CLASSICAL MECHANICS

D. A. Garanin

September 30, 2015

1 Introduction

Mechanics is part of physics studying motion of material bodies or conditions of their equilibrium. The latter is the subject of *statics* that is important in engineering. General properties of motion of bodies regardless of the source of motion (in particular, the role of constraints) belong to *kinematics*. Finally, motion caused by forces or interactions is the subject of *dynamics*, the biggest and most important part of mechanics.

Concerning systems studied, mechanics can be divided into mechanics of *material points*, mechanics of *rigid bodies*, mechanics of *elastic bodies*, and mechanics of fluids: *hydro*- and *aerodynamics*. At the core of each of these areas of mechanics is the equation of motion, Newton's second law. Mechanics of material points is described by ordinary differential equations (ODE). One can distinguish between mechanics of one or few bodies and mechanics of many-body systems. Mechanics of rigid bodies is also described by ordinary differential equations, including positions and velocities of their centers and the angles defining their orientation. Mechanics of elastic bodies and fluids (that is, mechanics of continuum) is more complicated and described by partial differential equation. In many cases mechanics of continuum is coupled to thermodynamics, especially in aerodynamics. The subject of this course are systems described by ODE, including particles and rigid bodies.

There are two limitations on classical mechanics. First, speeds of the objects should be much smaller than the speed of light, $v \ll c$, otherwise it becomes *relativistic mechanics*. Second, the bodies should have a sufficiently large mass and/or kinetic energy. For small particles such as electrons one has to use *quantum mechanics*.

Regarding theoretical approaches, mechanics splits into three parts: Newtonian, Lagrangian, and Hamiltonian mechanics. Newtonian mechanics is most straightforward in its formulation and is based on Newton's second law. It is efficient in most cases, especially for consideration of particles under the influence of forces. Lagrangian mechanics is more sophisticated and based of the least action principle. It is efficient for consideration of more general mechanical systems having constraints, in particular, mechanisms. Hamiltonian mechanics is even more sophisticated less practical in most cases. Its significance is in bridging classical mechanics to quantum mechanics.

In this course we will consider Newtonian, Lagrangian, and Hamiltonian mechanics, as well as some advanced additional topics.

Part I

Newtonian Mechanics

The basis of Newtonian mechanics are Newton's laws, especially second Newton's law being the equation of motion of a particle of mass *m* subject to the influence of a force **F**

$$m\ddot{\mathbf{r}} = \mathbf{F}. \tag{1}$$

Figure 1: Overview of mechanics Here $\ddot{\mathbf{r}} \equiv d^2\mathbf{r}/dt^2 \equiv \partial_t^2\mathbf{r}$ is the double time derivative of the position vector $\mathbf{r}$ of the particle, that is, its acceleration $\mathbf{a}$. This second-order differential equation can be written as the system of two first-order differential equations

$$\dot{\mathbf{r}} = \mathbf{v}$$

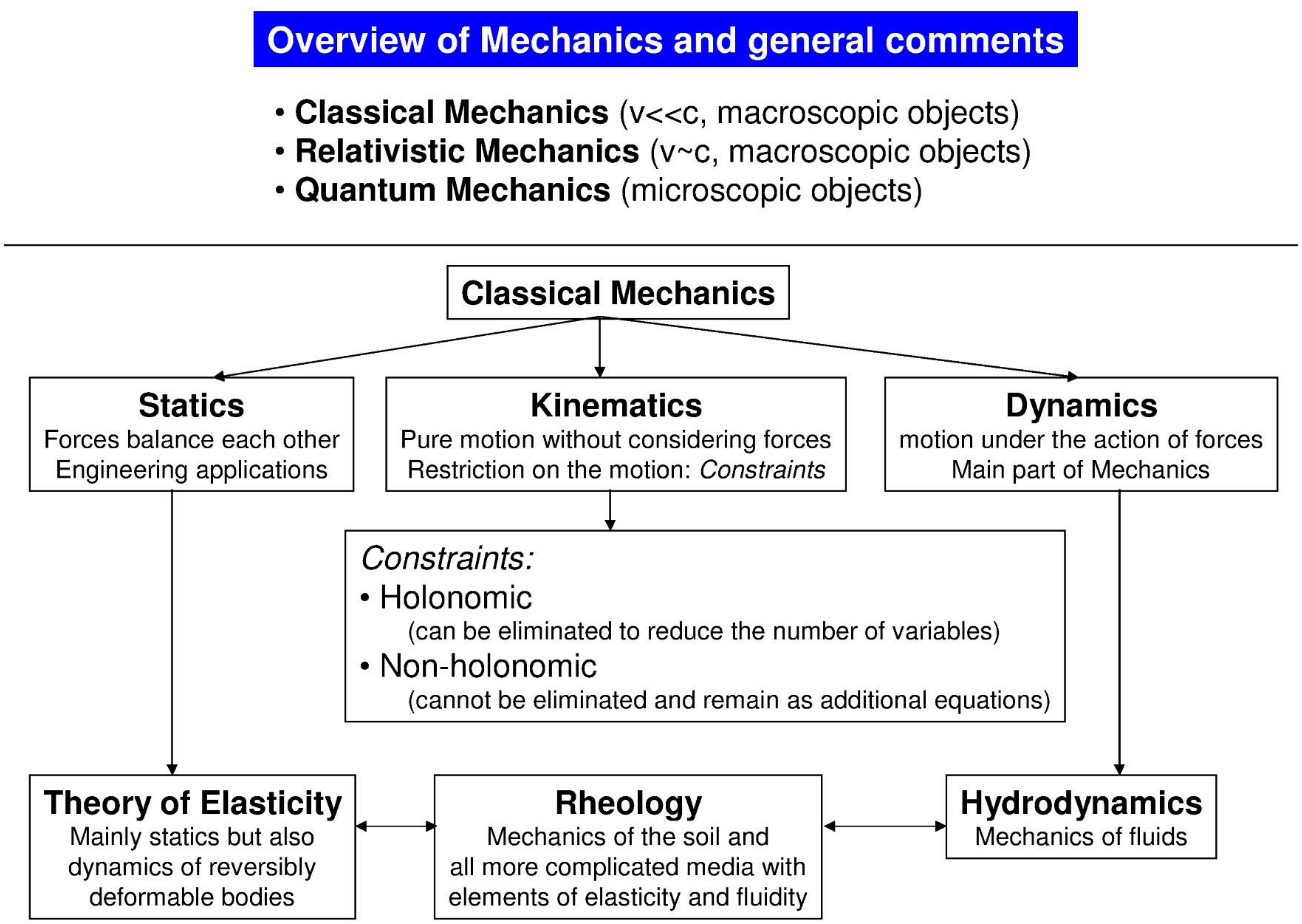
$$\dot{\mathbf{v}} = \mathbf{F}/m \tag{2}$$

where

$$\mathbf{v} = \dot{\mathbf{r}} \tag{3}$$

is particle's velocity. The force $\mathbf{F}$ can depend on particle's position and velocity, as well as explicitly on time: $\mathbf{F} = \mathbf{F}(\mathbf{r}, \mathbf{v}, t)$.

