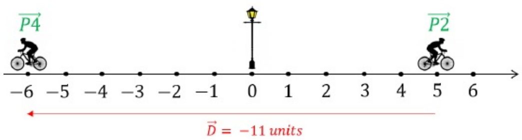

FIG. 4: Displacement of the cyclist between position P2 and position P4.

2. *Time*

Time is the ongoing sequence of events from the past, through the present, and into the future. While it is a fundamental part of existence, its definition varies significantly across different fields of study.

In Newtonian physics, time is defined as an absolute, universal entity that exists independently of space and any physical observer. Often referred to as "absolute time," it is conceptualized as a continuous and unchanging "flow" that serves as the fixed backdrop for all physical events.[3]

Units of time can be seconds (s), hours (hr), years (yr), eons.

3. *Velocity*

In Newtonian physics, velocity ($\vec{v}$) is a vector quantity representing the rate of change of an object's position with respect to time, encompassing both its speed (magnitude) and direction of motion. It is defined as the displacement divided by the time interval

$$\vec{v} = \frac{\text{final position} - \text{initial position}}{\text{final time} - \text{initial time}} = \frac{\Delta\vec{x}}{\Delta t} \tag{2}$$

Unlike speed, velocity changes if direction changes, even if the speed is constant.

4. *Mass*

In Newtonian physics, mass (m) is an intrinsic scalar property representing the quantity of matter in an object, its inertia (i.e. its resistance to changes in motion), and its strength in gravitational interactions. Newton originally defined it as the "quantity of matter" arising from its density and volume combined. In the International System of Units (SI) mass is measured in kg.

5. *Force*

In Newtonian physics, a force ($\vec{F}$) is a vector quantity – possessing both magnitude and direction – that acts as a push or pull on an object resulting from its interaction with another object. Forces cause objects to change their velocity (speed and/or direction), or shape. The SI unit is the Newton (N), where 1 N = 1 kg · m/s.

Among the most commonly experienced and well-known forces in daily life are: the gravitational force (acting between all objects with mass or energy, drawing them together) and the frictional force (that opposes motion between surfaces in contact). They represent two distinct categories of interaction: non-contact forces (gravity) and contact forces (friction).

For a set of forces acting on an object, the net force (or resultant force) is the vector sum of all individual forces,

$$\vec{F}_{\rm net} = \sum_{i=1}^{n} \vec{F}_i = \vec{F}_1 + \vec{F}_2 + \cdots + \vec{F}_n \,; \tag{3}$$

it represents the combined effect of all pushes and pulls on an object.[4]

6. *Newton's First Law*

Newton's First Law of Motion, also known as the principle of inertia, states that an object at rest stays at rest, and an object in motion stays in motion with the same speed and in the same direction, unless acted upon by an unbalanced external force. This means objects resist changes in their state of motion; a stationary object needs a force to start moving, and a moving one needs a force to stop or change direction.

The principle of inertia implies that there is no fundamental physical distinction between being at rest and moving at a constant velocity.

7. *Acceleration*

In Newtonian physics, acceleration ($\vec{a}$) is the vector rate of change of an object's velocity with respect to time, involving changes in speed, direction, or both. Defined as

$$\vec{a} = \frac{\text{final velocity} - \text{initial velocity}}{\text{final time} - \text{initial time}} = \frac{\Delta\vec{v}}{\Delta t} \,, \tag{4}$$

---

[3] Unlike in Einstein's theory of relativity [2], Newtonian time is not linked to space; they are distinct, separate containers where space is the "where" and time is the "when." Unlike spatial dimensions where one can move in any direction, time appears to move in only one direction: from past to future. This is the so-called "arrow of time" that is often explained by the second law of thermodynamics; for details see Sec. I C 2.

[4] The summation symbol $\sum$ represents a compact, efficient way to write the sum of a sequence of terms following a pattern. For the example given in Eq. (3), it indicates an addition over a set of forces, where the index starts at the lower bound ($i = 1$) and increases by 1 until the upper bound n is reached, substituting each value into $\vec{F}_i$.

it is directly proportional to the net external force acting on a body $\vec{F}_{\text{net}}$ and inversely proportional to its mass (m), per Newton's Second Law

$$\vec{F}_{\text{net}} = m\vec{a}. \tag{5}$$

Near the Earth's surface, the acceleration due to gravitational attraction is approximately $g \simeq 9.8\ \text{m/s}^2$ (or equivalently 32.2ft/s^2).

8. Newton's Third Law

Newton's Third Law states that forces always occur in pairs: if one object exerts a force on a second object, the second object simultaneously exerts a force of the same magnitude but in the opposite direction on the first object, acting on different bodies. A common example is that of a rocket pushing gas down (action) and the gas pushing the rocket up (reaction).

B. Basics of Energy and Its Various Forms

Energy is probably the most important concept in physics because it pervades all the branches of this discipline. One speaks of mechanical energy, gravitational energy, thermal energy, electric energy, sound energy, chemical energy, atomic energy, etc. But, what is energy? Energy is the fundamental capacity to do work, drive motion, and produce heat or light, existing in kinetic (motion) [3] and potential (stored) [4] forms.

1. Work

Work is the measure of energy transfer that occurs when an object is moved over a distance by an external force. Work is done only by the portion of the force that acts in the same direction as the object's movement. For example, if you pull a wagon at an angle, only the "forward" part of your pull counts toward work; the part pulling "up" does no work if the wagon does not move upward.

The most complete way to express work done by a constant force is:

$$W = F\, x\, \cos\theta\,, \tag{6}$$

where F is the magnitude of the force, x is the magnitude of the displacement, and θ is the angle between the force vector and the displacement vector.[5] This implies that if force and movement are in the same direction, work is simply

$$W = Fx\,, \tag{7}$$

because $\cos 0^\circ = 1$. If the force is perpendicular to the movement, then the work done by that specific force is zero, because $\cos 90^\circ = 0$. If the force opposes the movement, the work is negative because energy is being removed from the object,

$$W = -Fx\,; \tag{8}$$

note that $\cos 180^\circ = -1$.

The joule (J) is the SI derived unit for energy, defined as the work done by a force of one newton acting over a distance of one meter ($\text{J} = \text{N}\cdot\text{m}$).

2. Kinetic energy

Kinetic energy is the energy a body possesses as a result of its motion. It is formally defined as the work needed to accelerate a body from rest to its current velocity. In Newtonian mechanics, this work (and thus the kinetic energy) is equal to:

$$E_{\text{kinetic}} = \frac{1}{2} m v^2\,, \tag{9}$$

where m is the object's mass and v its speed. Having gained this energy during its acceleration, the body mantains this kinetic energy unless its speed changes. Negative work of the same magnitude would be required to return the body to a state of rest from that velocity.

3. Potential energy

Potential energy is energy stored in an object due to its position, configuration, or state. It is a property of a system, representing the potential for that system to perform work.

Examples include:

- The gravitational potential energy, which is stored due to an object's height in a gravitational field. Near the surface of the Earth, it is calculated as

$$E_{\text{gravitational}} = mgh\,, \tag{10}$$

where m is the object's mass, h its height, and g is the gravitational acceleration.

[5] Sine (sin) and cosine (cos) are fundamental trigonometric functions used to relate the angles of a right triangle to the ratios of its side lengths. In a right-angled triangle, for a given acute angle θ, the sin is the ratio of the side opposite the angle to the hypotenuse ($\sin\theta$ = opposite/hypotenuse), whereas the cos is the ratio of the side adjacent to the angle to the hypotenuse ($\cos\theta$ = adjacent/hypotenuse).

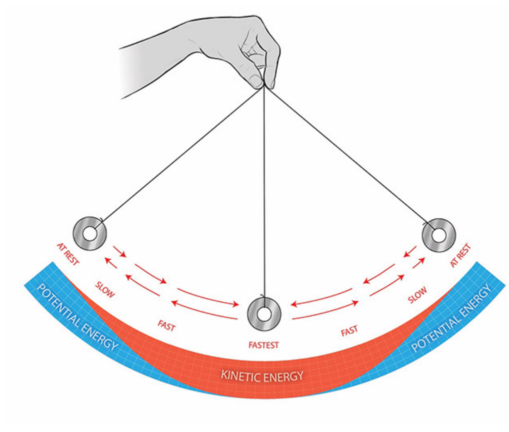

FIG. 5: A pendulum is a weight (bob) suspended from a fixed point so that it can swing freely back and forth under the influence of gravity. Its regular, predictable motion makes it useful in various applications, most famously in clocks.

- The eleastic potential energy, which is stored in objects that can be compressed or stretched, like springs. It is given by

$$E_{\text{elastic}} = \frac{1}{2}kx^2 \,, \tag{11}$$

where k is the spring constant and x is the displacement.

4. Conservation of mechanical energy

In an ideal system without friction or air resistance the total mechanical energy (kinetic + potential) stays constant, allowing energy to convert between kinetic and potential forms but never changing the total sum,

$$\Delta E_{\text{kinetic}} + \Delta E_{\text{potential}} = 0 \,. \tag{12}$$

This means that mechanical energy is never "lost"; it simply converts from one form to another. For example, a swinging pendulum (see Fig. 5) constantly swaps energy. It has maximum potential energy at its highest points and maximum kinetic energy at its lowest.

5. Power

Power is the rate at which work is done or energy is transferred over time. The average power is expressed by the formula

$$\langle P \rangle = \frac{\Delta W}{\Delta t} \,, \tag{13}$$

where $\langle \cdot \rangle$ denotes the average of the $\cdot$ quantity. Power is measured in Watts (W = J/s). It quantifies how quickly a specific amount of work is completed. It is essentially the speed of energy usage, with higher power meaning more work is done in less time. While "work" tells us how much energy is being used, power tells us how fast that energy is being used.

C. Macroscopic Properties and Microscopic Models

Macroscopic properties (e.g., pressure, temperature, volume) describe the bulk, observable state of matter, while microscopic models (e.g., atomic, molecular) explain these behaviors based on individual particle interactions. In this section we will dive deeper into these concepts, for further reading see e.g. [5].

1. Kinetic Theory

Kinetic theory bridges the gap between microscopic particle behavior and macroscopic bulk properties of matter (specifically gases) by explaining observables like pressure and temperature as results of molecular motion. Microscopic models assume molecules are in constant, random motion, exerting forces through collisions, while macroscopic properties are averages of these actions. For example, volume, density, and pressure are fundamental properties, defining how matter occupies space, packs together, and interacts with its surroundings. More concretely, the *volume* of a physical system is how much space it takes. We generally refer to it with the symbol V. The volume is usually obtained multiplying three different lengths for a given shape, so the units of volume are those of length^3.

Density is a measure of mass per unit volume, representing how tightly matter is packed within a substance. It is calculated by dividing an object's mass by its volume,

$$\text{density} \equiv \rho = m/V \,. \tag{14}$$

Pressure is the force exerted perpendicular to an object's surface per unit area A

$$\text{pressure} \equiv P = F/A \,. \tag{15}$$

While internal gas pressure can cause outward expansion (like in a balloon), pressure generally measures force concentration in any direction – inward, outward, or downward – depending on the context. It acts in all directions at once. Note that pressure depends on both the amount of force (e.g., gravity, molecular motion) and the area it is spread over. A smaller area results in higher pressure for the same force.

Atoms are the fundamental building blocks of ordinary matter, consisting of a dense, central nucleus of positively charged protons and neutral neutrons, surrounded by a cloud of negatively charged electrons [6–10]. Ranging in size around 10^{-10} m, they form the smallest unit of an element that retains its chemical properties.

The bodies that make up our everyday environment are made up of atoms that are attached to other atoms by bonds that hold them together, forming molecules.

Atoms are electrically neutral, with equal protons and electrons, whereas ions are atoms (or molecules) with a net electric charge due to the loss or gain of electrons. Positively charged ions (cations) lose electrons, while negatively charged ions (anions) gain them, and these charged particles are crucial for biological signaling and conducting electricity.

Protons and neutrons make up more than 99.9% of the atomic mass, while electrons occupy most of the volume, which is primarily empty space. The *atomic number* defines an element by its proton count, while isotopes are atoms of the same element (same atomic number/protons) that have different numbers of neutrons, leading to different mass numbers and atomic masses, like Carbon-12 (6 protons and 6 neutrons) and Carbon-14 (6 protons and 8 neutrons). So, the atomic number stays constant for an element, but isotopes vary in mass.

Solids, liquids, and gases are the three primary phases of matter, distinguished by particle arrangement. Solids have a fixed shape and volume with tightly packed, vibrating molecules. Liquids have a fixed volume but take the shape of their container. Gases have no fixed shape or volume, filling containers with freely moving particles.

Gases contain particles (atoms or molecules) that move around very quickly. The particles travel in straight lines until they bounce off another particle or a surface. Gas particles are widely spaced and tend to be only slightly attracted to each other. *Temperature* (T) is a macroscopic measure of the average kinetic energy of gas particles; increasing temperature increases molecular velocity. While a single atom (or molecule) has energy, it does not have temperature; temperature only describes the collective, average motion of a large number of particles. Throught these notes we will measure temperatures using the Kelvin (K) scale [11]. While the Fahrenheit [12] and Celcius [13] are relative temperature scales, Kelvin is absolute, anchored at 0 K which is the theoretical temperature at which molecular activity is nonexistent.

The properties of a gas are almost independent of its identity. This implies that gas molecules behave as if no other molecules are present. The ideal gas law

$$PV = Nk_{\rm B}T\,, \tag{16}$$

relates the pressure P, volume V, number of molecules N, and absolute temperature T, with k_B serving as the proportionality (Boltzmann) constant.

The *internal energy* (U) of a substance, arising from the random or disorganized molecular motion, is primarily dependent on temperature. The internal energy of a monatomic ideal gas is

$$U = \frac{3}{2}k_{\rm B}T \tag{17}$$

per molecule.

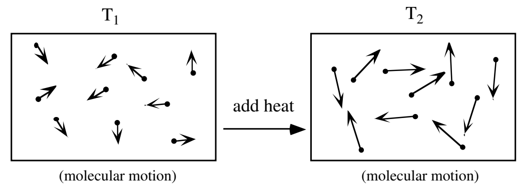

FIG. 6: Internal energy U is the total microscopic energy contained within a substance, directly proportional to its temperature. Higher temperature increases the molecular motion, directly raising the internal energy.

2. *Laws of Thermodynamics*

The four laws of thermodynamics are fundamental, universal principles governing energy, heat, and entropy interactions. They dictate that energy cannot be created or destroyed (*first law*) [14], entropy always increases in closed systems (*second law*) [15], absolute zero is unattainable (*third law*) [16], and define thermal equilibrium (*zeroth law*), mapping how energy moves and transforms throughout the universe. Understanding how these laws work requires introducing new concepts first.

A thermodynamic *system* is a specific, defined portion of matter or radiation in the universe chosen for study, separated from its surroundings by real or imaginary boundaries. It exchanges energy or matter with its environment, allowing analysis of properties like pressure, volume, and temperature via thermodynamic laws. Thermodynamic systems are classified by how they exchange energy and matter with their surroundings: open systems exchange both, closed systems exchange only energy, and isolated systems exchange neither.

In thermodynamics, *heat* Q is defined as the transfer of energy between two systems or a system and its surroundings due to a temperature difference [17, 18]. It is fundamentally energy in transit, moving from a hotter body to a cooler one until thermal equilibrium is reached, resulting in both bodies having the same temperature.

The *zeroth law* states that if two systems are both in thermal equilibrium with a third system, then they are in thermal equilibrium with each other. This law provides the basis for measuring temperature.

The *first law* is a variation of the law of conservation of energy that is adapted for thermodynamic systems. It can be stated as follows: The total energy of an isolated system is constant, where energy can be transformed from one form to another but can be neither created or destroyed. In an equation this is written as:

$$\text{Change in internal energy of a system} = \text{Heat added to the system} - \text{Work done by the system}\,, \tag{18}$$

or equivalently

$$\Delta U = Q - W\,. \tag{19}$$

The signs in Eq. (19) are critical and potentially confusing; thus, it is necessary to explicitly clarify that Q is positive if the heat is transferred to the system (see Fig. 6) and negative if it is transferred from the system, whereas W is postive if the work is done by the system and negative if it is done on the system. One of the implications of this law is that it is impossible to create perpetual motion machines that will produce work with no input of any energy source. Perpetual motion machines are an impossibility.

The second and third laws of thermodynamics concern themselves with the phenomenon known as entropy. *Entropy* was first introduced to describe the waste heat or energy loss from heat engines, and as such was related to the efficiency of mechanical devices in doing work. Later, this concept evolved into a concept of disorder, which related to the kinetic energy of particles in a system where eventually all particles will reach a state of high disorder and high entropy. More recently entropy is being interpreted as *energy dispersal* of a system, specifically the spread of temperature throughout a system as a function of temperature. Putting all this together, the change in entropy (ΔS) is defined as the heat transferred (Q) divided by the absolute temperature (T) at which the transfer occurs:

$$\Delta S = \frac{Q}{T}\,. \tag{20}$$

It measures the molecular disorder or energy dispersal in a system, typically expressed in Joules per Kelvin (J/K).

Bearing this in mind, the *second law* satates that in any spontaneous process, energy tends to disperse or spread out from a concentrated state to a more dilute, disordered, or lower-temperature state. This inevitable dispersal increases the total entropy of the universe, meaning energy becomes less available to do useful work. It establishes that heat flows spontaneously from hot to cold (arrow of time) and that no machine can be 100% efficient.

The *third law of thermodynamics* states that the entropy of a system will approach a constant as the system gets closer to absolute zero. The third law then implies that it is impossible for any system to ever reach absolute zero.

All in all, the *law of conservation of energy* states that energy cannot be created or destroyed, only transformed or transferred, keeping the total energy in an isolated system constant. It is illustrated by the example shown in Fig. 7.

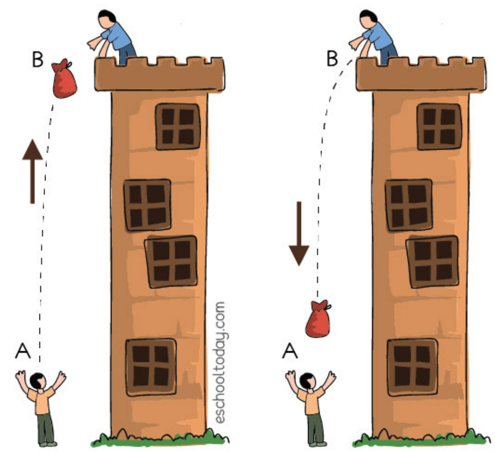

FIG. 7: Consider Mr. Green is throwing a bag of coins to Mr. Red, who is up a tower. As Mr. Green throws the bag from A to B, work is transferred from Mr. Green to the bag, which now has kinetic energy. As the bag moves upward, kinetic energy decreases, and gravitational energy increases. At the highest point B, the kinetic energy is zero and the gravitational potential energy is the highest. As the bag did not get to Mr. Red, it starts falling from point B downwards due to gravity. It starts falling slowly (kinetic energy is low) and then speeds up downwards. At point A, the bag is at full speed and the kinetic energy is the highest, while gravitational energy is nearly lost. When the bag of gold coins hits the ground, the kinetic energy is converted into heat and sound by the impact.