In the absence of the force (free particle), the solution of the above equation of motion is



$$\mathbf{v} = \mathbf{v}_0$$

$$\mathbf{r} = \mathbf{r}_0 + \mathbf{v}_0 t, \tag{4}$$

where $\mathbf{v}_0$ and $\mathbf{r}_0$ are integration constants of the ODE above, being physically the initial conditions: velocity and position at $t = 0$. This is first Newton's law, now having only a historical meaning.

Newton's second law is a differential equation for in general three-component vector variable $\mathbf{r} \equiv (x,y,z)$. Correspondingly Eq. (1) can be written in components as a system of three equations

$$m\ddot{x} = F_x$$

$$m\ddot{y} = F_y$$

$$m\ddot{z} = F_z. \tag{5}$$

In general, these three equations are not independent and may be coupled via the force depending on all three position and/or velocity components.

Systems of N particles are described by in generally coupled system of ODE's consisting of second Newton's laws for each particle

$$m_i\ddot{\mathbf{r}}_i = \mathbf{F}_i, \qquad i = 1,...,N. \tag{6}$$

Each force $\mathbf{F}_i$ can be represented as the sum of the external force $\mathbf{F}^{ext}_i$ and inter-particle interaction forces $\mathbf{f}_{ij}$,

$$\mathbf{F}_i = \mathbf{F}^{ext}_i + \sum_j \mathbf{f}_{ij}. \tag{7}$$

According to Newton's third law, interaction forces are anti-symmetric,

$$\mathbf{f}_{ij} = -\mathbf{f}_{ji}. \tag{8}$$

2 Mechanics of a single particle

Here we consider basic examples of solution of equations of motion for a single particle applying different basic mathematical methods.

2.1 Motion with a constant force

For $\mathbf{F}$ = const in Eq. (1) one readily finds the solution

$$\mathbf{v} = \mathbf{v}_0 + \mathbf{a}t$$

$$\mathbf{r} = \mathbf{r}_0 + \mathbf{v}_0 t + \frac{1}{2}\mathbf{a}t^2, \tag{9}$$

where $\mathbf{a} = \mathbf{F}/m$ is constant acceleration. The validity of this solution can be checked, in particular, by differentiation over time. The trajectory $\mathbf{r}(t)$ (a parabola) is confined to the plane, orientation of which is specified by the two vectors $\mathbf{v}_0$ and $\mathbf{a}$. Thus this motion is effectively two-dimensional. It is convenient to choose the coordinate system such that the trajectory is in xy plane while $z = 0$.

2.2 Motion with a viscous damping

If a body is moving in a viscous fluid or in the air at a small enough speed, it is experiencing a drag force. For a symmetric particle's shape, the drag force is opposite to its velocity,

$$\mathbf{F}_v = -\alpha\mathbf{v}. \tag{10}$$

Newton's second law with a drag force, is usually written it explicitly as,

$$m\ddot{\mathbf{r}} + \alpha\dot{\mathbf{r}} = \mathbf{F}, \tag{11}$$

where $\mathbf{F}$ means all other forces. This equation can be rewritten as

$$\ddot{\mathbf{r}} + \Gamma\dot{\mathbf{r}} = \mathbf{F}/m, \tag{12}$$

where $\Gamma \equiv \alpha/m$ is a characteristic relaxation rate, measured in s^{-1}. One can rewrite this equation via the velocity,

$$\dot{\mathbf{v}} + \Gamma \mathbf{v} = \mathbf{F}/m. \tag{13}$$

After solving this first-order ODE, one can find $\mathbf{r}(t)$ by simple integration, $\mathbf{r}(t) = \int dt\,\mathbf{v}(t)$.

Let us start with the uniform ODE with $\mathbf{F} = 0$, i.e.,

$$\dot{\mathbf{v}} + \Gamma \mathbf{v} = 0. \tag{14}$$

In this case the motion is one-dimensional along a line in 3d space. It is convenient to choose the coordinates so that the motion is along x axis, while $y = z = 0$. Using, for brevity, the notation $v_x = v$, one obtains the equation

$$\dot{v} + \Gamma v = 0 \tag{15}$$

that can be solved as follows:

$$\begin{aligned} \frac{dv}{v} &= -\Gamma dt \\ \int \frac{dv}{v} &= -\Gamma \int dt \end{aligned}$$

$$\ln v = -\Gamma t + \text{const} \tag{16}$$

that finally yields

$$v = v_0 e^{-\Gamma t}. \tag{17}$$

Returning to the general vector form, one obtains

$$\mathbf{v} = \mathbf{v}_0 e^{-\Gamma t}. \tag{18}$$

Now, integrating this equation yields

$$\mathbf{r} = \mathbf{r}_0 + \frac{\mathbf{v}_0}{\Gamma}\left(1 - e^{-\Gamma t}\right). \tag{19}$$

This method of solution works for a class of nonlinear first-order ODE.

For linear uniform ODE one can use a more powerful method based on searching the solution in the exponential form such as

$$\mathbf{v} \propto e^{\lambda t}. \tag{20}$$

Substitution into Eq. (14) yields the algebraic equation for λ

$$\lambda + \Gamma = 0 \tag{21}$$

having the solution $\lambda = -\Gamma$. Thus the solution of the ODE has the form

$$\mathbf{v} = \mathbf{C}e{-}\Gamma t = \mathbf{v}0e{-}\Gamma t \tag{22}$$

that coincides with Eq. (18).

In the presence of a constant force, e.g., gravity, Eq. (13) can be solved by a small modification of the method above. The problem has a non-trivial asymptotic solution that can be found by setting $\dot{\mathbf{v}} = 0$. This yields

$$\mathbf{v}_\infty = \mathbf{F}/(m\Gamma), \tag{23}$$

the stationary velocity at large times. Then one can rewrite Eq. (13) in terms of the new variable $\mathbf{u} \equiv \mathbf{v}-\mathbf{v}_\infty$ as

$$\dot{\mathbf{u}} + \Gamma\mathbf{u} = 0. \tag{24}$$

This is mathematically the same equation as before and it can be solved in a similar way. Finally one obtains

$$\mathbf{v} = \mathbf{v}_\infty + (\mathbf{v}_0 - \mathbf{v}_\infty)e^{-\Gamma t}. \tag{25}$$

Integration of this formula yields

$$\mathbf{r} = \mathbf{r}_0 + \mathbf{v}_\infty t + \frac{\mathbf{v}_0 - \mathbf{v}_\infty}{\Gamma}\left(1 - e^{-\Gamma t}\right). \tag{26}$$

Finally, Eq. (13) can be solved in quadratures for any time-dependent force $\mathbf{F}(t)$ by the method of variation of constants. This method allows to find the solution of a non-uniform ODE (or system of ODE), including ODE with coefficients depending on time, if the solution of the corresponding uniform ODE is known. Using this method, one searches for the solution of Eq. (13) in the form

$$\mathbf{v}(t) = \mathbf{C}(t)e^{-\Gamma t}. \tag{27}$$

Substituting this into Eq. (13), one obtains the equation for the “variable constant” $\mathbf{C}(t)$

$$\dot{\mathbf{C}}(t)e^{-\Gamma t} = \mathbf{F}(t)/m. \tag{28}$$

Solving for $\dot{\mathbf{C}}(t)$ and integrating the result, one obtains

$$\mathbf{C}(t) = \int dt' e^{\Gamma t'} \frac{\mathbf{F}(t')}{m}. \tag{29}$$

Substituting this into Eq. (27), one obtains

$$\mathbf{v}(t) = e^{-\Gamma t} \int dt' e^{\Gamma t'} \frac{\mathbf{F}(t')}{m}. \tag{30}$$

In this expression the indefinite integral contains an integration constant. One can work it out rewriting the result in terms of a definite integral as

$$\mathbf{v}(t) = \mathbf{v}_0 e^{-\Gamma t} + e^{-\Gamma t} \int_0^t dt' e^{\Gamma t'} \frac{\mathbf{F}(t')}{m}. \tag{31}$$

To check this formula, one can differentiate it over t, obtaining Eq. (13).

2.3 Harmonic oscillator

Harmonic oscillator in its basic form is a body of mass m attached to a spring with spring constant k. Also including viscous damping and generalizing Eq. (11), one writes down the equation

$$m\ddot{\mathbf{r}} + \alpha\dot{\mathbf{r}} + k\mathbf{r} = \mathbf{F}, \tag{32}$$

where **F** is an external force, as before. In the following we will consider the case of the body performing a linear motion along the x axis. Dividing by the mass, one obtains the equation

$$\ddot{x} + 2\Gamma\dot{x} + \omega_0^2 x = f, \tag{33}$$

where

$$\Gamma \equiv \frac{\alpha}{2m}, \qquad \omega_0 \equiv \sqrt{\frac{k}{m}}, \qquad f \equiv \frac{F}{m}. \tag{34}$$

Here we have defined Γ in a way different from above for the sake of simplicity of the formulas. ω_0 is the frequency of oscillations in the absence of damping.

Solution of the uniform equation ($f = 0$), in accordance with the general method, can be searched in the form

$$x(t) \propto e^{i\Omega t}. \tag{35}$$

The imaginary i has been inserted in anticipation of an oscillating motion of the system. Substituting this into Eq. (33), one obtains the quadratic equation

$$-\Omega^2 + 2i\Gamma\Omega + \omega_0^2 = 0 \tag{36}$$

having the solution

$$\Omega_\pm = i\Gamma \pm \tilde{\omega}_0, \qquad \tilde{\omega}_0 \equiv \sqrt{\omega_0^2 - \Gamma^2}. \tag{37}$$

Thus the solution of the ODE has the form

$$x(t) = C_+ e^{i\Omega_+ t} + C_- e^{i\Omega_- t}, \tag{38}$$

where $C_\pm$ are integration constants. Using the relation

$$e^{i\phi} \equiv \cos\phi + i\sin\phi, \tag{39}$$

one can rewrite the result in explicitly real form

$$x(t) = C_1 e^{-\Gamma t}\cos\tilde{\omega}_0 t + C_2 e^{-\Gamma t}\sin\tilde{\omega}_0 t, \tag{40}$$

where $C_{1,2}$ is another set of integration constants. The latter can be found from the initial conditions

$$x(0) = x_0, \qquad \dot{x}(0) = v_0 \tag{41}$$

that is,

$$x(0) = C_1 = x_0 \tag{42}$$

and

$$\dot{x}(0) = -\Gamma C_1 + \tilde{\omega}_0 C_2 = v_0. \tag{43}$$

One finds

$$C_1 = x_0, \qquad C_2 = \frac{v_0 + \Gamma x_0}{\tilde{\omega}_0}. \tag{44}$$

Thus

$$x(t) = x_0 e^{-\Gamma t} \cos \tilde{\omega}_0 t + \frac{v_0 + \Gamma x_0}{\tilde{\omega}_0} e^{-\Gamma t} \sin \tilde{\omega}_0 t. \tag{45}$$

Let us look at the solution. According to Eq. (37), in the absence of damping, $\Gamma = 0$, the body is oscillating with the frequency ω_0. Damping reduces oscillation frequency that turns to zero at $\Gamma = \omega_0$. In the strong-damping limit $\Gamma > \omega_0$ the motion of the body is aperiodic.