II. THE INS AND OUTS OF SOUND WAVES

A. Periodic Motion

The motion that repeats itself after fixed intervals of time is called periodic motion. For example, the oscillatory motion is a type of periodic motion. Oscillatory motion is defined as the to and fro motion of an object from its mean position. For instance, the oscillation of simple pendulum, vibrating strings of musical instruments, vibration prongs of tuning fork, etc. A simple form of oscillatory motion is the simple harmonic motion (SHM). In this motion, the restoring force is directly proportional to its displacement from its equilibrium position. The study of oscillatory motion is basic to physics; its concepts are required for the understanding of many physical phenomena. A wave is an example of oscillatory motion.

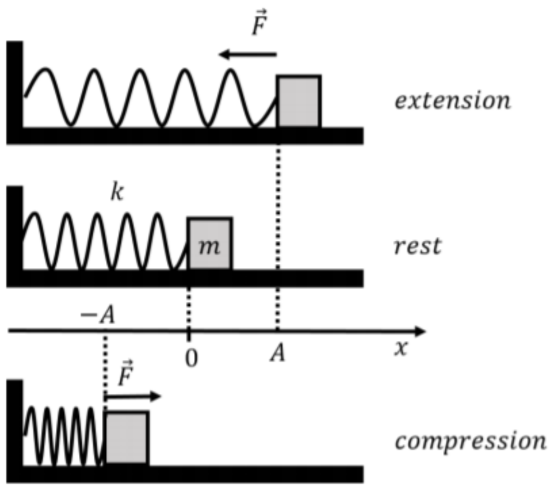

FIG. 8: SHM of a mass attached to a spring.

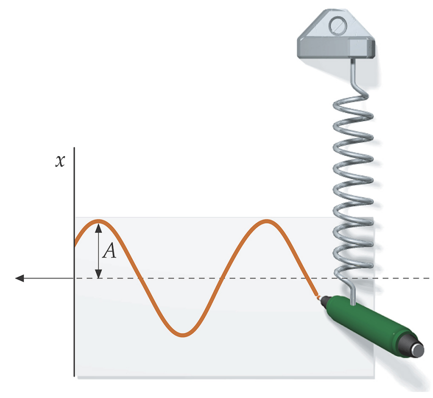

FIG. 9: A marking pen is attached to a mass on a spring and the paper is pulled to the left. As the paper moves with constant speed the pen traces out the displacement x as a function of time. In this example we have chosen x to be positive when the spring is compressed.

1. *Simple Harmonic Motion*

Two conditions must be met to produce SHM in a mechanical system.[6] First an equilibrium position must exist, *viz.* there must be one position at which the object executing the motion would be at rest if place there and to which it returns if displaced and release. Second, the force tending to pull the object back to its equilibrium position must be a *linear restoring force*, *viz.* the force must be linearly proportional to the distance of the object from the equilibrium position, as we explain next.

Consider the motion of a block of mass m that can slide without friction along a horizontal surface. The mass is attached to a spring of constant k, which is anchored to a wall on the other end. We introduce a one-dimensional coordinate system to describe the position of the mass, such that the axis is co-linear with the motion, the origin is located where the spring is at rest, and the positive direction corresponds to the spring being extended. The *spring-mass system* is illustrated in Fig. 8.

If the spring is compressed (or extended) by a distance A relative to the rest position, and the mass is then released, the mass will oscillate back and forth between $x = \pm A$, which is illustrated in Fig. 8. When the mass is at $x = \pm A$, its speed is zero ($v = 0$), as these points correspond to the location where the mass turns around. If it requires 1 N of force to compress the spring 1 cm, it will require 2 N for a 2-cm displacement, 3 N for a 3-cm displacement, and so forth, because for this system the restoring force is linear. Likewise, if the restoring force is linear it will require 1 N to stretch the spring 1 cm from the equilibrium position as well.

This is an example of a linear restoring force, with a force per unit of stretch of 1 N/cm. Such a linear system is said to obey Hooke's law

$$\vec{F} = -k\vec{x}, \tag{21}$$

where x is the displacement and k the spring constant [20].[7]

For example, say the mass on the spring is moved 3-cm to the right of its equilibrium position and released, it begins to accelerate leftward, passing through its equilibrium position, and executes SHM. At time $t = 0$ the mass is released from rest, 3-cm away ($x = +3$ cm) the equilibrium position $x = 0$. The mass starts moving, builds up speed as it passes through the equilibrium point, then slows down to zero velocity by the time it reaches the point $x = -3$ cm to the left of the equilibrium point. The mass then starts moving faster to the right until it goes through the equilibrium point, then slows down to zero velocity by the time it reaches its original position at $x = +3$ cm. The motion then repeats every 2 s.

The maximum displacement A of the mass from the equilibrium position (in either direction) is called the *amplitude*, and the repetition time τ is called the *period* of the SHM. For our example, $A = 3\ cm$ and $\tau = 2$ s.

Now, the period is related to the frequency f, which is the numbers of periods (or cycles) per second

$$f = 1/\tau\,. \tag{22}$$

A high-frequency oscillation has many vibrations per second; to fit many vibations into a 1-second time interval, the period must be short. Hence, a high-frequency

[6] The description of SHM builds upon the content of the excellent textbook by Berg and Stork [19].

[7] Hooke's law applies only within the elastic limit, where the spring return to its original shape once the force is removed. If this limit is exceeded, the spring undergoes permanent deformation.

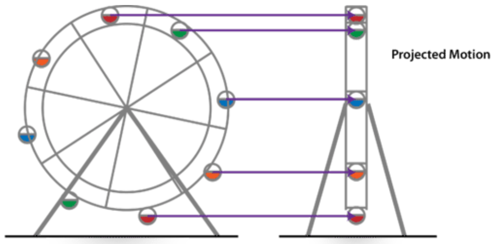

FIG. 10: SHM as the projection of uniform circular motion along a line in the plane of the paper.

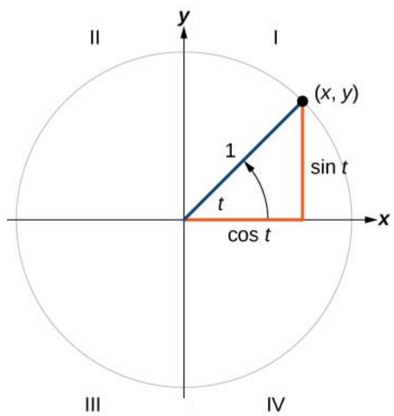

FIG. 11: A diagram of the unit circle showing the angular position θ changing with time (ωt), with all four quadrants I, II, III, and IV clearly marked. We have taken $\omega = 1$.

oscillation has a short period, whereas a low-frequency oscillation has a long period. In our example the frequency is found to be

$$f = \frac{1}{\tau} = \frac{1 \text{ cycle}}{2\,s} = 1/2 \text{ cycle per second} = 0.5\text{ Hz}\,. \quad (23)$$

The unit of frequency is the hertz (Hz), which is 1 cycle/s.

We can describe the motion of the mass using energy, since the mechanical energy of the mass is conserved. Recall that at any position x the mechanical energy $E_{\rm mechanical}$ of the mass will have a term from the potential energy $E_{\rm elastic}$, associated with the spring force, and kinetic energy $E_{\rm kinetic}$,

$$E_{\rm mechanical} = E_{\rm elastic} + E_{\rm kinetic}\,, \quad (24)$$

or equivalently

$$E_{\rm mechanical} = \frac{1}{2}kx^2 + \frac{1}{2}mv^2\,. \quad (25)$$

We can find the mechanical energy by evaluating the energy at one of the turning points. At these points, the kinetic energy of the mass is zero, so

$$E_{\rm mechanical} = E_{\rm elastic}\,(x = A) = \frac{1}{2}kA^2\,. \quad (26)$$

We can then write the expression for mechanical energy as:

$$E_{\rm mechanical} = \frac{1}{2}kA^2 = \frac{1}{2}kx^2 + \frac{1}{2}mv^2\,. \quad (27)$$

With a little bit of algebra, we can thus always know the speed of the mass at any position if we know the amplitude:

$$v = \sqrt{k(A^2 - x^2)/m}\,. \quad (28)$$

In Fig. 9 we show how we can experimentally obtain x versus t for a mass on a spring.

Some important details of SHM can be illustrated by comparison of circular harmonic motion, which is the motion of an object around a circle with constant speed. Uniform circular motion, view from a distant point in the plane of the motion is SHM as can be seen in Fig. 10. Indeed, SHM is the projection of uniform circular motion on a line in the plane of the motion.

Trigonometric functions are visualized on a unit circle (radius $r = 1$) by tracking the (x, y) coordinates of a point as it rotates counterclockwise. The y-coordinate represents the height and is associated to the $\sin\theta$ and the x-coordinate represents the horizontal position and is associated to the $\cos\theta$, where θ is the angular position measured from the positive horizontal x-axis. As the point moves, sin and cos oscillate between -1 and $+1$, repeating every 360° (or 2π radians). Consequently, the displacement in a SHM can be mathematically expressed as a sinusoidal function of time. Indeed, by taking $\theta = \omega t$ as indicated in Fig. 11, it is easily seen that a description of the curve shown in Fig. 9 is given by

$$x = A\sin(\omega t + \delta)\,, \quad (29)$$

where $\omega = 2\pi f$ is the angular frequency, and the phase constant δ, also known as the phase angle or phase shift, is a parameter in SHM that defines the initial state (position and velocity) of the oscillating mass at $t = 0$.[8]

Another example of SHM is a pendulum consisting of a mass or bob, hanging by a massless string from a fixed point and free to oscillate back and forth in a plane, as shown in Fig. 5. The equilibrium point is the point directly below the suspension point, where the pendulum will hang motionless. The force required to pull the pendulum away from its equilibrium position is linearly proportional to the displacement of the bob from equilibrium, so long as the amplitude is small. When displaced through some small angle and released from rest, the bob executes SHM in the plane of the paper. The motion can be described by a graph similar to the one in Fig. 9. The pendulum period is determined by its length

[8] We note in passing that δ shifts the SHM cosine or sine behavior along the horizontal axis, acting as a timing offset.

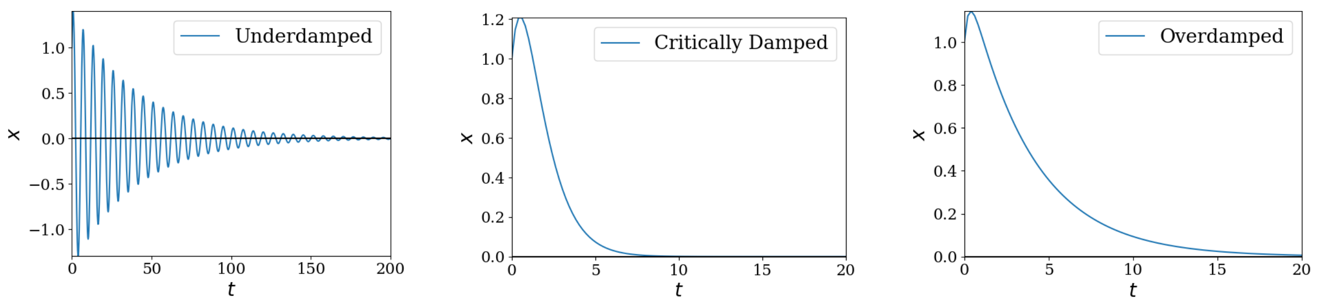

FIG. 12: Position (measured in m) versus time (measured in s) for an underdamped (left), critically damped (middle), and overdamp (right) oscillator.

(L) and the acceleration of gravity at the Earth surface (g) via

$$\tau = 2\pi \sqrt{L/g}\,, \tag{30}$$

and is independent of pendulum mass or amplitude. The energetic considerations have been described in Sec. I B 4.

2. *Damped and Driven Oscillations*

Damped and driven oscillations describe how real-world systems move when they lose energy to the environment and receive energy from an external source.

Damping occurs when resistive forces (like friction or air resistance) dissipate a system's energy. The amplitude of the oscillation decreases over time until the system eventually stops, see Fig. 12.

There are three main types of damping:

1. *Underdamped*, in which the system oscillates with a gradually decreasing amplitude;
2. *Critically damped*, in which the system returns to equilibrium as quickly as possible without oscillating;
3. *Overdamped*, in which the system returns to equilibrium slowly without oscillating.

For a linearly damped harmonic oscillator, the amplitude A decays exponentially over time according to a decay constant γ

$$A(t) = A_0 e^{-\gamma t} \tag{31}$$

with $e \simeq 2.72$. Substituting (31) into (26) we find that the mechanical energy decays as fast as

$$E_{\rm mechanical}(t) = \frac{1}{2} k A_0^2 e^{-2\gamma t} \tag{32}$$

This gives us the energy decay equation:

$$E_{\rm mechanical}(t) = E_0 e^{-2\gamma t} \tag{33}$$

The half-life of a damped oscillator is the time $t_{A_{1/2}}$ it takes for the amplitude of the oscillation to reduce to half of its initial value $A(t_{A_{1/2}}) = A_0/2$. It is instructive to determine $t_{E_{1/2}}$ defined by $E_{\rm mechanical}(t_{E_{1/2}}) = E_0/2$. For the amplitude, we have $e^{-\gamma t_{A_{1/2}}} = 1/2$, whereas for the energy, it follows that $e^{-2\gamma t_{E_{1/2}}} = 1/2$. Since both expressions equal 1/2, their exponents must be equal

$$\gamma t_{A_{1/2}} = 2\gamma t_{E_{1/2}} \tag{34}$$

Dividing both sides by γ we obtain

$$t_{A_{1/2}} = 2 t_{E_{1/2}} \Rightarrow t_{E_{1/2}} = t_{A_{1/2}}/2 \tag{35}$$

In summary, the energy reduces to half its value in $t_{A_{1/2}}/2$ because energy depends on the square of the amplitude. When the amplitude is squared, the exponential decay rate effectively doubles, meaning any percentage reduction in energy happens in half the time required for the same percentage reduction in amplitude.

A *driven* (or forced) oscillation occurs when a periodic external force is applied to a damped system to maintain its motion, causing it to vibrate at the driving frequency rather than its natural frequency. The phenomenon known as resonance occurs when a driven oscillator is made to vibrate at the same frequency as the driver oscillator. Resonance effects are characterized by vibrations with large amplitudes. Think about a child being pushed on a swing. The pusher (driver) pushes at the same frequency as the natural frequency of the oscillation of the swing (driven). The result is that the swing oscillates with large amplitude, since the energy it needs to gain altitude is being supplied at exactly the right time in the cycle.

B. Pressure Waves

A wave is a disturbance that transfers energy through space or a medium (solid, liquid, gas) without transporting matter. While the wave itself moves forward, the particles of the medium typically oscillate around a fixed equilibrium position and do not travel with the wave.

Waves are categorized according to the direction of oscillation:

- for transverse waves, the medium oscillates perpendicularly to the direction of energy travel (e.g., waves on a string);
- for longitudinal waves, the medium oscillates parallel to the direction of travel, creating regions of compression and rarefaction (e.g., sound waves in air).

This section covers the fundamental characteristics of longitudinal waves, specifically examining how particle oscillations occur parallel to energy transfer, resulting in regions of compression and rarefaction. We will analyze key properties such as wavelength, frequency, velocity, and amplitude.

1. Compression and Rarefaction Pulses

As we have seen in Sec. I C 1 , gas pressure represents the force exerted by molecules on their surroundings. The pressure distribution within a substance is not necessarily uniform. For example, localized pressure increases may occur near an active fan. Such disparities are transient and arise from differences in molecular density, thermal energy, or a combination of the two. Throughout this section, we focus exclusively on molecular concentration gradients, as temperature fluctuations in sound wave propagation are negligible by comparison and can be assumed constant. Molecular concentration is quantified as number density, defined as the total number of particles per volume,

$$n = N/V\,. \tag{36}$$

In Fig. 13 we show how two gas regions with disparate densities achieve equilibrium; the higher rate of molecular transfer from the denser region B to the sparser region A eventually results in a homogeneous state. Consequently, the system will, over time, spontaneously eliminate any initial pressure imbalances.

Gas pressure is often applied unidirectionally. For example, when inflating a balloon, molecules are forced into the interior from the outside in a specific direction. A fan operates similarly, creating directional airflow away from its blades.

An easy way to increase gas pressure is to force its molecules together, as shown in Fig. 14. When a device pushes molecules from the left, it creates a region of higher density (represented in red). Because molecules in this dense area move faster to the right than they do to the left, this compressed region propagates, or moves, through the gas. This moving area of high pressure is known as a compression pulse.

In a manner analogous to the process shown in Fig. 14, lowering the pressure (instead of raising it) reduces the density of gas molecules. This change spreads through the gas as a *rarefaction pulse*. Illustrated in Fig. 15, these blue regions of diminished density represent rarefaction, which is simply the inverse of compression.

2. Pressure Wave Dynamics and Fundamental Variables

The processes described in Sec. II B 1 can be generated repetitively, where compressions and rarefactions alternate in succession. This phenomenon, illustrated in Fig. 17, is known as a compression-rarefaction wave, or simply, a pressure wave. It is apparent from the figure that the gas number density n is spatially and temporally non-uniform. Since n depends on three position variables (one for each direction of space) and time, visualizing it requires simplification, as we can only plot two variables at once. To address this, we will reduce the problem to one dimension. By focusing on propagation along a single axis rather than all of space, we create a manageable model that still accurately describes realistic situations such as acoustic pipes (wind instruments) and vibrating strings.

The focus now shifts to how number density changes across space and time. Because plotting requires holding one variable constant, we can generate plots of n gainst time for a fixed position, or against position for a fixed instant in time. The side-by-side plots in Fig. 16 illustrate the spatial variation in number density for the wave depicted in Fig. 17. Each plot represents a fixed moment in time, essentially serving as a snapshot of the wave's entire length. The graphs in Fig. 18 illustrate density changes over time at a fixed location, functioning like a video of a single point along the wave. It is crucial to distinguish this from spatial representations, as the similarity between time-based and position-based plots of compression waves can easily lead to confusion. Always verify whether the horizontal axis represents time or position.

While wave representations in Figs. 16 and 18, depicted compression-rarefaction patterns, a key aspect was not explicitly mentioned: these waves are periodic, meaning the pattern repeats identically over time. The solid line in Fig. 19 represents the fundamental unit, which repeats indefinitely: a defining characteristic of all periodic waves.

In reality, the "perfect" square wave shown in Figs. 16 and 18 is impossible to achieve because it would require infinite bandwidth. Physical systems inherently act as filters that round off the sharp corners of a square wave, making it appear more sinusoidal (such as the SHM graph shown in Fig. 9) as high-frequency components are lost; see Fig. 20.

Because periodic waves follow a repeating pattern, they are straightforward to define. Even if the wave shapes shown in Fig. 20 look complex and hard to describe without a graph, they are characterized by simple, measurable properties: amplitude, period, frequency, and wavelength. We now turn to a detailed examination of these key properties:

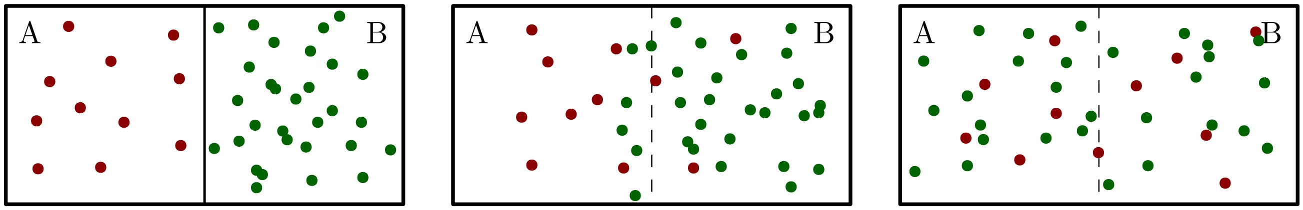

FIG. 13: Equilibration of two gases with unequal initial concentration. Because side B has a higher concentration of molecules than side A, the green molecules will migrate from B to A at a higher rate than red molecules traveling in the opposite direction.

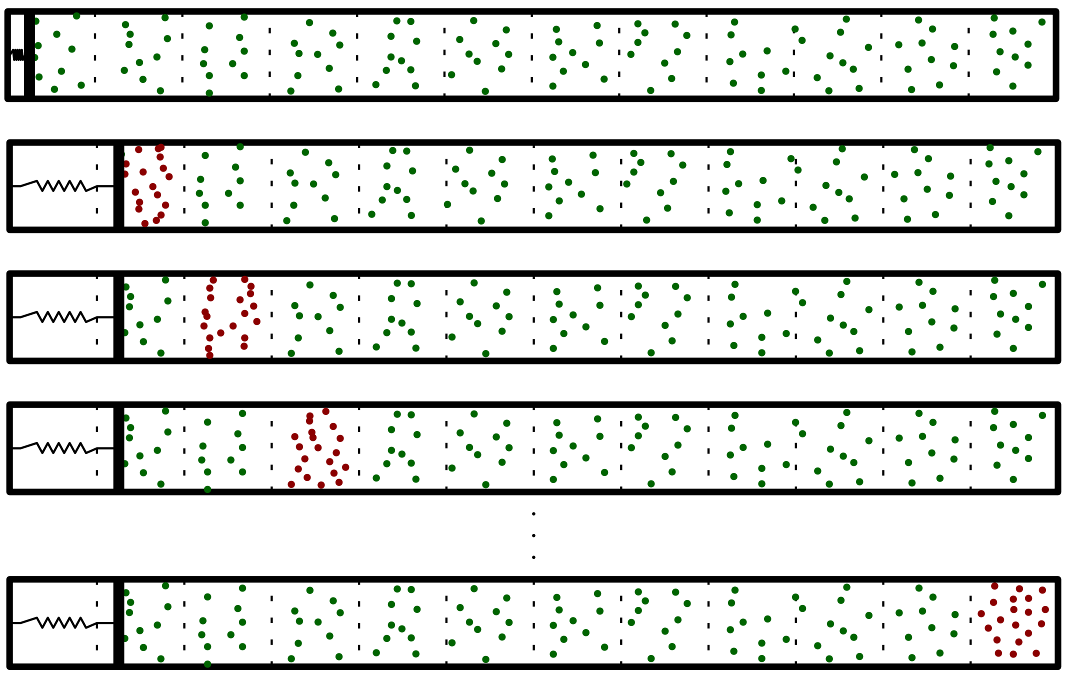

FIG. 14: Compression pulse (in red) moving towards the right.

- *Amplitude* (A) is the maximum displacement or distance moved by a point on a wave measured from its equilibrium (rest) position, see Fig. 21. It is directly related to the amount of energy the wave carries—a higher amplitude indicates more energy.
- *Frequency* (f) is the number of wave cycles that pass a specific point per second, measured in hertz (Hz). Higher frequency means more wave cycles pass in a given time.
- *Period* (τ) is the time it takes for one complete cycle of a wave to pass a point, see Fig. 22. It is the reciprocal of frequency (*viz.* $\tau = 1/f$) and is measured in seconds.
- *Wavelength* (λ) is the spatial distance between two corresponding points on consecutive waves, such as from crest to crest or trough to trough, see Fig. 22. It is measured in meters.

Figure 22 highlights the similarities and differences between period and wavelength.

3. *Speed of Wave Transmission*

The period and wavelength of a wave are linked by its propagation speed u. The relation between the times between peaks (period) and the distance between them (wavelength) is defined as

$$u = \lambda/\tau\,. \tag{37}$$

Since frequency is the reciprocal of the period, (37) can also be expressed as

$$u = \lambda f\,. \tag{38}$$

4. *Mapping Pressure Wave Oscillations*

A pressure wave can also be described in terms of particle displacement by relating the local pressure change

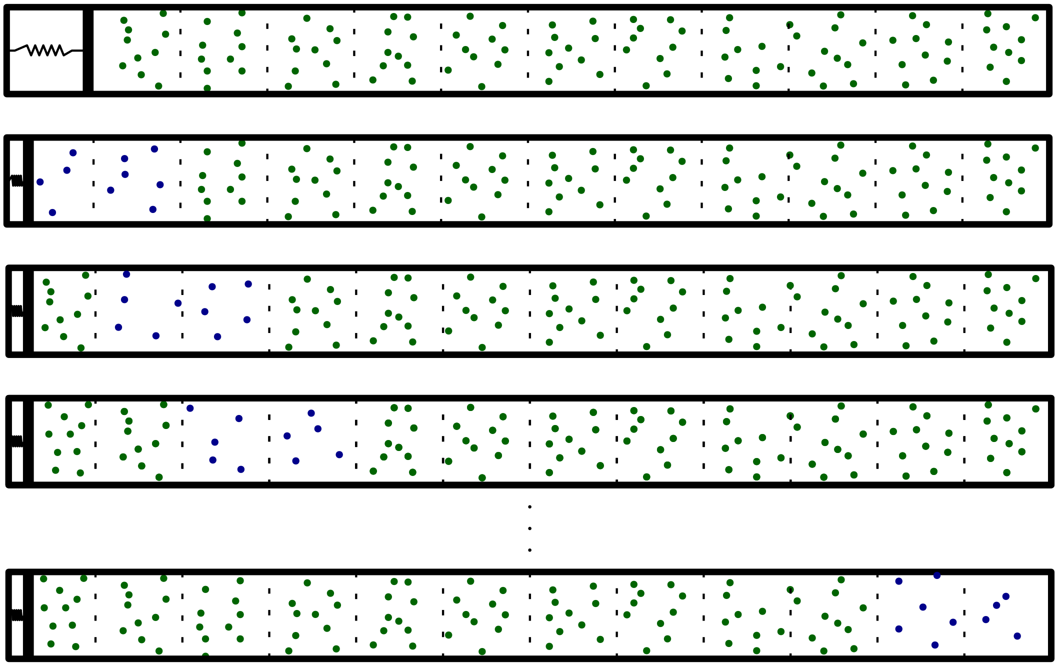

FIG. 15: Rarefaction pulse (in blue) moving towards the right.

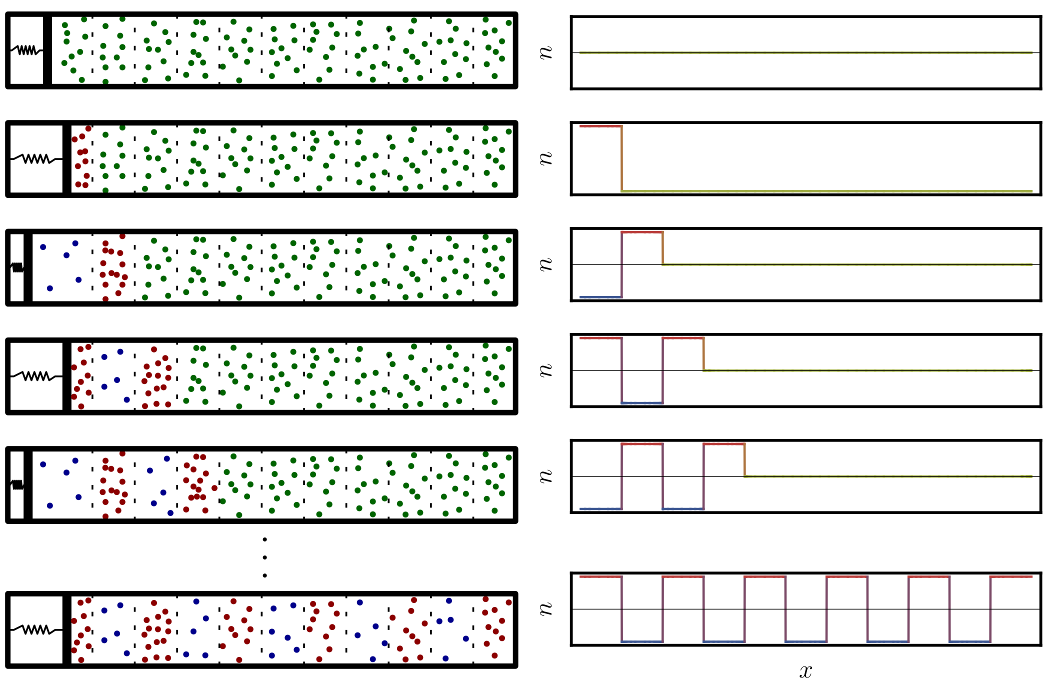

FIG. 16: Compression wave (left) and corresponding representation of the density as a function of the position (right).

to the gradient of the displacement wave.[9] Essentially, pressure fluctuations are caused by regions where particles in the medium are compressed together or spread apart.

More specifically, the pressure variation at a point is directly proportional to the negative spatial gradient of

[9] The gradient indicates how steep a slope is. Uphill slope is positive, whereas downhill slope is negative.

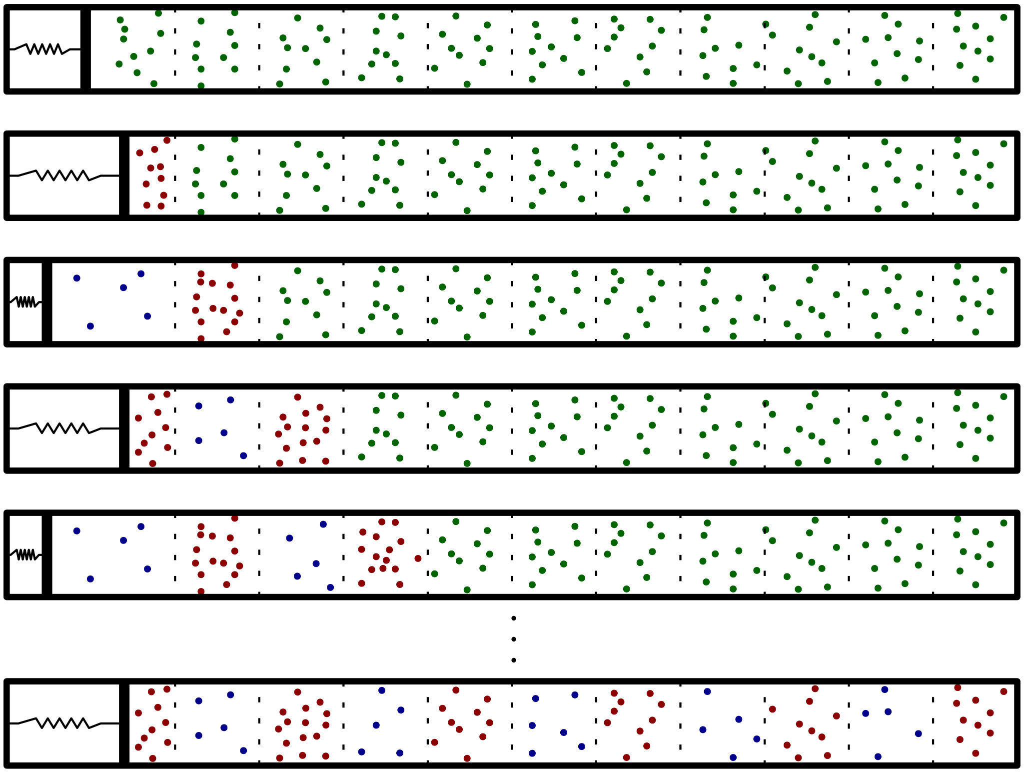

FIG. 17: A compression-rarefaction wave moving towards the right.

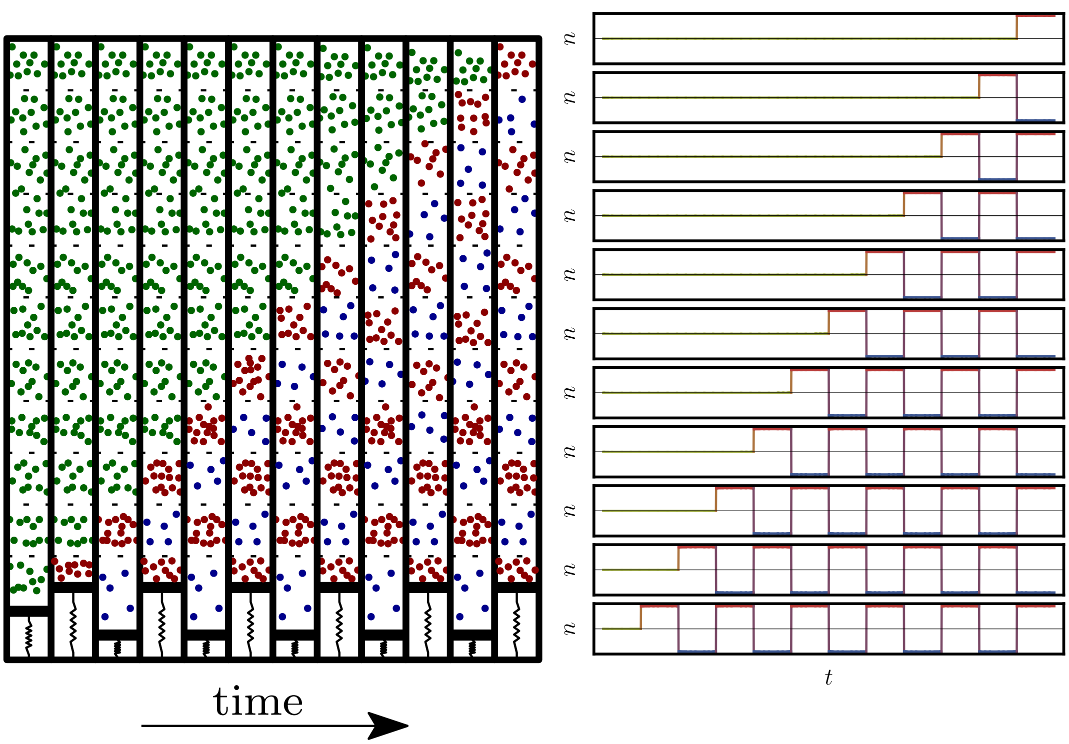

FIG. 18: Compression wave (left) and corresponding representation of the density as a function of the time (right).

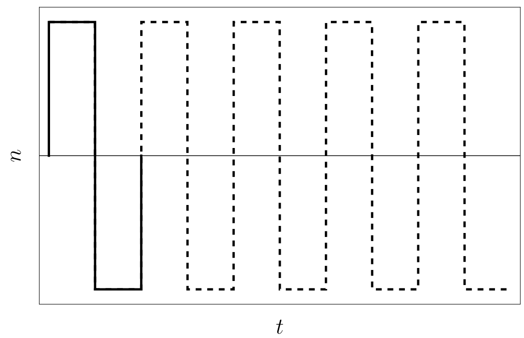

FIG. 19: The core, repeating unit, of a periodic wave.

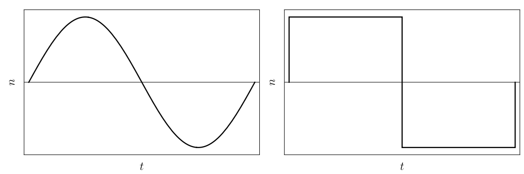

FIG. 20: Sinusoidal (left) and square (right) waveforms.

the displacement $x(t)$. Consequently, the pressure wave is 90° (or a quarter-cycle) out of phase with the displacement wave. This means that:

- maximum pressure (compression) occurs where the displacement gradient is steepest, which is at the points of zero displacement (equilibrium position);
- minimum pressure (rarefaction) also occurs at points of zero displacement, but where particles are moving away from each other;
- zero pressure change occurs at maximum or minimum displacement (turning points).

Notably, the maximum pressure fluctuation is directly proportional to the displacement amplitude A. These properties are encapsulated in Table I.

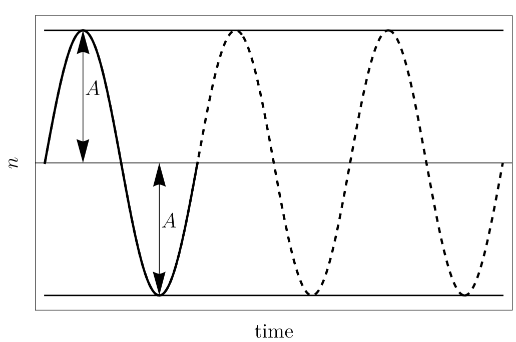

FIG. 21: A visual representation of the amplitude of a regular, periodic wave.

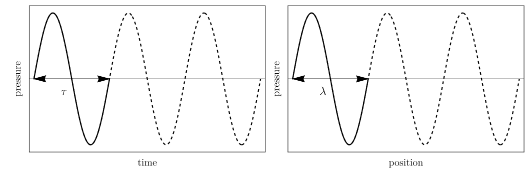

FIG. 22: Diagram showing the period (left) and wavelength (right) of a wave.

TABLE I: Displacement vs. pressure.

Point in wave	Displacement	Displacement gradient	Pressure
Compression	Zero	Negative	Maximum
Zero crossing	Maximum	Zero	Zero
Rarafaction	Zero	Positive	Minimum

C. Wave Superposition

Wave superposition dictates that the combined disturbance of multiple overlapping waves is the algebraic sum of their individual displacements. When waves occupy the same space, they interfere to produce a net, resultant wave without altering their original forms. In this section we examine this fundamental physical phenomenon.

1. *Constructive and Destructive Interference*

To understand the basic principles underlying interference. One of the most important properties of waves is the principle of superposition. The principle of superposition for waves states that when two waves occupy the same point, their effect on the medium adds algebraically. So, if two waves would individually have the effect +1 on a specific point in the medium, then when they are both at that point the effect on the medium is +2. If a third wave with effect −2 happens also to be at that point, then the total effect on the medium is zero. This idea of waves adding their effects, or canceling each other's effects, is the source of interference.

Constructive interference occurs when overlapping waves combine to form a resultant wave with a larger amplitude than the individual waves. Now, when two waves with the same frequency and phase (in-phase) meet, they undergo complete constructive interference. Because their crests and troughs perfectly align, their displacements add together at every point, resulting in a single wave with an amplitude equal to the sum of the individual amplitudes ($A_{\text{total}} = A_1 + A_2$).