Let us now consider the motion of the harmonic oscillator under the influence of external force. Using the method of variation of constants in Eq. (38), one searches for the solution in the form

$$x(t) = C_+(t) e^{i\Omega_+ t} + C_-(t) e^{i\Omega_- t}. \tag{46}$$

The “variable constants” satisfy the system of equations

$$\begin{aligned} \dot{C}_+(t) x_+(t) + \dot{C}_-(t) x_-(t) &= 0 \\ \dot{C}_+(t) \dot{x}_+(t) + \dot{C}_-(t) \dot{x}_-(t) &= f(t). \end{aligned} \tag{47}$$

Its solution is

$$\dot{C}_+(t) = \frac{\begin{vmatrix} 0 & x_-(t) \\ f(t) & \dot{x}_-(t) \end{vmatrix}}{\begin{vmatrix} x_+(t) & x_-(t) \\ \dot{x}_+(t) & \dot{x}_-(t) \end{vmatrix}} = \frac{-f(t) x_-(t)}{x_+(t)\dot{x}_-(t) - \dot{x}_+(t) x_-(t)}. \tag{48}$$

Using

$$x_+(t)\dot{x}_-(t) - \dot{x}_+(t) x_-(t) = i(\Omega_- - \Omega_+) e^{i(\Omega_- + \Omega_+)t} = -2i\tilde{\omega}_0 e^{-2\Gamma t} \tag{49}$$

one obtains

$$\dot{C}_+(t) = -\frac{i}{2}\frac{f(t)}{\tilde{\omega}_0} e^{i\Omega_- t + 2\Gamma t} = -\frac{i}{2}\frac{f(t)}{\tilde{\omega}_0} e^{-i\Omega_+ t} \tag{50}$$

and, similarly,

$$\dot{C}_-(t) = \frac{i}{2}\frac{f(t)}{\tilde{\omega}_0} e^{i\Omega_+ t + 2\Gamma t} = \frac{i}{2}\frac{f(t)}{\tilde{\omega}_0} e^{-i\Omega_- t}. \tag{51}$$

Integrating these two formulas and substituting the result in Eq. (46), one obtains

$$x(t) = x_{\text{free}}(t) + x_{\text{forced}}(t), \tag{52}$$

where $x_{\text{free}}(t)$ is the solution of the uniform ODE describing the free oscillator and given by Eq. (45) and

$$x_{\text{forced}}(t) = -\frac{i}{2\tilde{\omega}_0} e^{i\Omega_+ t} \int_0^t dt' f(t') e^{-i\Omega_+ t'} + \frac{i}{2\tilde{\omega}_0} e^{i\Omega_- t} \int_0^t dt' f(t') e^{-i\Omega_- t'} \tag{53}$$

is the response to the external force. The latter can be simplified as

$$
\begin{aligned}
(t) &= \frac{i}{2\tilde{\omega}_0} \int_0^t dt' f(t') \left[-e^{i\Omega_+(t-t')} + e^{i\Omega_-(t-t')} \right] \\
&= \frac{i}{2\tilde{\omega}_0} \int_0^t dt' f(t') e^{-\Gamma(t-t')} \left[-e^{i\tilde{\omega}_0(t-t')} + e^{-i\tilde{\omega}_0(t-t')} \right] \\
&= \frac{1}{\tilde{\omega}_0} \int_0^t dt' f(t') e^{-\Gamma(t-t')} \sin\left[\tilde{\omega}_0(t-t')\right].
\end{aligned}
$$

*x*forced

(54)

One can check that $x_{forced}(0) = \dot{x}_{forced}(0) = 0$, that is, the forced solution is independent of the initial conditions and does not change the form of $x_{free}(t)$.

Let us consider the important case of a sinusoidal force

$$f(t) = f_0 \sin\omega t, \tag{55}$$

applied starting from $t = 0$. To compute $x_{forced}(t)$, it is convenient to convert everything into the exponential form, after which integration simplifies:

$$
\begin{aligned}
(t) &= \frac{f_0}{4\tilde{\omega}_0} \int_0^t dt' \left(e^{i\omega t'} - e^{-i\omega t'} \right) e^{-\Gamma(t-t')} \left[-e^{i\tilde{\omega}_0(t-t')} + e^{-i\tilde{\omega}_0(t-t')} \right] \\
&= \frac{f_0}{4\tilde{\omega}_0} \int_0^t dt' e^{-\Gamma(t-t')} \left[-e^{i\omega t'} e^{i\tilde{\omega}_0(t-t')} + e^{i\omega t'} e^{-i\tilde{\omega}_0(t-t')} + c.c. \right]
\end{aligned}
$$

*x*forced

, (56)

where *c.c.* is complex conjugate. Further one proceeds as

$$
\begin{aligned}
(t) &= \frac{f_0}{4\tilde{\omega}_0}\int_0^t dt' \left[-e^{(-\Gamma+i\tilde{\omega}_0)t}e^{(\Gamma-i\tilde{\omega}_0+i\omega)t'} + e^{(-\Gamma-i\tilde{\omega}_0)t}e^{(\Gamma+i\tilde{\omega}_0+i\omega)t'}\right] + c.c. \\
&= \frac{f_0}{4\tilde{\omega}_0}\left[-e^{(-\Gamma+i\tilde{\omega}_0)t}\frac{e^{(\Gamma-i\tilde{\omega}_0+i\omega)t}-1}{\Gamma-i\tilde{\omega}_0+i\omega} + e^{(-\Gamma-i\tilde{\omega}_0)t}\frac{e^{(\Gamma+i\tilde{\omega}_0+i\omega)t}-1}{\Gamma+i\tilde{\omega}_0+i\omega}\right] + c.c. \\
&= \frac{f_0}{4\tilde{\omega}_0}\left[-\frac{e^{i\omega t}-e^{(-\Gamma+i\tilde{\omega}_0)t}}{\Gamma-i\tilde{\omega}_0+i\omega} + \frac{e^{i\omega t}-e^{(-\Gamma-i\tilde{\omega}_0)t}}{\Gamma+i\tilde{\omega}_0+i\omega}\right] + c.c.
\end{aligned}
$$

*x*forced

(57)