As an illustration, imagine two waves with displacements

$$x_1 = A_1 \sin(\omega_1 t + \delta_1) \tag{39}$$

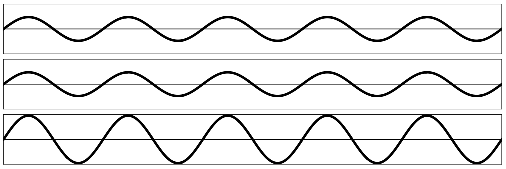

FIG. 23: Visual representation of the sum of two waves with complete constructive interference.

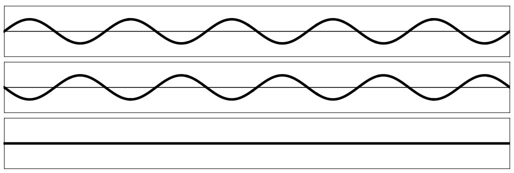

FIG. 24: Visual representation of the sum of two waves with complete destructive interference.

and

$$x_2 = A_2 \sin(\omega_2 t + \delta_2), \tag{40}$$

such that $\omega_1 = \omega_2 = \omega$ and $\delta_1 = \delta_2 = \delta$. Using the principle of superposition, the total displacement $x_{\rm total}$ is the sum of the individual displacements:

$$x_{\rm total} = x_1 + x_2 = A_1 \sin(\omega t + \delta) + A_2 \sin(\omega t + \delta)\,. \tag{41}$$

Because the $\sin(\omega t + \delta)$ term is common to both, we can factor it out to obtain

$$x_{\rm total} = (A_1 + A_2) \sin(\omega t + \delta)\,; \tag{42}$$

see Fig. 23.

Destructive interference occurs when overlapping waves subtract, reducing the overall amplitude. In particular, *complete destructive interference* occurs when two waves of identical frequency and amplitude are 180° (π radians) out of phase, meaning the crest of one aligns with the trough of another. As shown in Fig. 24, this results in the complete cancellation of the waves, producing a net amplitude of zero. In algebraic terms, this is expressed as

$$x_1 = A \sin(\omega t) \tag{43}$$

and

$$x_2 = A \sin(\omega t + \pi) = -A \sin(\omega t)\,, \tag{44}$$

where we used the trigonometric identity

$$\sin\theta = -\sin(\theta + \pi)\,. \tag{45}$$

Note that $x_1 + x_2 = 0$.

Usually, waves traveling in the same direction are neither perfectly in phase nor perfectly in counter-phase. That is, often, interference is neither constructive nor destructive, but somewhere in between. In Fig. 25 we show a large number of waves that are not aligned; rather, they exhibit random phase variations.

2. *Resolving Wave Patterns into Component Sinusoids*

Sinusoids occupy a foundational role in wave analysis, as Fourier analysis permits any complex waveform to be decomposed into constituent sine and cosine functions. More specifically, Fourier's theorem asserts that any periodic, continuous wave can be broken down into a sum of simple sinusoidal waves (partial waves) with specific amplitudes and phases. Imagine a highly complex, non-sinusoidal wave $W_{\rm non-sinusoidal}$. Fourier theorem states that

$$W_{\rm non-sinusoidal} = \sum_{n=1}^{\infty} w(A_n, f_n) \tag{46}$$

where each $w(A_n, f_n)$ is an individual sinusoidal paprtial wave with a specific frequency f_n and amplitude A_n. Recall that the $\sum$ symbol indicates that we are adding together multiple waves, so

$$W_{\rm non-sinusoidal} = w(A_1, f_1) + w(A_2, f_2) + \cdots\,. \tag{47}$$

Note that the sum consists of an infinite series of waves, which may appear computationally inefficient compared to simply using the single wave $W_{\rm non-sinusoidal}$. However, this is not a practical concern; because these waves are not equally significant, we do not need to use all of them.

Consider a wave decomposed into sinusoidal components with rapidly decreasing amplitudes (e.g., $A_1 = 100$, $A_2 = 10$, $A_3 = 1$). By the time you reach the tenth component ($A_{10} = 10^{-7}$), its amplitude is a billion times smaller than the first. Because its contribution is statistically negligible, you can discard it – along with all subsequent, even smaller waves – without significantly altering the result. Effectively, you can achieve a highly accurate approximation using only the first nine waves rather than an infinite series. While requirements vary, sometimes requiring 100, sometimes 1000, you do not need an infinite number of waves. Always prioritize retaining the waves with the highest amplitudes and discard the smaller ones.

As an illustration, consider the square wave at frequency f_0 shown in Fig. 26. By applying Fourier analysis (the derivation of which exceeds the scope of this course) the exact partial waves required to reconstruct this wave can be determined. Omitting the complicated derivation, these constituent waves are expressed as follows:

$$A_n = \frac{4}{\pi}\frac{1}{n} \tag{48}$$

and

$$f_n = n f_0\,, \tag{49}$$

with $n = 1, 3, 5, \cdots\cdots$.

It is straightforward to see that the first wave ($n = 1$) possesses the highest amplitude. Because the amplitude

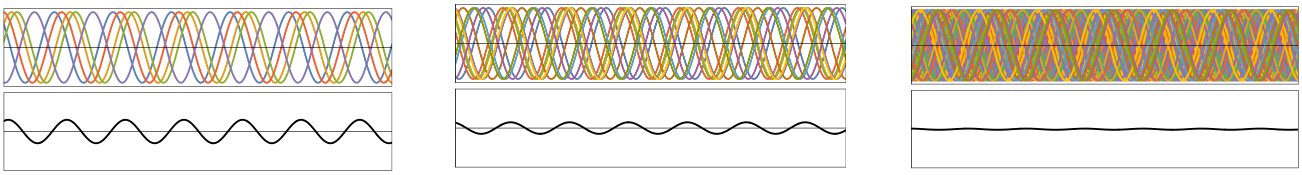

FIG. 25: Sum of 5 (left), 10 (middle), and 100 (right) randomly shifted waves.

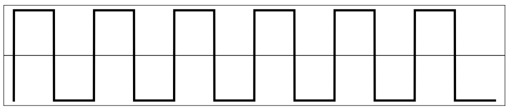

FIG. 26: Square wave.

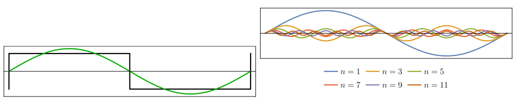

FIG. 27: Square wave with its fundamental harmonic (left) and the first six odd harmonics (right).

shrinks as n gets larger, the first partial wave is the most dominant, with all subsequent waves acting as minor refinements.

In the left panel of Fig. 27 we show the basic relationship between a square wave and its first partial wave $w(A_1, f_1)$, and in the right panel we display the first six partials ($n = 1, 3, 5, 7, 9, 11$). When these waves are combined,

$$W_{\text{square}} \approx \sum_{n=1}^{M} w(A_n, f_n) \tag{50}$$

the resulting sums for one, two, three, and four components are shown in the left panel of Fig. 28, whereas the addition of 10, 100 and 10000 of those waves are shown in the right panel of the figure.

D. Standing Waves

Up to now, we have been studying waves that propagate continuously through a medium, but we have not discussed what happens when waves encounter the boundary of the medium. Waves do interact with boundaries of the medium, and all or part of the wave can be reflected. For example, when you stand some distance from a rigid cliff face and yell, you can hear the sound waves reflect off the rigid surface as an echo. The result of successive reflections is the formation of standing waves (also known as a stationary waves). In this section, we explore wave behavior at boundaries and the key characteristics of standing waves.

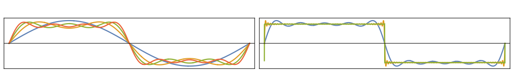

FIG. 28: Cumulative square wave approximation using $M = 1, 3, 5, 7$ (left) and $M = 10, 100, 10000$ (right).

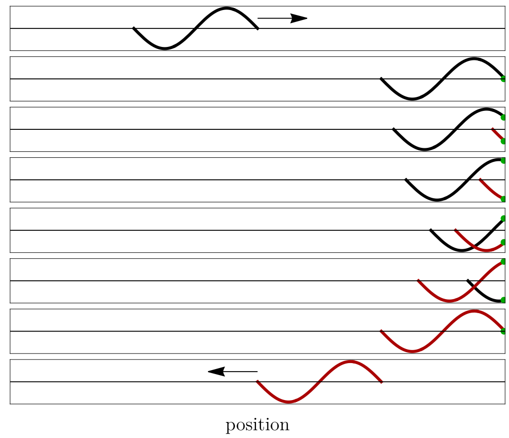

FIG. 29: Reflection of a wave with phase reversal.

1. *Phase Reversal*

When a pressure wave encounters a boundary, such as a wall or the open end of a tube, it reflects, traveling back in the opposite direction. Beyond simply reversing direction, the reflected wave can differ from the initial wave depending on the boundary conditions.

To visualize this, imagine a short pulse consisting of a single compression followed by a rarefaction. As the compression strikes a reflective surface, the local density increases, creating a new, reverse-traveling wave. The critical question is whether this reflected pulse begins with the same compression or with a *phase-reversed* rarefaction.

The phenomenon shown in Fig. 29 in which a compressed wave reflects as a rarefied wave is known as phase reversal. Conversely, a reflection without phase reversal means the wave returns without any alteration to its value, as shown in Fig. 30.

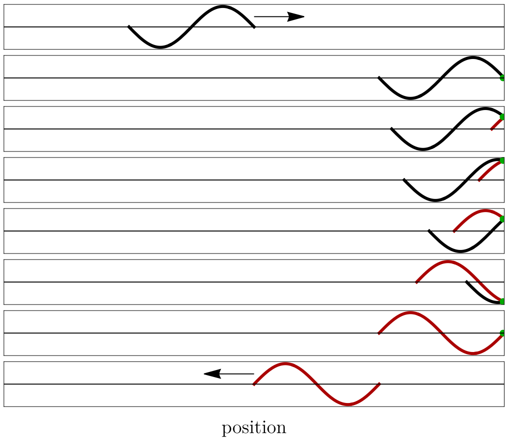

FIG. 30: Reflection of a wave without phase reversal.

2. *Waves on Cavities and Strings*

Building on our previous mathematical models, we will now explore physical systems that behave as one-dimensional (1D) environments. While everything in our world is technically 3D, we can simplify our description of objects that are significantly smaller in certain dimensions—much like treating a thin sheet of paper as 2D or a strand of hair as 1D. We will focus on two primary examples to study wave propagation:

1. Strings (solids): Long, thin systems like guitar strings.

2. Cavities (gases): Narrow enclosures filled with air, such as a flute.

While the 1D approximation is slightly less perfect for cavities than for strings, both provide an excellent framework for understanding standing waves and how sound is produced in physical structures.

Standing waves require a resonant system where waves can continuously reflect back and forth. Consequently, the strings or cavities described here cannot be infinite; they require a specific length with boundaries on both sides, creating a confined space for waves to oscillate. Figures 29 and 30 illustrate this reflection behavior within these confined systems.

As discussed previously, reflection can induce *phase reversal*, where a compression transforms into a rarefaction, and vice versa. In air systems, compression usually reflects as compression, and rarefaction as rarefaction, as seen in Fig. 30. However, if phase reversal occurs, these characters flip, a phenomenon that also occurs on vibrating strings; see Fig. 29.

Whether a string or cavity end induces this phase reversal depends on its boundary conditions. These ends can be categorized into two types:

1. for strings, a *fixed end* (where the string is attached to a rigid support) or a *loose end* (where the end is free to move, such as floating in the air);

2. for cavities, ends are either closed (closed by a rigid wall) or open to the atmosphere.

While the physics behind why specific ends cause or avoid phase reversal requires deeper analysis, we can take the following behaviors as given:

1. strings
    - fixed end → phase is reversed
    - loose end → phase is not reversed

2. cavities
    - open end → phase is reversed
    - closed end → phase is not reversed

3. *Nodes and Antinodes*

Nodes and antinodes are specific points along a standing wave. They represent the two extremes of how the medium (like a guitar string or air in a pipe) behaves during vibration.

A node is a point along a standing wave where the wave has zero amplitude. This point point appears to be standing still; there is no displacement of the medium. Nodes are the result of complete destructive interference, where two waves of the same frequency traveling in opposite directions perfectly cancel each other out.

An antinode is a point along a standing wave where the wave has maximum amplitude. This point undergos the most intense vibration, moving back and forth between large positive and negative displacements. Antinodes occur due to constructive interference, where the overlapping waves reinforce each other.

We have seen that resonant air columns can be analyzed by pressure and air displacement. A key rule is that a displacement node is always a pressure antinode (maximum pressure change), and vice versa. At a displacement node, air is squeezed and expanded, creating maximum pressure variation.

4. *Effects of the Ends*

Within a cavity, oppositely traveling waves are interdependent, originating from the reflection of an initial wave at both boundaries. These reflections create distinct boundary conditions at the ends, which are essential for understanding how standing waves are formed. To analyze this, we must separately examine systems with open (fixed) versus closed (free) boundaries.

When a cavity end is open, the reflected wave undergoes a phase reversal, meaning a compression returns as a rarefaction (and vice-versa), as shown in Fig. 29. Because the incident and reflected waves are equal in magnitude but opposite in sign, they cancel each other

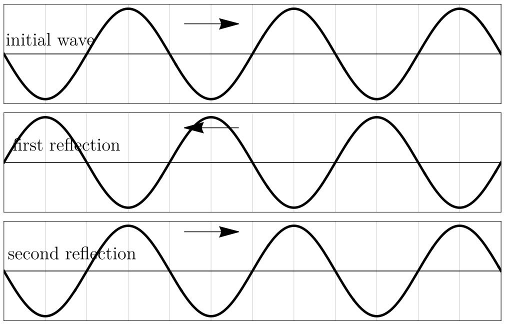

FIG. 31: Description of the alignment of waves upon reflections. Aligned wave reflection.

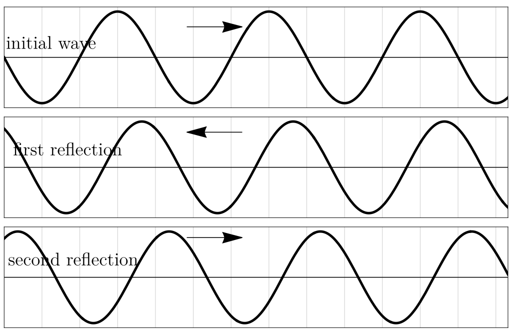

FIG. 32: Description of the alignment of waves upon reflections. Not aligned wave reflection.

out. The green dots in the figure illustrate that their sum is always zero, creating a *node* at the open end.

When a cavity end is closed, reflected waves do not undergo a phase reversal. This means an incoming compression reflects as a compression, and a rarefaction reflects as a rarefaction (as shown in Fig. 30). Unlike an open end, where waves cancel out, the incoming and reflected waves at a closed end reinforce each other. Because they have the same value, they create a maximum oscillation between compression and rarefaction, producing a pressure *antinode*.

For waves bouncing inside a cavity to produce constructive interference, they must have a wavelength that ensures consecutive reflections align properly.

In Fig. 31 we show an initial wave and its second reflection, both moving right, perfectly aligned. This

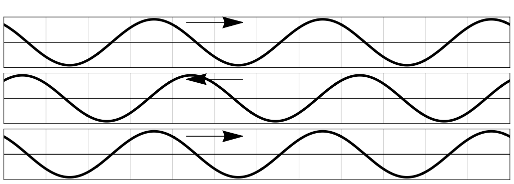

FIG. 33: Wave and its reflection for a tube with both closed ends.

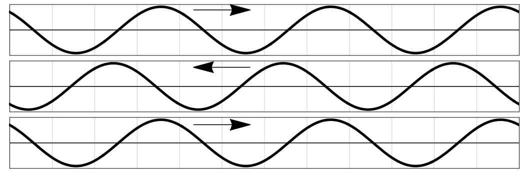

FIG. 34: Wave and its reflection for a tube with both open ends.

leads to fully constructive interference, where their amplitudes combine. In Fig. 32 we show an example where the initial wave and its second reflection are misaligned, causing them to cancel each other out over multiple reflections.

Let's explore the various options for pressure waves bouncing inside a cavity.

- *Equal ends*. For a cavity with fixed boundaries at both ends, waves reflect without inverting phase (compressions remain compressions). For open ends, reflection causes phase reversal (compressions become rarefactions). As shown in Figs. 33 and 34, although the first reflection differs, the second reflection is aligned with the incident wave in both cases, leading to constructive interference. This resonance occurs only when the cavity length is an exact multiple of half-wavelengths, as demonstrated by the comparison in Fig. 35. These examples illustrate a general rule relating wave wavelength to the cavity length required for constructive interference, which will be explored further below.

- *Different ends.* When a cavity has one open and one closed end, only one of the two reflections undergoes a phase reversal. Both configurations of the open-closed ends (left-right and right-left) are illustrated in Figs. 36 and 37. While the first reflection differs depending on which end is closed, the second reflection is identical in both scenarios, allowing them to be analyzed together.

  For the initial wave and its second reflection to align and create constructive interference, the cavity length must equal an integer number of full wavelengths plus an additional quarter or three-quarters of a wavelength. As shown in Fig. 38, this alignment occurs at 2.25 and 2.75 wavelengths, but fails at 2.5. This specific behavior follows a broader rule that we will explore in the next section.

5. Visualizing Standing Waves: Graphical Analysis and Governing Equations

Using the results from the previous subsection, we can now determine the conditions required for the formation of standing waves. By applying a graphical method, we

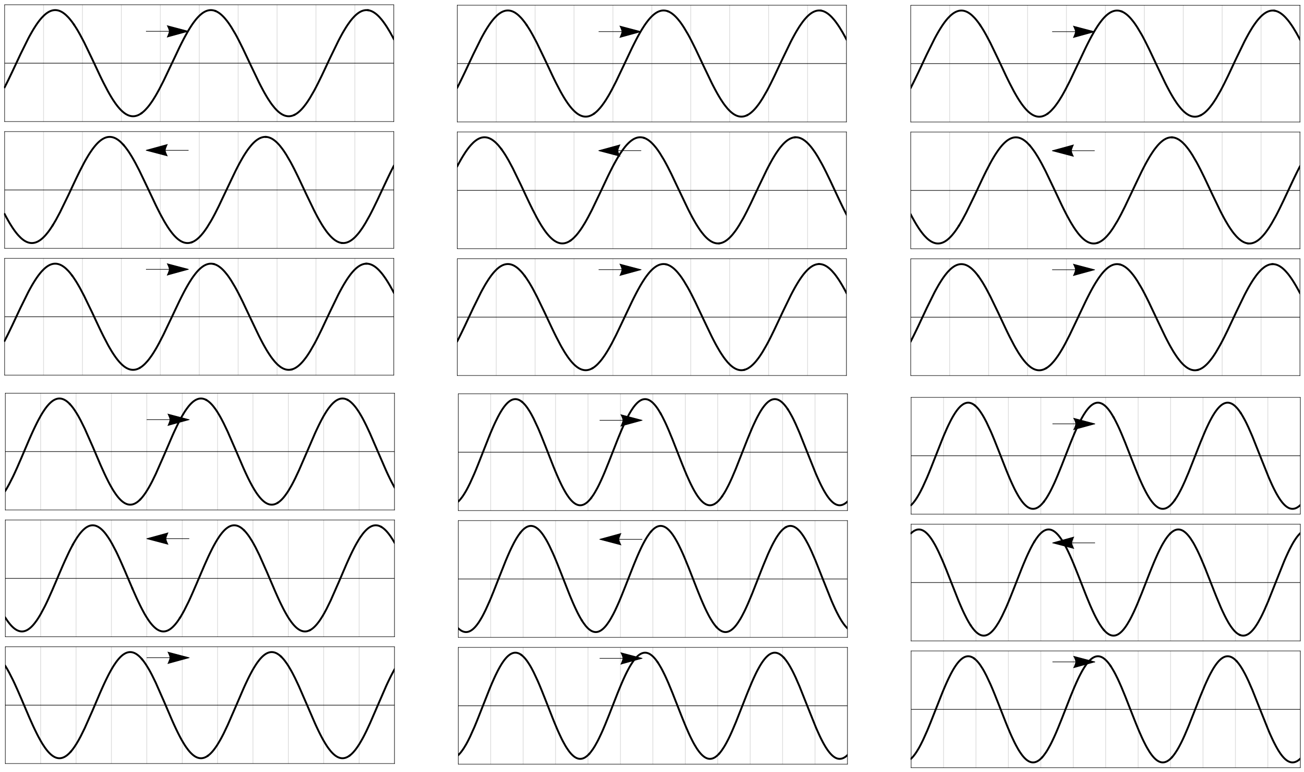

FIG. 35: Comparison of multiple cases of reflection in a cavity with equal ends. Each of the (left, middle, right) panels contains three rows. From top tp bottom and left to right: closed-closed ends, 2.5 wavelengths; open-open ends, 2.5 wavelengths; closed-closed ends, 2.75 wavelengths; open-open ends, 2.75 wavelengths; closed-closed ends, 3 wavelengths; open-open ends, 3 wavelengths.

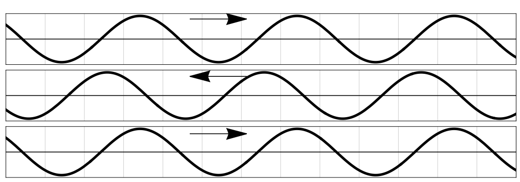

FIG. 36: Wave and its reflection for a tube with open-closed ends.

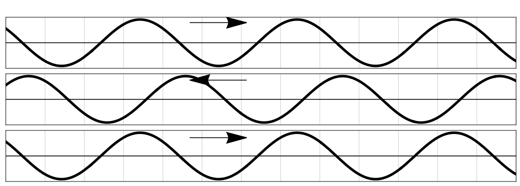

FIG. 37: Wave and its reflection for a tube with closed-open ends.

can sketch waves that satisfy two key boundary conditions: pressure nodes form at open ends, while antinodes form at closed ends.[10] We will examine four cavity configurations: open-open, closed-closed, open-closed, and closed-open (noting that the latter two are functionally identical, as one can be rotated into the other). Diagrams are provided to illustrate these, with Fig. 39 showing open-open cavities (nodes at both ends), Fig. 40 showing closed-closed cavities (antinodes at both ends), and Fig. 41 illustrating open-closed cavities.

Based on the Figs. 39, 40, and 41 a pattern emerges that allows for the classification of standing waves within different types of cavities.

Figures 39 and 40 demonstrate that open-open and closed-closed cavities share similar wave patterns, specifically in the number of waves they can support. In both scenarios, the cavity holds an integer or half-integer number of wavelengths: one-half, one, one-and-a-half, and so on $(0.5, 1, 1.5, 2, 2.5, \cdots$ full waves). Be-

---

[10] Recall that an open cavity end acts as an antinode for displacement, where air particles can move freely, resulting in maximum amplitude. In pressure terms, this is a point of minimum amplitude (a pressure node). It is a *free* end, allowing the standing wave to have maximum displacement at that point. A closed cavity constitutes a node for air displacement standing waves, as air molecules cannot move back and forth at the boundary. Consequently, it serves as a pressure antinode (maximum pressure variation).

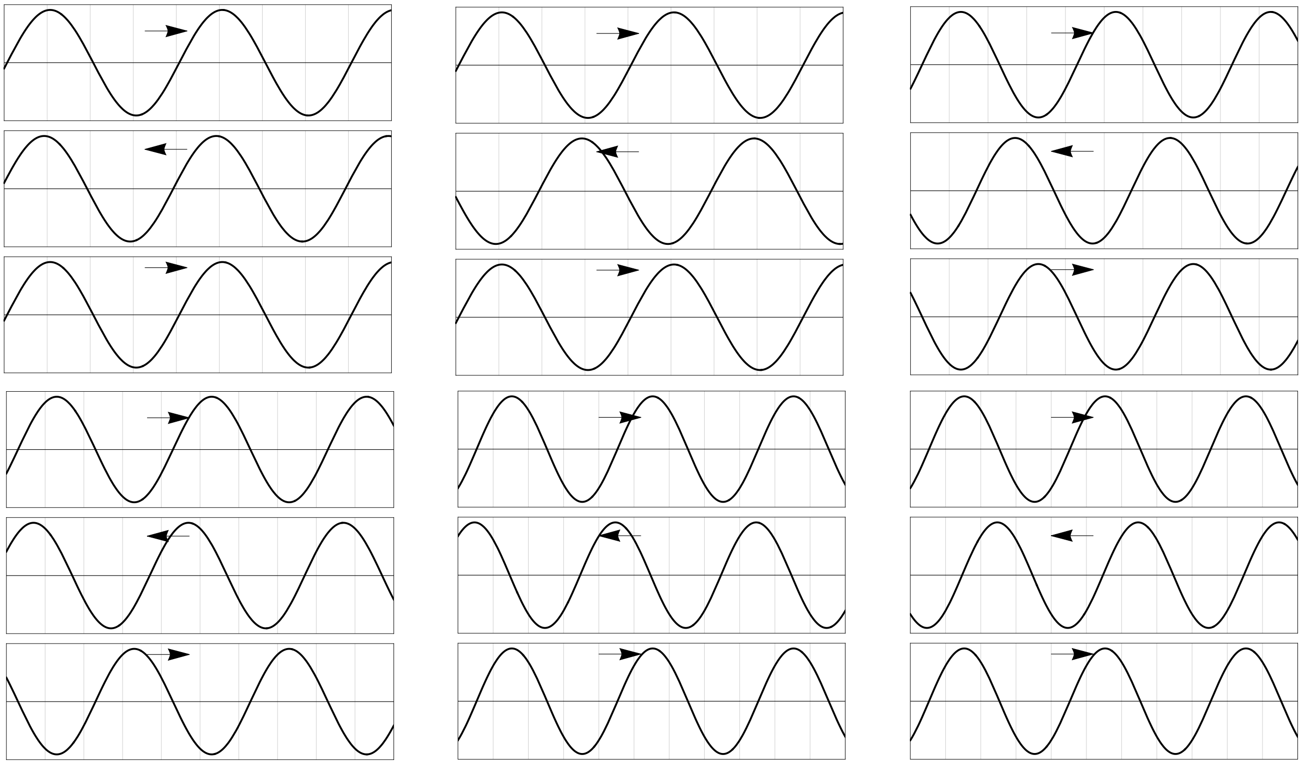

FIG. 38: Comparison of multiple cases of reflection in a cavity with unequal ends. Each of the (left, middle, right) panels contains three rows. From top tp bottom and left to right: open-closed ends, 2.25 wavelengths; closed-open ends, 2.25 wavelengths; open-closed ends, 2.50 wavelengths; closed-open ends, 2.50 wavelengths; open-closed ends, 2.75 wavelengths; closed-open ends, 2.75 wavelengths.

cause these cavities have matching boundary conditions at both ends, the relationship between the cavity length (L) and the wavelength (λ) is consistent for both types:

$$\frac{L}{\lambda} = 0.5, 1, 1.5, 2, 2.5, \cdots, \tag{51}$$

which can be expressed in general as

$$L = \frac{n}{2}\,\lambda\,, \tag{52}$$

with $n = 1, 2, 3, \cdots$.

Figure 41 demonstrates that with mismatched end conditions, the cavity supports resonances at odd multiples of a quarter-wavelength $(0.25, 0.75, 1.25, 1.75, 2.25\cdots)$ full wave periods. The resonant wavelength (λ) is therefore related to the cavity length (L) by:

$$\frac{L}{\lambda} = 0.25, 0.75, 1.25, 1.75, 2.25\cdots, \tag{53}$$

which can be expressed in general as

$$L = \frac{n}{4}\,\lambda\,, \tag{54}$$

with $n = 1, 3, 5, \cdots$.

A cavity can support various standing waves, known as harmonics, depending on the integer n used in the resonance formulas. The simplest of these ($n = 1$) is the fundamental harmonic. For a cavity with identical ends, this fundamental mode determines the longest possible resonant wavelength

$$\frac{L}{\lambda_1} = \frac{1}{2} \Rightarrow \lambda_1 = 2L \tag{55}$$

which is twice the length of the cavity. In the case of a cavity with unequal ends, the fundamental harmonic has a wavelength given by

$$\frac{L}{\lambda_1} = \frac{1}{4} \Rightarrow \lambda_1 = 4L\,. \tag{56}$$

For pressure waves, it is more practical to express the cavity length in terms of frequency rather than wavelength. Substitting the wave speed equation $u = \lambda f$ into (52) and (54) we obtain

$$f_n = \frac{nu}{2L}, \quad n = 1, 2, 3, \cdots \tag{57}$$

and

$$f_n = \frac{nu}{4L}, \quad n = 1, 3, 5, \cdots, \tag{58}$$

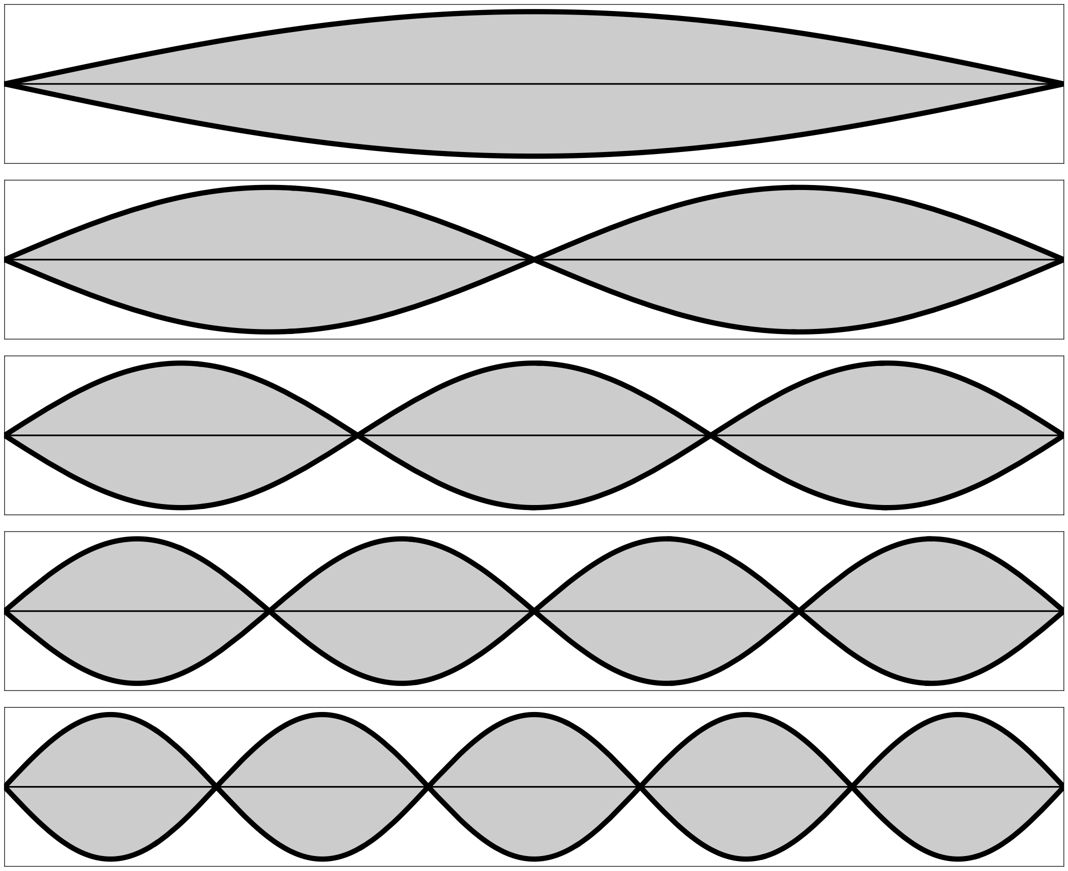

FIG. 39: Acoustic pressure modes in an open-open cavity, for n = 1, 2, 3, 4, 5 (top to bottom).

respectively. It is straightforward to see that for equal ends the fundamental frequency is

$$f_1 = \frac{u}{2L}\,, \tag{59}$$

whereas for unequal ends we have

$$f_1 = \frac{u}{4L}\,. \tag{60}$$

Ultimately, we can see that all frequencies (f_n) are simply integer multiples of the fundamental frequency f_1 using

$$f_n = n f_1\,. \tag{61}$$

The relation (61) holds for both boundary conditions, where equal ends allow for all harmonics ($n = 1, 2, 3, \cdots$) and unequal ends allow only for odd harmonics ($n = 1, 3, 5, \cdots$).

A cavity with identical ends produces harmonics at every integer multiple of the fundamental frequency (e.g., 50, 100, 150 Hz). Conversely, a cavity with dissimilar ends generates only odd-numbered multiples (e.g., 50, 150, 250 Hz).

E. Craftting Sound Anatonmy

Sound is generated when an object vibrates, forcing surrounding particles to move and create pressure waves. It travels fastest through solids, then liquids, and slowest through gases. In this section, we examine the fundamentals of sound waves.

1. *Sound Energy*

Sound energy is essentially mechanical energy created by vibrating objects. These vibrations travel through air, water, or solids as waves that we can hear. Defined by their pitch, volume, and length, these waves consist of both moving particles and the pressure they create. Beyond just being what we hear, this energy is a fundamental tool for communication, music, and technology.

Sound energy is classified by frequency into three distinct ranges: infrasound (under 20 Hz), audible sound (20 Hz to 20 kHz), and ultrasound (over 20 kHz). Each category serves specific roles, with audible sound enabling human communication, infrasound facilitating geological and environmental studies, and ultrasound providing vital tools for medical and industrial imag-

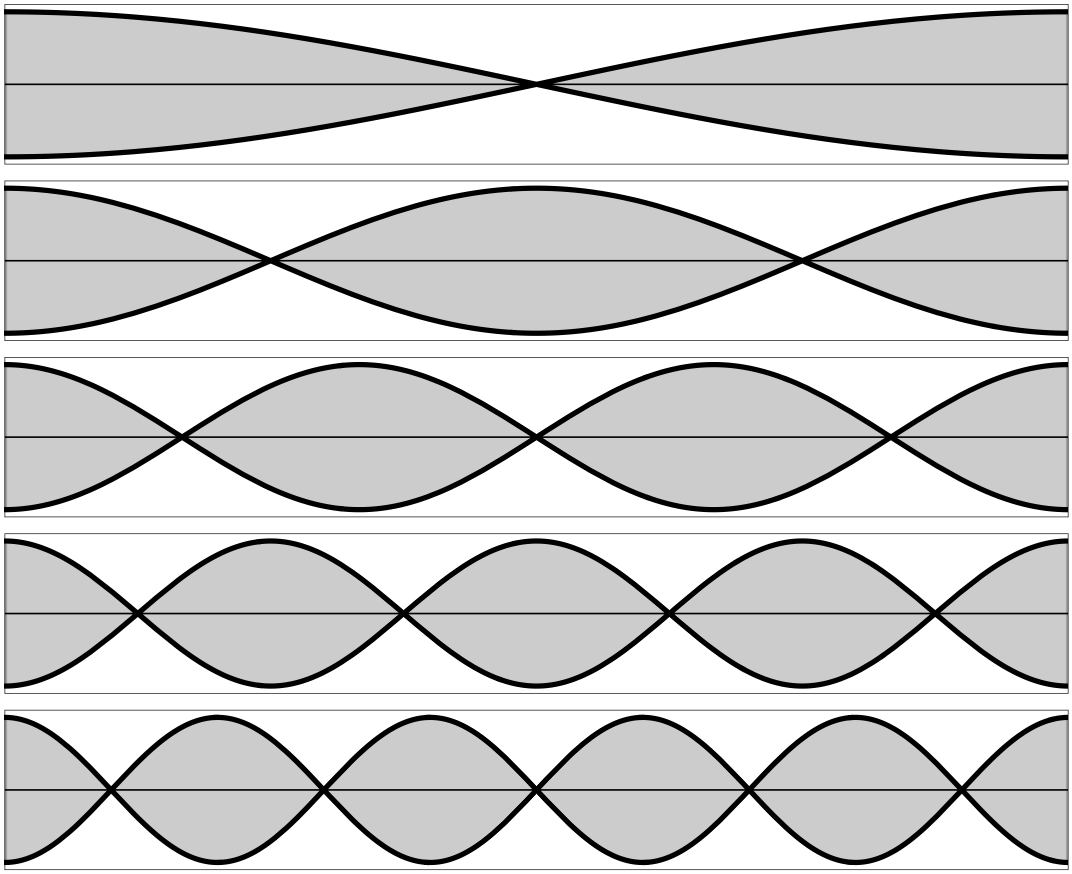

FIG. 40: Acoustic pressure modes in a closed-closed cavity for n = 1, 2, 3, 4, 5 (top to bottom).

ing.

Our ears turn sound waves into electrical signals for the brain. Sound vibrates the eardrum and middle ear bones, sending energy to the inner ear. There, cells convert these vibrations into signals that travel along the auditory nerve, allowing us to hear the world around us.

2. *Overtone Series and Musical Instruments*

Complex sounds (like human voices, musical instruments, or rushing water) are actually blends of multiple sine waves. This specific mix of waves at any moment is called a spectrum, and each individual wave within it is a partial.

The lowest frequency is the fundamental, while the higher waves are labeled overtones and numbered in order. For instance, in a sound made of 100 Hz, 260 Hz, and 400 Hz waves, the fundamental is 100 Hz and the overtones are 260 Hz and 400 Hz. More concretely, 260 Hz is called the first overtone while 400 Hz is called the second overtone.

In music theory, the set of overtones is a key component that defines a sound's timbre, which is why different instruments sound unique even when playing the same pitch. Importantly, this harmonic structure is not fixed; individual partials can change over time. In fact, some overtones may fade away entirely while new ones emerge as the sound evolves.