The first term in this expression is the so-called resonant term in which the denominator becomes small for ω close to ω_0. The other term is non-resonant term that differs from the first one by replacement $\tilde{\omega}_0 \Rightarrow -\tilde{\omega}_0$. It is sufficient to compute one of these terms, then the other one can be easily obtained from the first one. Let us calculate the resonance term. Adding *c.c.* amounts to doubling the real part of the expression and annihilating its imaginary part. Shortcutting the non-resonant term as *...,* one proceeds as

$$
\begin{aligned}
(t) &= -\frac{f_0}{4\tilde{\omega}_0}\frac{e^{i\omega t}-e^{(-\Gamma+i\tilde{\omega}_0)t}}{\Gamma-i\tilde{\omega}_0+i\omega} + c.c. + \ldots \\
&= -\frac{f_0}{4\tilde{\omega}_0}\frac{\left[e^{i\omega t}-e^{(-\Gamma+i\tilde{\omega}_0)t}\right](\Gamma+i\tilde{\omega}_0-i\omega)}{(\omega-\tilde{\omega}_0)^2+\Gamma^2} + c.c. + \ldots \\
&= -\frac{f_0}{4\tilde{\omega}_0}\frac{\Gamma\left[e^{i\omega t}-e^{(-\Gamma+i\tilde{\omega}_0)t}+c.c.\right] + i(\tilde{\omega}_0-\omega)\left[e^{i\omega t}-e^{(-\Gamma+i\tilde{\omega}_0)t}-c.c.\right]}{(\omega-\tilde{\omega}_0)^2+\Gamma^2} + \ldots \\
&= -\frac{f_0}{2\tilde{\omega}_0}\frac{\Gamma\left[\cos(\omega t)-\cos(\tilde{\omega}_0 t)\,e^{-\Gamma t}\right] - (\tilde{\omega}_0-\omega)\left[\sin(\omega t)-\sin(\tilde{\omega}_0 t)\,e^{-\Gamma t}\right]}{(\omega-\tilde{\omega}_0)^2+\Gamma^2} + \ldots
\end{aligned}
$$

*x*forced

(58)

At the times longer than the relaxation time of the oscillator

$$\tau \equiv \frac{1}{\Gamma} \tag{59}$$

the terms in the above formula that are oscillating at oscillator's own frequency $\tilde{\omega}_0$ die out and only the forced terms oscillating at the frequency ω remain,

$$x_{\rm forced}(t) = \frac{f_0}{2\tilde{\omega}_0}\left[-\frac{\Gamma\cos(\omega t) + (\omega - \tilde{\omega}_0)\sin(\omega t)}{(\omega - \tilde{\omega}_0)^2 + \Gamma^2} + \frac{\Gamma\cos(\omega t) + (\omega + \tilde{\omega}_0)\sin(\omega t)}{(\omega + \tilde{\omega}_0)^2 + \Gamma^2}\right]. \quad (60)$$

Here the terms with cos(ωt) is shifted by quarter of the period with respect to the harmonic force. Exactly at resonance, $\omega = \tilde{\omega}_0$, only this term in the resonant part of the expression becomes dominant and reaches its maximum. Near the resonance, $|\omega - \tilde{\omega}_0| \sim \Gamma \ll \omega_0$, the resonant term is large, and the much smaller non-resonant term can be neglected.

The power absorbed by the oscillator can be calculated via the work of the external force done on the oscillator,

$$P_{\rm abs} = \frac{1}{T}\int_0^T dt \frac{dA}{dt} = \frac{1}{T}\int_0^T dt F(t)\dot{x}(t), \quad (61)$$

where $T \equiv 2\pi/\omega$. In the stationary state near the resonance, $x(t)$ is given by the first term of Eq. (60) and $F(t) = mf_0 \sin\omega t$. One obtains

$$P_{\rm abs} = \frac{1}{T}\int_0^T dt\, \frac{mf_0^2}{2\omega_0}\frac{\Gamma}{(\omega - \omega_0)^2 + \Gamma^2}\omega\sin^2\omega t \cong \frac{mf_0^2}{4}\frac{\Gamma}{(\omega - \omega_0)^2 + \Gamma^2}. \quad (62)$$

Here approximation $\omega \cong \tilde{\omega}_0 \cong \omega_0$ was used that is mandatory since the whole approach ignoring the nonresonant term is valid near the resonance only under that condition that the resonance is narrow, $\Gamma \ll \omega_0$.

The stationary (settled) state of the oscillator under the influence of a harmonic force can be obtained in a much easier way by just searching the solution in the form

$$x(t) = A\sin\omega t + B\cos\omega t. \quad (63)$$

Pre-calculating

$$\dot{x}(t) = A\omega\cos\omega t - B\omega\sin\omega t \quad (64)$$

$$\ddot{x}(t) = -\omega^2 x(t) \quad (65)$$

and inserting it into Eq. (33) using Eq. (55), one obtains

$$-A\omega^2\sin\omega t - B\omega^2\cos\omega t + 2\Gamma A\omega\cos\omega t - 2\Gamma B\omega\sin\omega t + A\omega_0^2\sin\omega t + B\omega_0^2\cos\omega t = f_0\sin\omega t. \quad (66)$$

Equating the coefficients in front of sinωt and cosωt, one obtains the system of linear equations

$$\begin{aligned} -A\omega^2 - 2\Gamma B\omega + A\omega_0^2 &= f_0 \\ -B\omega^2 + 2\Gamma A\omega + B\omega_0^2 &= 0. \end{aligned} \quad (67)$$

From the second equation one obtains $A = B(\omega^2 - \omega_0^2)/(2\Gamma\omega)$. Substituting it into the first equation yields

$$-B\frac{(\omega^2 - \omega_0^2)^2}{2\Gamma\omega} - 2\Gamma\omega B = f_0 \quad (68)$$

and

$$B = -f_0\frac{2\Gamma\omega}{(\omega^2 - \omega_0^2)^2 + 4\Gamma^2\omega^2}, \qquad A = -f_0\frac{\omega^2 - \omega_0^2}{(\omega^2 - \omega_0^2)^2 + 4\Gamma^2\omega^2}, \quad (69)$$

so that

$$x(t) = -f_0\frac{2\Gamma\omega\cos\omega t + (\omega^2 - \omega_0^2)\sin\omega t}{(\omega^2 - \omega_0^2)^2 + 4\Gamma^2\omega^2}.$$

(70)

This solution contains both resonant and nonresonant terms. Near resonance one can write

$$\omega^2 - \omega_0^2 = (\omega - \omega_0)(\omega + \omega_0) \cong 2\omega_0 (\omega - \omega_0) \quad (71)$$

and thus

$$x(t) \cong -\frac{f_0}{2\omega_0} \frac{\Gamma \cos \omega t + (\omega - \omega_0) \sin \omega t}{(\omega - \omega_0)^2 + \Gamma^2}. \quad (72)$$

This coincides with the resonant term in Eq. (60). In general, one should be able to demonstrate that Eqs.
(60) and (70) that are both exact solutions of the problem, are identical.

2.4 Charged particle in a magnetic field

Moving charged particle experiences the so-called Lorentz force from the magnetic field B,

$$F_L = q\,[\mathbf{v} \times \mathbf{B}]. \quad (73)$$

Since this force depends on the velocity, it is convenient to write the equation of motion in terms of the velocity, as was done in the case of the viscous drag force,

$$m\dot{\mathbf{v}} = q\,[\mathbf{v} \times \mathbf{B}]. \quad (74)$$

In the basic case of uniform magnetic field that will be considered below, z axis will be chosen along $\mathbf{B}$, so that $B_z = B$ and $B_x = B_y = 0$. The equation of motion in components has the form

$$\dot{v}_x = \omega_c v_y \quad (75) \quad \dot{v}_y = -\omega_c v_x \quad (76) \quad \dot{v}_z = 0, \quad (77)$$

where

$$\omega_c \equiv \frac{qB}{m} \quad (78)$$

is the cyclotron frequency. Differentiating the first equation over time and substituting the second equation, one obtains the second-order ODE for v_x

$$\ddot{v}_x + \omega_c^2 v_x = 0. \quad (79)$$

This is the harmonic-oscillator equation considered in the pervious section. Thus v_x will be oscillating in time. The same equation can be obtained for v_y, thus v_y will be oscillating, too. On the other hand, motion in the direction of the field is a free motion,

$$\dot{v}_z = 0, \qquad v_z = \text{const} = v_{z0}, \qquad z = z_0 + v_{z0}t. \quad (80)$$

The solution of Eq. (79) can be obtained from Eq. (40):

$$v_x(t) = C_{x1} \cos\omega_c t + C_{x2} \sin\omega_c t. \quad (81)$$

Integration constants obtained from the initial conditions have the form

$$Cx1 = vx(0) = vx0 \quad (82)$$

$$C_{x2} = \dot{v}_x(0)/\omega_c = v_y(0) = v_{y0}, \tag{83}$$

thus the final solution has the form

$$v_x(t) = v_{x0}\cos\omega_c t + v_{y0}\sin\omega_c t. \tag{84}$$

Similarly one obtains

$$v_y(t) = v_{y0}\cos\omega_c t - v_{x0}\sin\omega_c t. \tag{85}$$

Let us calculate

$$\begin{aligned} v_x^2 + v_y^2 &= (v_{x0}\cos\omega_c t + v_{y0}\sin\omega_c t)^2 + (v_{y0}\cos\omega_c t - v_{x0}\sin\omega_c t)^2 \\ &= v_{x0}^2\cos^2\omega_c t^2 + 2v_{x0}v_{y0}\cos\omega_c t\sin\omega_c t + v_{y0}^2\sin^2\omega_c t + v_{x0}^2\sin^2\omega_c t^2 - 2v_{x0}v_{y0}\sin\omega_c t\cos\omega_c t + v_{y0}^2\cos^2\omega_c t \\ &= v_{x0}^2 + v_{y0}^2 = \text{const}. \end{aligned} \tag{86}$$

Thus the vector (v_x,v_y) is rotating at the constant rate ω_c while the kinetic energy $E_k = m\mathbf{v}^2/2$ is conserved. Time dependence of x and y can be obtained by integration of the results for v_x and v_y. One obtains

$$\begin{aligned} x &= x_0 + \frac{v_{x0}}{\omega_c}\sin\omega_c t - \frac{v_{y0}}{\omega_c}\cos\omega_c t \\ y &= y_0 + \frac{v_{y0}}{\omega_c}\sin\omega_c t + \frac{v_{x0}}{\omega_c}\cos\omega_c t. \end{aligned} \tag{87}$$

This trajectory is circular with the center at an arbitrary (x_0,y_0). The radius of the circle R can be found by the calculation similar to the above:

$$R^2 = (x-x_0)^2 + (y-y_0)^2 = \frac{v_{x0}^2 + v_{y0}^2}{\omega_c^2}. \tag{88}$$

Thus the cyclotron radius is given by

$$R = \frac{v_\perp}{\omega_c} = \frac{mv_\perp}{qB}, \tag{89}$$

where $v_\perp = \sqrt{v_x^2 + v_y^2}$.

3 Momentum and angular momentum

Momentum of a particle is defined as

$$\mathbf{p} = m\mathbf{v}. \tag{90}$$

It can be used to write Newton's second law in the form

$$\dot{\mathbf{p}} = \mathbf{F}. \tag{91}$$

In the absence of forces momentum is conserved, as well as the velocity. The above is not essentially new.

Significance of momentum becomes apparent if one considers a system of interacting particles and defines the total momentum as

$$\mathbf{P} = \sum_i \mathbf{p}_i. \tag{92}$$

i

Differentiating it over time one obtains

$$\dot{\mathbf{P}} = \sum_i \dot{\mathbf{p}}_i = \sum_i \mathbf{F}_i. \qquad (93)$$

Separating the forces into external and internal, Eq. (7), and using Newton's third law, Eq. (8), one can see that only external forces change the momentum of the system,

$$\dot{\mathbf{P}} = \sum_i \mathbf{F}_i^{\mathrm{ext}}. \qquad (94)$$

In an isolated system there are only interaction forces between the particles that are changing momenta of individual particles, whereas the total momentum is conserved. Conservation of the total momentum is important, in particular, in collisions. During collisions usually systems can be considered as effectively isolated because collisions occur during a very short time, so that external forces cannot change momenta significantly during collisions. To the contrary, internal forces during collisions are large and important. Angular momentum of a particle is defined by

$$\mathbf{l} = [\mathbf{r} \times \mathbf{p}], \qquad (95)$$

where **r** is the position vector of the particle defined with respect to a particular frame or coordinate system. Thus angular momentum depends on the position of the origin of the frame. The time derivative of the angular momentum is given by

$$\dot{\mathbf{l}} = [\dot{\mathbf{r}} \times \mathbf{p}] + [\mathbf{r} \times \dot{\mathbf{p}}] = [\mathbf{r} \times \mathbf{F}] \equiv \tau. \qquad (96)$$

Here τ is the torque that also depends on the choice of the frame origin. If the force is central, i.e., directed everywhere away from or towards a particular central point, one can choose a frame having the origin at this point, then the torque is zero and angular momentum is conserved.