Overtones are categorized as harmonic if they are integer multiples of the fundamental frequency and inharmonic if they are not. While natural environmental sounds like rain or thunder are typically inharmonic, musical instruments – especially those utilizing vibrating strings (e.g., guitar, violin) or air columns (e.g., trumpet, saxophone, flute) – produce harmonic overtones because the physical length of the instrument constrains the waveforms to whole-number ratios

3. *Speed of Sound in Air*

Sound waves propagate through space at a finite speed determined by intermolecular collisions within a gas, representing an intrinsic property of the gas rather than a factor of the source's velocity. As molecular interaction times limit propagation speed, pressure waves produced by moving sources (such as a speaker on a vehicle) travel at the same speed through the medium as waves

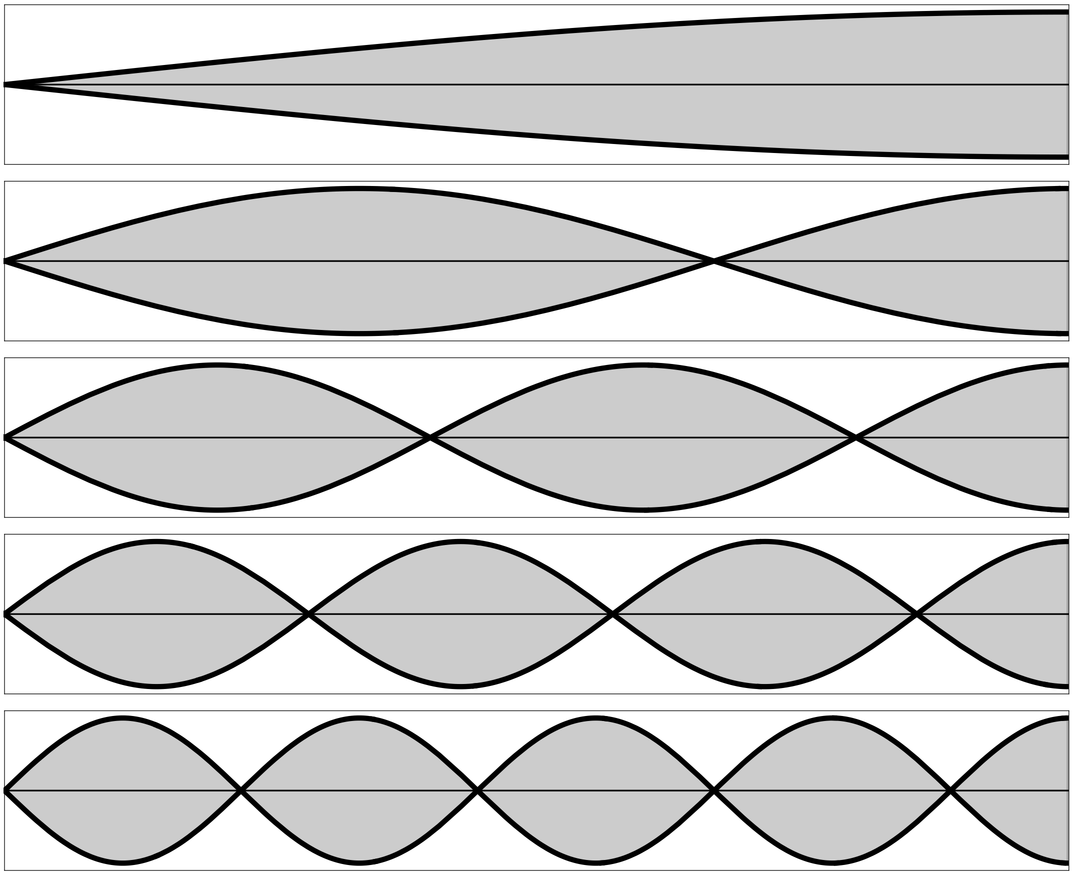

FIG. 41: Acoustic pressure modes in an open-closed cavity for n = 1, 3, 5, 7, 9 (top to bottom). The open end is on the left and the closed end is on the right.

produced by stationary sources.

We have seen that the speed of pressure waves in a gas depends on the frequency of molecular collisions. When molecules collide more often, compression pulses can more effectively shift them into adjacent areas. Since collision rates are driven by molecular velocity, which is itself a function of temperature, a direct link exists between a gas's temperature and its ability to propagate pressure waves. Specifically, the speed of pressure waves in an ideal gas, expressed in terms of the Boltzmann constant k_B is given by

$$\text{speed of pressure waves} \equiv u = \sqrt{\frac{\gamma\, k_B T}{m}} \qquad (62)$$

where T is the absolute temperature (measured in K), m is the mass of one molecule (measured in kg), and γ is the *adibatic gas index* a dimensionless parameter (always > 1; e.g., 1.4 for air) that characterizes the behavior of an ideal gas during an adiabatic process, where no heat is transferred. [11]

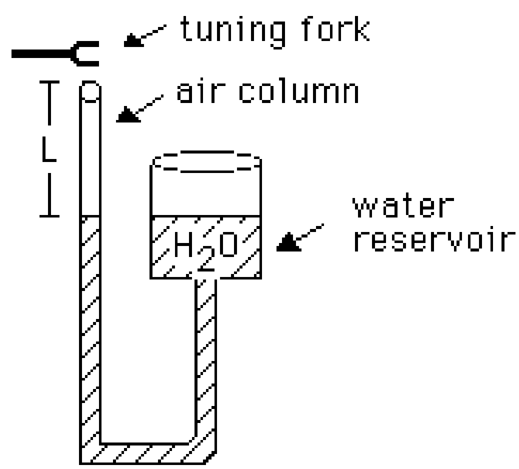

FIG. 42: Experimental set up to measure lengths of air column where resonances occur. Two connected pipes (one fixed and one movable). Moving the reservoir container up and down changes the water level inside the measuring pipe, allowing precise adjustment of the air column.

[11] A point worth notng at this juncture is that transverse waves on

Sound waves travel faster in solids than in liquids or gases primarily because the particles in a solid are tightly packed together and strongly bonded, allowing vibrational energy to transfer rapidly.

Measuring the speed of sound in air using harmonics is typically done with a resonance tube (an open-closed cavity) and a tuning fork of known frequency f, see Fig. 42. By finding the lengths of air column where resonance occurs we can calculate the wavelength and subsequently the speed of sound. The first harmonic occurs when the column air is 1/4 of the wavelength of the sound,

$$L_1 = \frac{1}{4}\lambda \tag{63}$$

The next point of maximum loudness (first overtone) occurs when the air column is 3/4 of the wave of sound,

$$L_2 = \frac{3}{4}\lambda\,. \tag{64}$$

This implies that

$$\Delta L = L_2 - L_1 = \lambda/2 \tag{65}$$

and so the speed of sound in air is

$$u = \lambda f = 2\Delta L\, f\,. \tag{66}$$

Using two harmonics eliminates the need to calculate the bordoer effects at the top of the tube.[12] For example, a sound wave of 512 Hz propagating inside the resonant tube at 284.31 K (which corresponds to about 11°C and 52°F) leads to $L_1 \simeq 16$ cm and $L_2 \simeq 49$ cm, yielding $u \simeq 338$ m/s [22]. The speed of sound in air, as a function of temperature, is measured to be

$$u \approx 331.3\sqrt{\frac{T}{273.15\ \mathrm{K}}}\ \mathrm{m/s}\,. \tag{67}$$

Measuring the speed of sound with toilet rolls is an alternative hands-on experiment that uses the principle of destructive interference to determine the sound's wavelength. In this setup, you typically use a pair of wired earbuds (two coherent sources) inside a tube (made of toilet rolls), looking for the *quiet spots* where the sound almost disappears. When the two earbuds are separated by exactly half a wavelength ($\lambda/2$), the peak of the wave from one earbud meets the trough of the wave from the other. They cancel each other out, creating a *node* of silence. By measuring the distance between two consecutive quiet spots we can find the wavelength, and then use it to calculate the speed u.[13]

In sound speed experiments, the difference between *destructive* and *constructive* interference depends on whether you are measuring phase cancellation between two sources (the toilet roll method) or resonance in a standing wave (the water tube method).

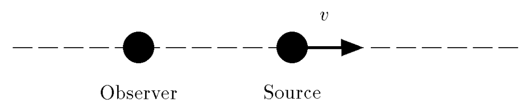

FIG. 43: A source moves away from a stationary observer.

F. Doppler Effect

The Doppler effect is the change in frequency of a wave in relation to an observer moving relative to the source, causing higher pitch/frequency as it approaches and lower as it recedes [21]. The effect only occurs when the wave source and observer are moving toward or away from each other. A classic example is a police car siren that sounds higher-pitched as it approaches and lower-pitched after it passes. For the scope of this course, the source velocity is considered insignificant relative to the speed of sound.[14]

1. *Source in Motion and Observer at Rest*

First, assume that the source is moving and the observer is standing still (relative to the air), with all motion taking place along a line, see Fig. 43. Hereafter, u stands for the velocity of sound waves, v for the recession velocity of the source, $\tau_S \equiv \Delta t_S$ for the period of the wave at the source, and $\tau_O = \Delta t_O$ for the period of the wave

a string are caused by a perpendicular vibration and depend on different properties of the string itself, rather than the bulk material properties of the medium. Indeed the speed of a transverse wave on a string is $u = \sqrt{T/\mu}$ where T is the string tension (i.e. force applied to stretch it) and μ the linear mass density (or mass per unit length) of the string.

[12] The standing wave does not end exactly at the top of the tube, but slightly above it. The end correction is approximately 0.3 times inside diameter of the tube.

[13] In greater detail, place one toilet roll on a flat surface. Put one earbud at the very bottom, facing up. Stack the other two rolls on top to create a long tunnel. Attach some Blu-Tack to the cable near the second earbud to keep it hanging straight. Play a constant 3 kHz tone through both earbuds using a smartphone. To find the quietest points, lower the second earbud into the tube. Slowly raise it until the sound becomes very quiet (this is a point of destructive interference). Mark the cable at the top of the tube with a pen or Blu-Tack. Continue raising it until you find the next quietest point and mark that on the cable as well. Pull the cable out and measure the distance between your two marks. This distance is the wavelength λ [23].

[14] This section's explanation of the Doppler effect draws from Guth's superb lecture notes [24].

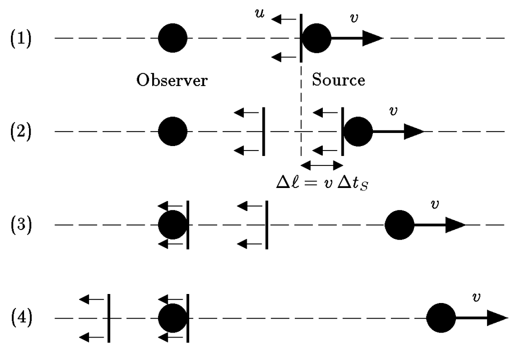

FIG. 44: Successive wavefronts emitted by a source moving with speed v to the right, while the observer is at rest.

as observed. Now consider the sequence illustrated in Fig. 44:

1. The source emits a wave crest.
2. At a time Δt_S later, the source emits a second wave crest. During this time interval the source has moved a distance $\Delta\ell = v\Delta t_S$ further away from the observer.
3. The stationary observer receives the first wave crest.
4. At some time Δt_O after (3), the observer receives the second wave crest. Our goal is to find Δt_O.

The time at which the first wave crest is received depends of course on the distance between the source and the observer, which was not specified in the description above.

We are interested, however, only in the time difference Δt_O between the reception of the first and second wave crests. This time difference does not depend on the distance between the source and the observer, since both wave crests have to travel this distance. The second crest, however, has to travel an extra distance

$$\Delta\ell = v\,\Delta t_S\,, \tag{68}$$

since the source moves this distance between the emission of the two crests. The extra time that it takes the second crest to travel this distance is $\Delta\ell/u$, so the time between the reception of the two crests is

$$\begin{aligned}\Delta t_O &= \Delta t_S + \frac{\Delta\ell}{u}\\ &= \Delta t_S + \frac{v\,\Delta t_S}{u}\\ &= \left(1+\frac{v}{u}\right)\Delta t_S\,.\end{aligned} \tag{69}$$

Equation (69) can be recast as

$$f_O = f_S\,\frac{1}{1+v/u} \tag{70}$$

where $f_O = 1/\tau_O = 1/\Delta t_O$ and $f_S = 1/\tau_S = 1/\Delta t_S$. Because the denominator in in the left-hand-side of (70) is greater than 1, the pitch heard f_O is lower than f_S.

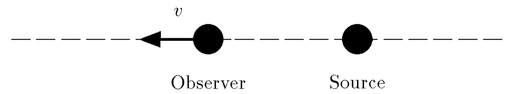

FIG. 45: The receiver is moving with respect to the emitter.

Now, if a source is moving toward a stationary observer, the direction of v changes. Duplicating the steps above for a source moving to the left, it is easily seen that the shift in the frequency is given by

$$f_O = f_S\,\frac{1}{1-v/u}\,. \tag{71}$$

Because the denominator in the left-hand-side of (71) is less than 1, the observed frequency f_O is greater than the source frequency.

In large-scale performances like marching bands, the physical movement of musicians can cause slight pitch shifts (the "out-of-tune" effect) perceived by the audience. In elite performances aiming for higher precision, the rapid movement causing Doppler shifts is accompanied by high-volume, complex sounds, making the subtle pitch variations (typically only fractions of a hertz) inaudible or negligible to the audience.

2. *Observer in Motion and Source at Rest*

Suppose now that the source stands still, but the observer is receding at a speed v, as shown in Fig. 45.

In this case, the sequence (shown in Fig. 46) becomes:

1′. The source emits a wave crest.

2′. At a time Δt_S later, the source emits a second wave crest. The source is standing still.

3′. The moving observer receives the first wave crest.

4′. At a time Δt_O after (3′), the observer receives the second wave crest. During the time interval between (3′) and (4′), the observer has moved a distance $\Delta\ell = v\Delta t_O$ further from the source.

Using the same strategy as in the first case, we note that in this case, the second wave crest must travel an extra distance $\Delta\ell = v\,\Delta t_O$. Thus,

$$\Delta t_O = \Delta t_S + \frac{\Delta\ell}{u} = \Delta t_S + \frac{v\Delta t_O}{u}\,. \tag{72}$$

In this case Δt_O appears on both sides of the equation, but after a bit of algebra we can easily solve for Δt_O to find

$$\Delta t_O = \left(1-\frac{v}{u}\right)^{-1}\,\Delta t_S\,, \tag{73}$$

or equivalently

$$f_O = \left(1-\frac{v}{u}\right)\,f_S\,, \tag{74}$$

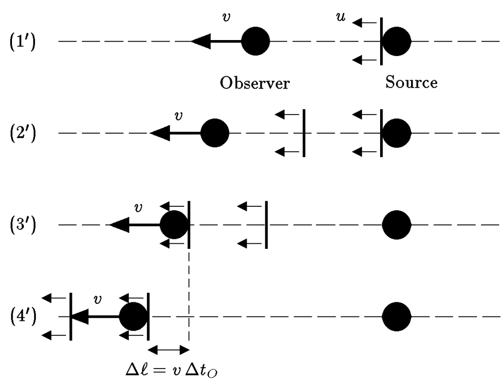

FIG. 46: Successive wavefronts emitted by a source at rest, while the observer is moving with speed v to the left.

with $f_O = 1/\tau_O = 1/\Delta t_O$ and $f_S = 1/\tau_S = 1/\Delta t_S$. Because $(1 - v/u) < 1$, the pitch heard f_O is lower than f_S.

Now, if an observer is moving toward a stationary source, the direction of v changes. Duplicating the steps above for an observer moving to the right, it is easily seen that the shift in the frequency is given by

$$f_O = \left(1 + \frac{v}{u}\right) f_S \,. \qquad (75)$$

Because $(1 + v/u) > 1$, the pitch heard f_O is higher than f_S.

III. ACOUSTIC LOGARITHMIC SCALES

Acoustics utilizes logarithmic scales (decibels (dB) for intensity and octaves for frequency) to match the human ear's logarithmic perception of sound. Decibels measure amplitude, where a 10 dB increase feels twice as loud, while octaves group frequencies by doubling, splitting sound into manageable, musically relevant, and often logarithmic octave bands. In this section, we examine the application of logarithmic scales in acoustics.

A. Logarithm Rules

In this section, we briefly detour into the mathematical rules governing logarithms.

1. Understanding Logarithms Intuitively

While this section covers the logarithm rules, we will start with a review of exponents. Because logarithms and exponents are closely linked, understanding the latter is necessary for mastering the former.

In an expression like 2^4, the number 4 is the exponent (or power) whereas the number 2 is the base. For example, in $81 = 9^2$, the number 2 is the exponent and 9 is the base.

To motivate the study of logarithms, consider the relationship between exponents and multiplication. We know that $16 = 2^4$ and $8 = 2^3$. If we wish to multiply 16 and 8, the standard method is direct long-multiplication, which yields 128. However, this process becomes increasingly tedious as numbers grow larger. Alternatively, let's work out this calculation more efficiently by utilizing the properties of powers. First, express each number as a power of a common base. In this case, both 16 and 8 are powers of 2:

$$16 = 2^4 \quad \text{and} \quad 8 = 2^3 \qquad (76)$$

Instead of multiplying the decimal values, multiply the exponential expressions. According to the product rule for exponents, $a^m \cdot a^n = a^{m+n}$. By adding the exponents, we simplify the operation:

$$16 \cdot 8 = 2^4 \cdot 2^3 = 2^{4+3} = 2^7 \,. \qquad (77)$$

Finally, convert the resulting power back into a standard integer. Since $2^7 = 128$, we arrive at the same product without performing traditional long-multiplication.

In a similar manner, dividing 16 by 8 can be expressed as $2^4/2^3$. . Following the laws of indices, we subtract the exponents $(4 - 3)$ to find the resulting power of 1. This gives 2^1 , or simply 2. By utilizing a lookup table of the powers of 2, we could easily determine that $2^7 = 128$, which allows us to find the product of $16 \cdot 8$ through simple addition of their exponents $(4 + 3 = 7)$. This fundamental insight inspired Napier and Briggs to invent logarithms, providing a method to simplify complex calculations by substituting multiplication with addition and division with subtraction [25, 26].

The exponential equation $16 = 2^4$ can be rewritten in an equivalent logarithmic form as $\log_2 16 = 4$, which is read as: log base 2 of 16 equals 4. In this conversion, the base of the exponent (2) becomes the base of the logarithm, while the exponent (4) equals the result of the logarithm. Because these forms are equivalent, they imply one another:

$$16 = 2^4 \Leftrightarrow \log_2 16 = 4 \,. \qquad (78)$$

2. Understanding Logarithms: Definition and Essential Laws

The two sets of statements given in (78), one using exponents and the other logarithms, are equivalent. Generally, we *define the logarithm as*:

$$\text{If } x = a^n \quad \text{then equivalently} \quad \log_a x = n \,. \qquad (79)$$

Note that for any base a we have $a = a^1$ and so $\log_a a = 1$.

The *first law of logarithms* states that the logarithm of a product is equivalent to the sum of the logarithms of the individual factors. To understand this law assume

$$x = a^n \quad \text{and} \quad y = a^m \,, \qquad (80)$$

that rearranges to the following logarithmic forms

$$\log_a x = n \quad \text{and} \quad \log_a y = m\,. \tag{81}$$

Using the rule of exponents

$$xy = a^n \cdot a^m = a^{n+m}\,. \tag{82}$$

Converting the expression $xy = a^{n+m}$ into its logarithmic equivalent yields $log_a(xy) = n + m$. By substituting the original definitions $n = log_a x$ and $m = log_a y$, we arrive at the identity

$$\log_a(xy) = \log_a x + \log_a y\,. \tag{83}$$

Therefore, multiplying two numbers and taking the log equals adding the logs of those numbers.