The total angular momentum of the system is defined by

$$\mathbf{L} = \sum_i [\mathbf{r}_i \times \mathbf{p}_i]. \qquad (97)$$

Its time derivative reads

$$\dot{\mathbf{L}} = \sum_i [\mathbf{r}_i \times \mathbf{F}_i] = \sum_i \left[\mathbf{r}_i \times \mathbf{F}_i^{\mathrm{ext}}\right] + \sum_i \left[\mathbf{r}_i \times \sum_j \mathbf{f}_{ij}\right]. \qquad (98)$$

In the interaction term, renaming i and j and using Newton's third law, one can write

$$\dot{\mathbf{L}}_{\mathrm{int}} = \sum_{ij} [\mathbf{r}_i \times \mathbf{f}_{ij}] = \sum_{ij} [\mathbf{r}_j \times \mathbf{f}_{ji}] = -\sum_{ij} [\mathbf{r}_j \times \mathbf{f}_{ij}]. \qquad (99)$$

One can symmetrize $\dot{\mathbf{L}}_{\mathrm{int}}$ as

$$\dot{\mathbf{L}}_{\mathrm{int}} = \frac{1}{2} \sum_{ij} [(\mathbf{r}_i - \mathbf{r}_j) \times \mathbf{f}_{ij}]. \qquad (100)$$

In nature interaction forces between particles are directed along the line connecting the particles. Thus the cross product in the formula above disappears. The change of the angular momentum of a system is entirely due to external torques,

$$\dot{\mathbf{L}} = \sum_i \langle \mathbf{r}_i \times \mathbf{F}_i^{ext} \rangle = \sum_i \tau_i^{ext} = \tau. \tag{101}$$

Consider another frame with the origin shifted by the vector **a**, so that the positions of the particles in the old frame are given by

$$\mathbf{r}_i = \mathbf{r}'_i + \mathbf{a}. \tag{102}$$

For the total angular momentum one obtains

$$\mathbf{L} = \sum_i \left[(\mathbf{r}'_i + \mathbf{a}) \times \mathbf{p}_i \right] = \mathbf{L}' + [\mathbf{a} \times \mathbf{P}]. \tag{103}$$

Thus, if the total momentum **P** is zero, total angular momentum does not depend on the position of frame's origin. If the system is close to the origin of the primed frame but far from the origin of the original frame, one can say that the term [**a** × **P**] is the angular momentum corresponding to the motion of the system as the whole in the original frame.

Similarly for the torque one can write

$$\boldsymbol{\tau} = \sum_i \left[(\mathbf{r}'_i + \mathbf{a}) \times \mathbf{F}_i \right] = \boldsymbol{\tau}' + \mathbf{a} \times \sum_i \mathbf{F}_i. \tag{104}$$

If the net force acting on the system is zero, torque does not depend on the position of the frame's origin, $\tau = \tau^0$.

4 Work, energy and potential forces

Infinitesimal work done by the force **F** on a particle undergoing displacement $d\mathbf{r}$ is defined by

$$\delta A = \mathbf{F} \cdot d\mathbf{r}. \tag{105}$$

In this formula δA is used instead of dA since work is not a function and the above is not its differential. Power W is work done per unit of time,

$$W = \frac{\delta A}{dt} = \mathbf{F} \cdot \frac{d\mathbf{r}}{dt} = \mathbf{F} \cdot \mathbf{v}. \tag{106}$$

If **F** is acting on a free particle, the latter will accelerate or decelerate, and the work on the particle will change its kinetic energy defined by

$$E_k = \frac{m\mathbf{v}^2}{2}. \tag{107}$$

Indeed, the infinitesimal change of the energy can be related to the infinitesimal work as

$$dE_k = m\mathbf{v} \cdot d\mathbf{v} = m\frac{d\mathbf{v}}{dt} \cdot \mathbf{v}dt = \mathbf{F} \cdot d\mathbf{r} = \delta A. \tag{108}$$

Potential forces are forces that can be expressed via gradients of position-dependent potential energy U,

$$\mathbf{F} = -\nabla U(\mathbf{r}) = -\frac{\partial U}{\partial \mathbf{r}}. \quad (109)$$

Work done by potential forces is independent of the trajectory and depends only on the initial and final positions,

$$A_{12} = \int_{\mathbf{r}_1}^{\mathbf{r}_2} \mathbf{F} \cdot d\mathbf{r} = -\int_{\mathbf{r}_1}^{\mathbf{r}_2} \frac{\partial U}{\partial \mathbf{r}} \cdot d\mathbf{r} = U(\mathbf{r}_1) - U(\mathbf{r}_2). \quad (110)$$

The work on a closed path is zero for potential forces,

$$\oint \mathbf{F} \cdot d\mathbf{r} = 0, \quad (111)$$

because the change of U is zero. Eq. (110) can be rewritten as the definition of potential energy via the force,

$$U(\mathbf{r}) = U(\mathbf{r}_0) - \int_{\mathbf{r}_0}^{\mathbf{r}} \mathbf{F} \cdot d\mathbf{r}. \quad (112)$$

Here $U(\mathbf{r}_0)$ is the (arbitrary) value of potential energy at the reference point $\mathbf{r}_0$. Potential energy a secondary or auxiliary physical quantity, not directly measurable and introduced via the force as a primary quantity, and it is defined up to an arbitrary additive constant. Using Eq. (112), one has first to be sure that the integral does not depend on the path. To check whether **F** is a potential force, one can use the Stokes' theorem to express the closed-path line integral as the flux through the surface spanned by the closed path,

$$\oint \mathbf{F} \cdot d\mathbf{r} = \int_S \mathrm{rot}\mathbf{F} \cdot d\mathbf{S}. \quad (113)$$

If

$$\mathrm{rot}\mathbf{F} = \begin{vmatrix} \mathbf{e}_x & \mathbf{e}_y & \mathbf{e}_z \\ \partial_x & \partial_y & \partial_z \\ F_x & F_y & F_z \end{vmatrix} = 0, \quad (114)$$

F is a potential force.