The *second law of logarithms* states that the logarithm of a number raised to an exponent is evaluated by multiplying the exponent by the logarithm of the base number. Assume $x = a^n$, or equivalently $\log_a x = n$. Now, raise both sides of $x = a^n$ to the power m,

$$x^m = (a^n)^m\,. \tag{84}$$

Using the rules of exponents we can write (84) as

$$x^m = a^{nm}\,. \tag{85}$$

Taking the quantity x^m as a single term, the logarithmic form is

$$\log_a x^m = nm = m \log_a x\,. \tag{86}$$

The *third law of logarithms* states that the logarithm of a quotient is equal to the difference between the logarithms of the numerator and the denominator. Assuming again that the relations (80) and (81) hold, we can write using the rules of exponents

$$\frac{x}{y} = \frac{a^n}{a^m} = a^{n-m} \tag{87}$$

or equivalently, in logarithmic form

$$\log_a \frac{x}{y} = n - m\,, \tag{88}$$

which from (81) can be rewritten as

$$\log_a \frac{x}{y} = \log_a x - \log_a y\,. \tag{89}$$

Recall that any non-zero base raised to the power of zero equals, i.e. $a^0 = 1$. In logarithmic form, this identity is expressed as:

$$\log_a 1 = 0\,. \tag{90}$$

The logarithm of 1 in any base is 0.[15]

[15] The natural logarithm (lnx) is a logarithm with the base e, where is an irrational number approximately equal to 2.71828.

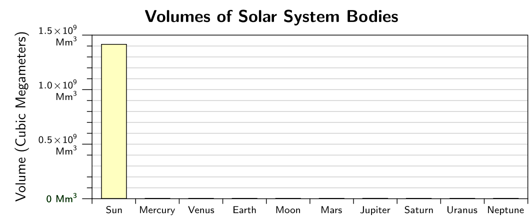

FIG. 47: Volumes of bodies in the solar system in linear scale.

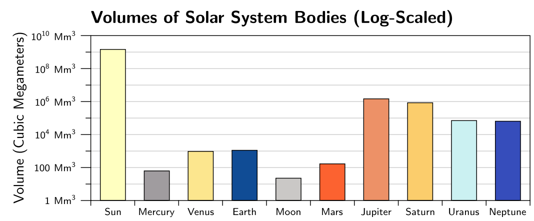

FIG. 48: Volumes of bodies in the solar system in $\log_{10}$ scale.

3. *Tales from Interplanetary Space*

Logarithms are incredibly useful for comparing, at a glance, items that vary greatly in size. For example, if we wanted to graph the volumes of major solar system bodies using cubic megameters (Mm^3), a useful unit roughly equivalent to the distance from Madrid to Paris, we would face massive scaling issues. Logarithms allow us to visualize these vast differences clearly.

The traditional bar graph in Fig. 47 illustrates a common data visualization issue: when one value (the Sun) is vastly larger than the others, it becomes impossible to distinguish the sizes of the remaining objects. We cannot easily compare e.g., Saturn to Neptune; the graph only highlights that the Sun is huge, which is not particularly informative. By using a logarithmic scale, where each vertical notch represents a 10-fold increase rather than a linear step, we gain much better insight. This method allows us to visualize data spanning several orders of magnitude. In Fig. 48 we show the same data as in Fig. 47 but the vertical axis is now in $\log_{10}$ scale. Although the Sun remains the largest, we can now compare the planets effectively. Jupiter and Saturn are roughly a thousand times smaller than the Sun. Uranus and Neptune are ten times smaller than Jupiter/Saturn. Earth is approximately a million times smaller than the Sun, a distinction hidden in Fig. 47.

Log-scaled graphs are powerful tools, but they can be deceptive. A casual glance might suggest Earth is one-sixth the size of the Sun because its bar appears to be one-sixth the height, but that is incorrect; it is actually six whole orders of magnitude smaller. Logarithmic scales should be used when the data spans massive ranges, providing clarity that linear scales cannot.

B. Decibel Scale

The human ear can detect sounds ranging from a pin drop to a roaring jet, a trillion-fold difference in power. The decibel scale compresses this enormous range into a manageable set of numbers (typically 0 to 140) that matches our biological perception. In this section, we will break down the decibel scale.[16]

1. Loudness

As previously noted, the human ear possesses remarkable sensitivity, with the threshold of hearing for a young, healthy ear starting at a physical intensity of 1×10^{-12} W/m^2. While this physical measure is known as Sound Intensity (I), loudness is typically expressed in Sound Intensity Level (L), measured in decibels (dB), primarily because human perception of sound is not linear.

Our ears do not perceive a doubling of physical intensity as *twice as loud*. Generally, the sound intensity must increase about tenfold before it is perceived as twice as loud (a rule of thumb that varies by frequency). If physical sound intensity (W/m^2) were used, the scale would need to span over 14 orders of magnitude (from 1 to $100{,}000{,}000{,}000{,}000$) to cover the range from the faintest to loudest sounds.

The decibel scale is preferred because it is logarithmic and compact (0 dB to 140 dB for the same range), closely aligning with our perception of loudness. Therefore, you can consider sound intensity as the objective physical measurement, and the decibel level as the subjective, psychological measure of loudness.

The equation that relates sound intensity to sound intensity level is given by

$$L = 10 \log_{10} \frac{I_2}{I_1}\,, \tag{91}$$

where L is the number of decibels I_2 is greater than I_1, I_2 is the higher sound intensity being compared, and I_1 is the lower sound intensity being compared. In other words, the decibel is a dimensionless unit (one-tenth of a bel) representing the base-10 logarithm of a ratio of intensity levels.[17]

Keep in mind that intensity (I) is measured in Watts/meter^2 (W/m^2), representing the raw power of sound. The L in the formula represents the decibel difference between two sounds. Typically, I_1 is set to the threshold of human hearing. Therefore, common decibel measurements (such as 20 dB for a whisper, 80 dB for traffic, or 105 dB for a loud cheer) are measured relative to this faint baseline (e.g., “80 dB” means 80 dB above the threshold). It is important to note that 0 dB does not indicate total silence, but rather the quietest sound a human with perfect hearing can detect. Check out Table II for common examples.

TABLE II: Decibel levels for typical sounds.

Source of sound	L (dB)	I (W/m^2)
Threshold of hearing	0	1×10^{-12}
Breathing	20	1×10^{-10}
Whispering	40	1×10^{-8}
Talking softly	60	1×10^{-6}
Lound conversation	80	1×10^{-4}
Yelling	100	1×10^{-2}
Loud concert	120	1
Jet takeoff	140	1×10^{2}

2. Inverse Square Law

Sound intensity is significantly affected by your distance from the source; a whisper audible at one meter becomes entirely lost across a football field. The reason for this is that sound energy radiates outward in all directions, forming an expanding sphere, with a surface $= 4\pi r^2$.

At one meter, that energy is concentrated on a small area of 4π m^2. By two meters, the same amount of power must cover 16π m^2, which is four times the area. Because the same energy is spread over a larger area, the intensity at two meters is only one-quarter of the original strength. Similarly, at three meters, the area covered is nine times larger (36π m^2), making the intensity only one-ninth of the original.

This phenomenon is known as the *inverse square law*. In simple terms: if you double the distance from the source, the sound becomes four times less intense; if you triple the distance, it is nine times less intense. Essentially, the intensity changes by the square of the reciprocal of the distance change.

For example, if the sound intensity of a screaming baby were 10^{-2} W/m^2 at 2.5 m away and we want to determine what would it be at 6 m away, the reasoning is as follows:

- the distance from the source of sound is greater by a factor of $6.0/2.5 = 2.4$;
- the sound intensity is decreased by $1/2.4^2 = 0.174$;
- the new sound intensity is 10^{-2} W/m$^2 \cdot 0.174 = 1.74 \times 10^{-3}$ w/m^2.

---

[16] This section aligns closely with the insightful discussion presented by Lapp [27].

[17] The decibel scale was developed by engineers at Bell Telephone Laboratories in the 1920s to measure power loss in telecommunication systems. It is named after Alexander Graham Bell, with the official, widely adopted definition appearing in 1928, replacing earlier transmission units.

TABLE III: Sound intensity levels (measured at 10 m away for various instruments.

Orchestral Instrument	L (dB)
Violin (at its quietest)	34.8
Clarinet	76.0
Trumpet	83.9
Cymbals	98.8
Bass drum (at its loudest)	103

Now, to calculate the total power output of a sound source, we multiply its sound intensity (in Watts per square meter) by the surface area of the sphere to which the sound has spread. For the baby produces a sound intensity of 10^{-2} W/m^2 at a distance of 2.5 m, the total power output is calculated by multiplying this intensity by the area of a sphere with a 2.5 m radius, i.e., $P = 0.785$ W. The power output calculated 6 m away it is obviously the same, because the power output depends on the baby, not the position of the observer. This means we can always equate the power outputs that are measured at different locations:

$$P_1 = P_2 \Rightarrow I_1 r_1^2 = I_2 r_2^2 \,. \tag{92}$$

It is easily seen that as the radius increases, the intensity decreases in inverse proportion to its square.

3. *Orchestral Sound*

A live orchestra fills a concert hall with rich sound, created by the unique vibrations of each individual instrument. While the audience hears immense volume, the actual acoustic power is generated by the physical effort of the musicians—a finger pressing a piano key, air passing over a clarinet reed, or cymbals crashing together. Surprisingly, only about 1% of this input energy is converted into the actual sound wave emitted by the instrument. However, because the human ear is incredibly sensitive, it requires very little energy to perceive sound. For example, a full 75-piece orchestra actually produces a total output of approximately $P_{\text{orchestra}} \approx 67$ W.

To find the sound intensity level at $r = 10$ m, first calculate the sound intensity at that distance

$$I_{\text{orchestra}} = \frac{P_{\text{orchestra}}}{4\pi r^2} \approx 0.056 \text{ W/m}^2 \,. \tag{93}$$

The sound intensity level would then be

$$L = 10 \log_{10}\left(\frac{I_{\text{orchestra}}}{I_0}\right) \approx 107.5 \text{ dB} \,. \tag{94}$$

Table III provides the decibel levels for various orchestral instruments measured from a distance of 10 m, which can be used to solve the problems in Sec. IV.

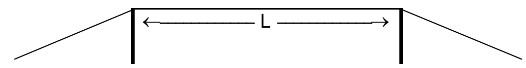

FIG. 49: Single string guitar of size L.

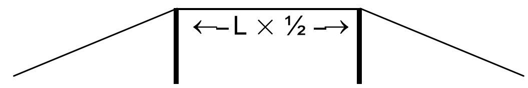

FIG. 50: Single string guitar of size $L/2$.

C. Fom Wave to Note

The octave serves as the fundamental building block of music, representing a unique, almost universal perceptual boundary where a note and its higher/lower counterpart sound functionally identical, yet exist at different pitch registers. To wrap up our exploration of sound and music, this section focuses on the octave and its constituent intervals.

1. *The Harmony of Numbers: Pythagoras*

Pythagoras was the first to scientifically define a major scale by discovering that harmonious musical intervals are rooted in simple mathematical ratios

$$1:1 \quad 2:1 \quad 3:1 \,. \tag{95}$$

In what follows, we concentrate on the 1:1 and 2:1 ratio.

To ideally reconstruct Pythagoras's experiment, we visualize a single-string guitar (monochord) by stretching a string between two bridges separated by a distance L, as shown in Fig. 49. Plucking the string now produces a specific pitch.

Next, we build an identical instrument, but with the bridge distance reduced to $L/2$ as shown in Fig. 50. The key to achieving harmonic resonance lies in maintaining uniform tension across strings of different lengths. According to modern physical principles, a string half the length of another will vibrate at twice the frequency, creating a 2:1 ratio. Because this relationship is fundamental to pitch recognition, both tones are perceived as the same

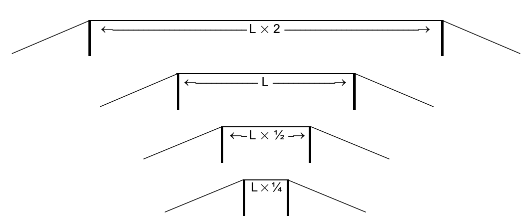

FIG. 51: Pythagorean harmonic ratios.

musical note (e.g., 440 Hz and 880 Hz). This principle of equivalence is what enables collective singing, allowing individuals with different vocal registers to sing the same melody in unison.

This procedure is reversible and repeatable, allowing for the construction of the sequence of single-string guitars shown in Fig. 51.

Given the right conditions, these strings generate this ratio of frequencies

$$1:2 \quad 1:1 \quad 2:1 \quad 4:1\,. \tag{96}$$

These ratios follow a natural, ascending order, establishing a natural ordinal scale. This scale is part of the following infinite, ordered sequence

$$\cdots \boxed{\frac{1}{4}} < \boxed{\frac{1}{2}} < \boxed{\frac{1}{1}} < \boxed{\frac{2}{1}} < \boxed{\frac{4}{1}} < \boxed{\frac{8}{1}} \cdots. \tag{97}$$

By assigning the ratio 1/1 to a reference frequency of 220Hz, the following derivative scale is generated

$$\cdots \boxed{\substack{55\\ \text{Hz}}} < \boxed{\substack{110\\ \text{Hz}}} < \boxed{\substack{220\\ \text{Hz}}} < \boxed{\substack{440\\ \text{Hz}}} < \boxed{\substack{880\\ \text{Hz}}} < \boxed{\substack{1760\\ \text{Hz}}} \cdots. \tag{98}$$

Note that this scale is theoretically infinite, even though it rapidly extends beyond the range of human hearing in both directions.

2. *The Octave: Music's Fundamental Interval*

So far, our scale has an ordered ranking without defined start or end points. We can turn this into an interval scale by assuming the distance between each consecutive ratio or frequency is the same, i.e.,

$$\boxed{\frac{1}{4}} - \boxed{\frac{1}{2}} = \boxed{\frac{1}{2}} - \boxed{\frac{1}{1}} = \boxed{\frac{1}{1}} - \boxed{\frac{1}{2}} = 1 \text{ octave} \tag{99}$$

or

$$\boxed{\substack{55\\ \text{Hz}}} - \boxed{\substack{110\\ \text{Hz}}} = \boxed{\substack{110\\ \text{Hz}}} - \boxed{\substack{220\\ \text{Hz}}} = \boxed{\substack{220\\ \text{Hz}}} - \boxed{\substack{440\\ \text{Hz}}} = 1 \text{ octave}\,. \tag{100}$$

The resulting interval-unit is musically fundamental, and is called an *octave*. It is defined by the gap between a pitch and the next one sharing the same note name. In terms of frequency, this happens whenever a pitch's vibration rate doubles. For instance, the note A at 220 Hz and the next A at 440 Hz are exactly one octave apart; likewise, the jump from 440 Hz to 880 Hz forms another octave.

It's important to recognize that the specific numerical difference between frequencies doesn't actually matter in music; what counts is the ratio between them. We could think of a musical scale as a logarithmic scale. For instance, if we convert those frequency ratios into base-2 logarithms, we get a scale that looks much more familiar

$$\cdots \boxed{-2} < \boxed{-1} < \boxed{0} < \boxed{+1} < \boxed{+2} \cdots; \tag{101}$$

specifically, values are spaced apart by simple subtraction; for instance, a difference of 1 represents a distance of exactly one octave.

3. *Making Sense of Cents in Musical Scales*

We have seen that an octave is the interval between two pitches where the higher frequency is double the lower (2:1 ratio). In Western music, an octave consists of 12 semitones (or half-steps), each 100 cents apart, totaling 1200 cents. A cent is a logarithmic unit for measuring intervals, where 100 cents make 1 semitone, and 5–10 cents is the minimum difference people usually detect.

In the context of measuring musical intervals, one must bear in mind the following "equations"

$$\boxed{\text{intervals = ratios}} \tag{102}$$

and

$$\boxed{\text{adding intervals = multiplying ratios}}\,. \tag{103}$$

For example,

$$\begin{aligned} 3 \text{ octaves} &= 1 \text{ octave} + 1 \text{ octave} + 1 \text{ octave} \\ &= \quad 2/1 \quad \times \quad 2/1 \quad \times \quad 2/1 \\ &= 8/1\,. \end{aligned} \tag{104}$$

Likewise, when we define a semitone as a twelfth of an octave, we are referring to the following relationship,

$$\begin{aligned} s+s+s+s+s+s+s+s+s+s+s+s &= \text{octave} \\ r\times r\times r\times r\times r\times r\times r\times r\times r\times r\times r\times r &= 2/1\,, \end{aligned} \tag{105}$$

where s is the semitone and r is the corresponding ratio. To put it concisely:

$$r^{12} = 2\,, \tag{106}$$

which implies

$$r = 2^{1/12}\,. \tag{107}$$

Therefore, the frequency ratio r for a semitone is an irrational number, approximately 1.05946309. Using the same logic, we can determine that a single cent corresponds to a ratio of roughly 1.0005777895.

In Table IV we give a summary of the interval units. When working with musical intervals, it is important to remember that units like cents, semitones, and octaves represent ratios rather than fixed frequencies. To apply these intervals to a specific frequency (such as 440 Hz), you use multiplication or division rather than simple addition. For example, moving up or down an octave means multiplying or dividing the frequency by 2. In the same way, adjusting a frequency by one cent involves multiplying or dividing it by the ratio $2^{1/1200}$.

4. *The Math Behind the Major Musical Scale*

While contemporary music divides the octave into semi-tones and cents, this is a relatively modern development. Long before this, Pythagoras established a system based on natural harmonies, utilizing ratios of 2:1

TABLE IV: Short list of interval units.

Unit Name	Relative Value	Pitch Ratio
octave	fundamental	2/1
semi-tone	1/12 of an octave	$2^{1/12}$
cent	1/100 of a semi-tone	$2^{1/1200}$

(octave) and 3:2 (fifth). Specifically, Pythagorean tuning is derived by placin the 3:2 ratio between 1:1 and 2:1, resulting in an original, ordered scale

$$\boxed{\frac{1}{1}} < \boxed{\frac{3}{2}} < \boxed{\frac{2}{1}} \tag{108}$$

Then, we take the ratio (2/1)/(3/2) = 4/3, which can be depicted as

$$\begin{array}{ccccc} \boxed{\frac{1}{1}} & & \boxed{\frac{3}{2}} & & \boxed{\frac{2}{1}} \\ | & -- \quad -- & \| & -- \quad -- & | \\ & 3/2 & & 4/3 & \end{array} \tag{109}$$

or a microtonal scale defined by cents (¢),

$$\begin{array}{ccccc} \boxed{\frac{1}{1}} & & \boxed{\frac{3}{2}} & & \boxed{\frac{2}{1}}. \\ | & -- \quad -- & \| & -- \quad -- & | \\ & 702\ ¢ & & 498\ ¢ & \end{array} \tag{110}$$

We take our new ratio of 4/3 and place it between 1/1 and 3/2 to create this ordinal scale

$$\boxed{\frac{1}{1}} < \boxed{\frac{4}{3}} < \boxed{\frac{3}{2}} < \boxed{\frac{2}{1}}. \tag{111}$$

We next measure the interval between 4/3 and 3/2, which is calculated as (3/2)/(4/3) = 9/8 and leads to

$$\begin{array}{ccccccc} \boxed{\frac{1}{1}} & & \boxed{\frac{4}{3}} & & \boxed{\frac{3}{2}} & & \boxed{\frac{2}{1}}, \\ | & -- \quad -- & \| & -- \quad -- & \| & -- \quad -- & | \\ & 4/3 & & 9/8 & & 4/3 & \end{array} \tag{112}$$

or a scale with intervals quantified in cents

$$\begin{array}{ccccccc} \boxed{\frac{1}{1}} & & \boxed{\frac{4}{3}} & & \boxed{\frac{3}{2}} & & \boxed{\frac{2}{1}}. \\ | & -- \quad -- & \| & -- \quad -- & \| & -- \quad -- & | \\ & 498\ ¢ & & 204\ ¢ & & 498\ ¢ & \end{array} \tag{113}$$

In Table V we are adding three specifically named intervals (the perfect fifth, perfect fourth, and perfect second) to our list of interval units.

TABLE V: Extended list of interval units.

Unit Name	Pitch Ratio	Cent Value
octave	2:1	1200
perfect fifth	3:2	702
perfect fourth	4:3	498
perfect second	9:8	204
semi-tone	$2^{1/12}$	100
cent	$2^{1/1200}$	1

For the next step of the Pythagorean construction, we use the 9/8 ratio (204 ¢) as our base increment to fill in four more ratios

$$\boxed{\frac{1}{1}} \ll \boxed{\frac{9}{8}} \ll \boxed{\frac{81}{64}} < \boxed{\frac{4}{3}} \ll \boxed{\frac{3}{2}} \ll \boxed{\frac{27}{16}} \ll \boxed{\frac{243}{128}} < \boxed{\frac{2}{1}}.$$

Calculating the final two increments is simple using cents: both are exactly 90. While the resulting ratio (256/243) is cumbersome, we will label this 90 ¢ interval a *half-step* and the standard 204 ¢ interval a *whole-step*, despite 90 being slightly less than half of 204

$$\begin{array}{c} \boxed{\frac{1}{1}} \ll \boxed{\frac{9}{8}} \ll \boxed{\frac{81}{64}} < \boxed{\frac{4}{3}} \ll \boxed{\frac{3}{2}} \ll \boxed{\frac{27}{16}} \ll \boxed{\frac{243}{128}} < \boxed{\frac{2}{1}}, \\ |---|---|---|---|---|---|---| \\ 204\ ¢ \quad 204\ ¢ \quad 90\ ¢ \quad 204\ ¢ \quad 204\ ¢ \quad 204\ ¢ \quad 90\ ¢ \end{array}$$

where $\ll$ indicates a whole-step and $<$ indicates a half-step. Finally, we arrive at the Pythagorean construction of the major scale, represented in its familiar form

$$\begin{array}{c} \boxed{\text{do}} \ll \boxed{\text{re}} \ll \boxed{\text{mi}} < \boxed{\text{fa}} \ll \boxed{\text{sol}} \ll \boxed{\text{la}} \ll \boxed{\text{ti}} < \boxed{\text{do}}. \\ |---|---|---|---|---|---|---| \\ 204\ ¢ \quad 204\ ¢ \quad 90\ ¢ \quad 204\ ¢ \quad 204\ ¢ \quad 204\ ¢ \quad 90\ ¢ \end{array}$$

In musical terms, this corresponds to the white keys of a piano, beginning with C

$$\boxed{C} \ll \boxed{D} \ll \boxed{E} < \boxed{F} \ll \boxed{G} \ll \boxed{A} \ll \boxed{B} < \boxed{C} \tag{114}$$

and it is known as the C-major scale.

IV. ELEVEN WICKED QUESTIONS TO PONDER

1. An echo off a building is heard 5 s after the sound is created. If the average air temperature is 20°C, how far is the building from the observer?

2. Consider a lightly damped system in which the amplitude of the oscillator takes times $t_{A_{1/2}}$ to fall to half its initial value. How long does it take the total mechanical energy of the system to fall to half its initial value?

3. Show that the period of a simple pendulum is given by (30).

4. A 75-cm-long string guitar of mass 2.1 g is near an open-closed tube that is also 75 cm long. How much tension should be in the string if it is to produce resonance in (its fundamental mode) with the third harmonic on the tube?

5. A 50-dB sound wave strikes an eardrum whose area is 5×10^{-5} m^2. How much energy is absorbed by the eardrum per second?

6. How many clarinets would it take to equal the acoustic power of a pair of cymbals?

7. How far would you have to be from the violin (when it's at its quietest) in order to barely detect its sound?

8. If the sound emerges from the trumpet at 0.5 m from the trumpet player's ear, how many decibels does he experience during his trumpet solo?

9. If the orchestra conductor wanted to produce the sound intensity of the entire orchestra, but use only violins (at their quietest) to produce the sound, how many would need to be used? How about if it were to be done with bass drums (at their loudest)?

10. The conductor has become concerned about the high decibel level and wants to make sure he does not experience more than 100 dB. How far away from the orchestra must he stand?

11. A typical marching cadence is 120-135 steps per minute, with a standard step size of 22.5 inches (57 cm). This results in a speed of roughly 0.5 m/s to 1.5 m/s directly toward or away from a listener. Show that for a marching band moving at typical speeds, the doppler shift is usually right at the threshold of human detection (typically 3 cents).

Acknowledgments

Figures in Secs. II B, II C, and II D are courtesy of Jorge Fernandez Soriano. The work of L.A.A. is supported by the U.S. National Science Foundation under Grant No. PHY-2412679.

[1] I. Newton, Philosophiæ Naturalis Principia Mathematica, (1687). Reprinted on *On the Shoulders of Giants: The Great Works of Physics and Astronomy*, (Ed. S. Hawking, Running Press, Philadelphia, 2002) ISBN 0-7624-1348-4; p.733.

[2] A. Einstein, On the electrodynamics of moving bodies, Annalen Phys. **17**, 891 (1905) [Annalen Phys. **14**, 194 (2005)]. doi:10.1002/andp.200590006

[3] G. G. Coriolis, Du Calcul de l'Effet des Machines (Paris, Carilian-Goeury, 1829).

[4] W. J. M. Rankine (1853) On the general law of the transformation of energy, Proceedings of the Philosophical Society of Glasgow **3** No. 5 (1853).

[5] L. A. Anchordoqui and G. V. Anchordoqui, Understanding the wicked world that surrounds you, [arXiv:1711.07445 [physics.ed-ph]].

[6] J. Dalton, New system of chemical philosophy (R. Bickerstaff, Strand, London, 1808).

[7] J. J. Thomson, Cathode rays, Phil. Mag. Ser. 5 **44**, 293 (1897). doi:10.1080/14786449708621070

[8] E. Rutherford, The scattering of alpha and beta particles by matter and the structure of the atom, Phil. Mag. Ser. 6 **21**, 669 (1911). doi:10.1080/14786440508637080

[9] N. Bohr, On the Constitution of atoms and molecules, Phil. Mag. Ser. 6 **26**, 1 (1913). doi:10.1080/14786441308634955

[10] N. Bohr, On the Constitution of Atoms and Molecules II: Systems containing only a single nucleus, Phil. Mag. Ser. 6 **26**, 476 (1913). doi:10.1080/14786441308634993

[11] Sir. W. Thomson (Lord Kelvin) On an absolute thermometric scale Phil. Mag. [series 3] **33**, 313 (1848).

[12] D. G. Fahrenheit Experiments performed concerning the degree of heat of some boiling liquids, Phil. Trans. R. Soc. Lond. **33**, 1 (1724).

[13] A. Celsius, Observations about two stable degrees on a thermometer, Proceedings of the Royal Swedish Academy of Sciences **3**, 171 (1742).

[14] R. Clausius, The mechanical theory of heat, with its applications to the steam engine and to the physical properties of bodies, (Edited by T. Archer Hirst, John Van Voorst, London, 1867).

[15] S. Carnot, Reflections on the motive power of heat and on machines fitted to develop that power (J. Wiley & Sons, New York, 1897).

[16] W. Nernst, The New Heat Theorem: Its Foundations in Theory and Experiment, (Methuen & Co. Ltd., London, 1926).

[17] B. Thompson (Count Rumford) An inquiry concerning the source of the heat which is excited by friction., Phil. Trans. R. Soc. Lond. **88**, 80 (1798).

[18] J. P. Joule, On the mechanical equivalent of heat Phil. Trans. R. Soc. Lond. **140**, 61 (1850).

[19] R. E. Berg and D. G. Stork, The Physics of Sound, (Prentice Hall College Div., 1982) ISBN 10: 0136742831.

[20] R. Hooke, De Potentia Restitutiva, or of Spring. Explaining the Power of Springing Bodies Printed for John Martyn Printer to the Royal Society, London (1678).

[21] C. Doppler, About the coloured light of the binary stars and some other stars of the heavens Abh. Königl. Böhm. Ges. Wiss. **2**, 465 (1843).

[22] L. A. Anchordoqui, Measuring the speed of sound in air with standing waves, https://youtu.be/PsDtxkDRAV4.

[23] National Physical Laboratory (R. Underwood and L. Bevilaqua), Measuring the speed of sound with toilet rolls,

https://www.youtube.com/watch?v=ihKYVqkajig.
[24] A. Guth, The Early Universe, MIT Fall 2018 course lectures.
[25] J. Napier, Mirifici Logarithmorum Canonis Descriptio [The Description of the Wonderful Rule of Logarithms] (in Latin), (Edinburgh, Scotland: Andrew Hart, 1614).
[26] H. Briggs, Logarithmorum chilias prima, (London, 1617).
[27] D. R. Lapp, The Physics of Music and Musical Instruments (Tufts University, Medford, MA, 2003).

Answers and Comments on the Questions

1. An echo is a reflected sound wave. For an observer to hear an echo, the sound must travel from the observer to the building and then bounce back to the observer. This means the total distance the sound travels is twice the distance between the observer and the building. Since the total distance travelled by the wave is $d_{\rm tot} = ut = 1715$ m, the The building is 857.5 m away from the observer.

2. For a damped harmonic oscillator, the amplitude A decays exponentially over time according to a decay constant γ

$$A(t) = A_0 e^{-\gamma t}\,, \tag{115}$$

where $e \simeq 2.72$. Substituting (31) into the mechanical energy relation (26) we find that the mechanical energy decays as fast as

$$E_{\rm mechanical}(t) = \frac{1}{2} k A_0^2 e^{-2\gamma t}\,. \tag{116}$$

This gives us the energy decay rate

$$E_{\rm mechanical}(t) = E_0 e^{-2\gamma t} \tag{117}$$

Next, we compare the half-life times. Let $t_{A_{1/2}}$ be the time where $A(t_{A_{1/2}}) = A_0/2$ and $t_{E_{1/2}}$ be the time where $E_{\rm mechanical}(t_{E_{1/2}}) = E_0/2$. For the amplitude, we have $e^{-\gamma t_{A_{1/2}}} = 1/2$ whereas for the energy, $e^{-2\gamma t_{E_{1/2}}} = 1/2$. Since both expressions equal 1/2, their exponents must be equal $\gamma t_{A_{1/2}} = 2\gamma t_{E_{1/2}}$. Dividing both sides by γ, we obtain

$$t_{A_{1/2}} = 2 t_{E_{1/2}} \Rightarrow t_{E_{1/2}} = t_{A_{1/2}/2}\,. \tag{118}$$

The energy reduces to half its value in $t_{A_{1/2}}/2$ because energy depends on the square of the amplitude. In summary, when the amplitude is squared, the exponential decay rate effectively doubles, meaning any percentage reduction in energy happens in half the time required for the same percentage reduction in amplitude.

3. Imagine a point P moving in a circle of radius R at a constant speed v (uniform circular motion). Look at only the horizontal motion of this point (its shadow on the x-axis). The shadow moves back and forth exactly like a pendulum. The total time for one full circle is the period τ. For uniform circular motion, the acceleration is always directed toward the center and is given by $a_c = v^2/R$. Using the relation $v = 2\pi R/\tau$, the circular motion shadow acceleration can be expressed as

$$a = (2\pi)^2 x/\tau^2\,. \tag{119}$$

Next, we look at the actual forces on a pendulum of length L at a small angle θ. Gravity pulls the bob down with force mg, but only the component tangent to the arc (the "restoring force") causes motion: $F = mg \sin theta$. For small angles, geometry tells us that $\sin\theta \approx x/L$, where x is the horizontal displacement. The restoring acceleration is then

$$a = F/m = g\sin\theta \approx g/L\,. \tag{120}$$

To find the period, set the circular motion shadow acceleration equal to the acceleration of the physical pendulum. Cancel the displacement x from both sides and solving for τ we find that the period of a simple pendulum is given by (30).

4. The frequency of the guitar string is to be the same as the third harmonic ($n = 3$) of the open-closed tube. The resonance frequencies of an open-closed tube of length L are

$$f_n = \frac{nu}{4L}\,, \tag{121}$$

with $n = 1, 3, 5$ and the frequency of the stretched string is

$$f = \frac{1}{2L}\sqrt{\frac{T}{m/L}}, \tag{122}$$

where $u = 343$ m/s, $L = 75$ cm and $m = 2.1$ g. Equate the two frequencies and solve for the tension

$$T = \frac{9u^2 m}{4L} = 740\ \text{N}. \tag{123}$$

5. The energy absorbed per second is the power of the wave, which is the intensity times the area. Now,

$$I = 10^{L/10} I_0 = 10^{-7}\ \text{W/m}^2 \tag{124}$$

where we have taken $L = 50$ dB and $I_0 = 10^{12}$ W/m^2. Since the area of the eardrum is 5×10^{-5} m^2, the power is $P = 5 \times 10^{-12}$ W.

6. To find out how many times more powerful one sound is than another, first determine the difference between their sound intensity levels:

$$\Delta L \ = \ L_{\text{cymbals}} - L_{\text{clarinet}} = 98.8\ \text{dB} - 76\ \text{dB} = 22.8\ \text{dB}. \tag{125}$$

Next we relate the decibel difference to the power ratio. Specifically, the relationship between a change in decibels (ΔL) and the ratio of acoustic powers (n) is given by

$$\Delta L = 10 \log_{10} n. \tag{126}$$

To solve for the number of clarinets, we rearrange the formula

$$n = 10^{\Delta L/10}. \tag{127}$$

Substituting (125) into (127) we find that it would take 191 clarinets to match the acoustic power of the cymbals.

7. The *barely detectable* limit for human hearing is generally defined as the threshold of hearing, which corresponds to 0 dB. Sound level decreases as you move further from a point source. The relationship between sound level in decibels and distance r is governed by

$$\Delta L = -10 \log\left[\left(\frac{r_2}{r_1}\right)^2\right] = 20 \log\left(\frac{r_2}{r_1}\right). \tag{128}$$

Taking, $L_1 = 34.8$ dB as the initial source level, $r_1 = 10$ m, and $L_2 = 0$ dB as the target sound level, we rearrange (128) to solve for

$$r_2 = r_1 10^{\Delta L/20} \simeq 550\ \text{m}. \tag{129}$$

8. The intensity of sound from a point source decreases with the square of the distance. In decibels this relationship is expressed as:

$$L_2 = L_1 + 20 \log_{10}\left(\frac{r_1}{r_2}\right), \tag{130}$$

where $L_1 = 83.9$ dB is the initial sound level, $r_1 = 10$ m is the initial distance, and $r_2 = 0.5$ m is the new distance. Plug the given values in (130) we obtain 110 dB.

9. To determine how many violins or bass drums are needed to match the sound intensity of an entire orchestra, we must compare their sound power. Sound power is an additive property, meaning the total power of n identical sources is simply n times the power of a single source. Let's first calculate a single violin power. The intensity leve of a single violin at its quietest is $L = 34.8$ dB at a distance of 10 m. We first find the sound intensity I using the reference intensity $I_0 = 10^{-12}$ W/m^2,

$$L = 10 \log_{10}\left(\frac{I}{I_0}\right) \Rightarrow I = I_0 10^{L/10} \approx 3 \times 10^{-9}\ \text{W/m}^2. \tag{131}$$

Assuming the violin radiates sound spherically, the sound power P_{violin} is calculated by multiplying intensity by the surface area of a sphere ($4\pi r^2$) at $r = 10$ m,

$$P_{\text{violin}} = I \cdot 4\pi r^2 \approx 3.79 \times 10^{-6}\ \text{W}\,. \tag{132}$$

The total power for a 75-piece orchestra is $P_{\text{orchestra}} = 67$ W, so the number of violins needed is

$$n_{\text{violins}} = P_{\text{orchestra}}/P_{\text{violins}} = 17,678,100\,; \tag{133}$$

i.e., the orchestra conductor would need approximately 17.7 million violins played at their quietest to match the power of the full orchestra.

A bass drum at its loudest typically produces a sound intensity level of approximately $L = 103$ dB at 10 m. We can use the same logic to find its power

$$P_{\text{drum}} \approx 25\ \text{W} \tag{134}$$

and the number of drums

$$n_{\text{drum}} = P_{\text{orchestra}}/P_{\text{drum}} \approx 2.7\,; \tag{135}$$

i.e., the orchestra conductor would need approximately 3 bass drums at its loudest to match the power of the full orchestra.

10. First, we convert the desired decibel level $L = 100$ dB into sound intensity I using the threshold of hearing I_0 as a reference

$$L = 10 \log_{10}\left(\frac{I}{I_0}\right) \tag{136}$$

so

$$I = 10^{L/10} I_0 = 0.01\text{W/m}^2 \tag{137}$$

Assuming the sound spreads uniformly in all directions (spherical propagation), we use the inverse square law for intensity

$$I = \frac{P}{4\pi r^2}\,. \tag{138}$$

Rearranging to solve for the distance

$$r = \sqrt{\frac{P}{4\pi I}} \approx 23\ \text{m}\,. \tag{139}$$

11. If the source (band) is moving and the observer is stationary the Dopple shift is estimated to be

$$f_O = f_S\,\frac{u}{u - v}\,, \tag{140}$$

and so the ratio of the frequencies is

$$f_O/f_S \approx 1.00439 \tag{141}$$

where we have taken the maximum speed of a typical marcher ($v \approx .1.5$ m/s), the speed of sound $u = 343$ m/s at 20°C.

The human ear perceives pitch changes in cents

$$n = 1200 \cdot \log_2\left(\frac{f_O}{f_S}\right) \approx 7.6\ \text{cents}. \tag{142}$$

However, this assumes maximum speed. Duplicating the calculation assuming the marcher is moving at 0.5 m/s (a more relaxed pace), the shift is roughly 2.5 cents flat (moving away) or sharp (moving toward).