Not all forces are potential. For instance, friction forces are not. An example of potential forces is gravity near the Earth's surface,

$$U = mg\mathbf{r} \cdot \mathbf{e}_z + \mathrm{const}, \quad (115)$$

where $\mathbf{e}_z$ is the unit vector directed up. The general gravitational force between two masses has the potential energy

$$U = G\frac{mM}{|\mathbf{r}|}. \quad (116)$$

Here the origin of the coordinate system is put at the big mass M (e.g., the sun) and $\mathbf{r}$ points to the position of the small mass m (e.g., the Earth). The gravitational force acting on m has the form

$$\mathbf{F} = -\frac{\partial U}{\partial \mathbf{r}} = -G\frac{mM}{r^2}\frac{\mathbf{r}}{r}. \qquad (117)$$

A special kind force if Lorentz force, Eq. (73). As it is not doing any work, there is no potential energy associated with it. For the same reason this force does not lead to dissipation of energy as friction forces. The total energy of the particle

$$E = \frac{m\mathbf{v}^2}{2} + U(\mathbf{r}) \qquad (118)$$

can be shown to be dynamically conserved (i.e., to be an *integral of motion*), if there are no other forces such as friction. Indeed,

$$\dot{E} = m\dot{\mathbf{v}} \cdot \mathbf{v} + \frac{\partial U}{\partial \mathbf{r}} \cdot \dot{\mathbf{r}} = (m\dot{\mathbf{v}} - \mathbf{F}) \cdot \mathbf{v} = 0 \qquad (119)$$

via Newton's second law. Because of conservation of the total energy, systems with potential forces are called conservative. In the presence of the viscous friction force, Eq. (10), one obtains

$$\dot{E} = m\dot{\mathbf{v}} \cdot \mathbf{v} + \frac{\partial U}{\partial \mathbf{r}} \cdot \dot{\mathbf{r}} = \left(-\alpha\mathbf{v} - \frac{\partial U}{\partial \mathbf{r}} + \frac{\partial U}{\partial \mathbf{r}}\right) \cdot \mathbf{v} = -\alpha\mathbf{v}^2, \qquad (120)$$

dissipation of energy.

Let us obtain the total energy E as the first integral of Newton's second law. Dot-multiplying it by $\mathbf{v}$ and manipulating the expressions, one obtains

$$0 = \left(m\dot{\mathbf{v}} + \frac{\partial U}{\partial \mathbf{r}}\right) \cdot \mathbf{v} = \frac{m}{2}\frac{d\mathbf{v}^2}{dt} + \frac{\partial U}{\partial \mathbf{r}} \cdot \dot{\mathbf{r}} = \frac{d}{dt}\left(\frac{m\mathbf{v}^2}{2} + U(\mathbf{r})\right). \qquad (121)$$

Integrating this, one obtains Eq. (118) with E = const as an integral of motion.

The total energy of a system of interacting particles has the form

$$E = \sum_i \frac{m_i \mathbf{v}_i}{2} + U(\{\mathbf{r}_i\}), \qquad (122)$$

where the potential energy depends, in general, on the positions of all particles. In the absence of nonconservative forces, conservation of this many-body energy follows from Newton's second law, as above. In most cases, potential energy includes one-particle and two-particle terms,

$$U(\{\mathbf{r}_i\}) = \sum_i U_0(\mathbf{r}_i) + \frac{1}{2}\sum_{ij} V(|\mathbf{r}_i - \mathbf{r}_j|). \qquad (123)$$

The factor 1/2 in the interaction term is inserted to compensate for double counting of interacting pairs *ij* and *ji*, so that the interaction energy between each two particles is just V. Using

$$\frac{\partial |\mathbf{r}|}{\partial \mathbf{r}} = \frac{\mathbf{r}}{|\mathbf{r}|}, \qquad (124)$$

one obtains the corresponding force,

$$\mathbf{F}_i = -\frac{\partial U}{\partial \mathbf{r}_i} = -\frac{\partial U_0}{\partial \mathbf{r}_i} + \sum_j \mathbf{f}_{ij}, \tag{125}$$

where interaction forces are given by

$$\mathbf{f}_{ij} = -V'(|\mathbf{r}_i - \mathbf{r}_j|)\frac{\mathbf{r}_i - \mathbf{r}_j}{|\mathbf{r}_i - \mathbf{r}_j|}. \tag{126}$$

Here V' is the derivative of the function over its argument. As it was said in the comment to Eq. (100), interaction forces are directed along the line connecting the two particles.

5 Center of mass, reduced mass

For any system of point masses one can define the center of mass (CM) or center of inertia as

$$\mathbf{R} = \frac{1}{M}\sum_i m_i \mathbf{r}_i, \tag{127}$$

where $M = \sum_i m_i$ is the total mass. For a solid body CM is defined by a corresponding integral. The velocity of CM is related to the total momentum,

$$\mathbf{V} = \frac{1}{M}\sum_i m_i \mathbf{v}_i = \frac{\mathbf{P}}{M}. \tag{128}$$

This formula evokes an image of a system considered as one body of mass M moving with the velocity $\mathbf{V}$. Dynamics of CM is due to the external forces,

$$M\dot{\mathbf{V}} = \dot{\mathbf{P}} = \sum_i \mathbf{F}_{\mathrm{ext}\,i}, \tag{129}$$

where Eq. (94) was used. CM of an isolated system is moving with a constant velocity.

In many cases it is convenient to put the origin of the coordinate system at the center of mass since it leads to simplifications. In particular, in the CM frame the system is at rest as the whole and there is only internal motion. Let the primed frame in Eq. (102) be CM frame, so that

$$\mathbf{r}_i = \mathbf{r}'_i + \mathbf{R} \tag{130}$$

and

$$\mathbf{v}_i = \mathbf{v}'_i + \mathbf{V}. \tag{131}$$

Position of the center of mass in the CM frame is zero,

$$\mathbf{R}' = \frac{1}{M}\sum_i m_i \mathbf{r}'_i = \frac{1}{M}\sum_i m_i (\mathbf{r}_i - \mathbf{R}) = \mathbf{R} - \mathbf{R} = 0. \tag{132}$$

Total momentum in the CM frame is zero, too,

$$\mathbf{P}' = \sum_i m_i \mathbf{v}'_i = \sum_i m_i (\mathbf{v}_i - \mathbf{V}) = \mathbf{P} - \mathbf{P} = 0. \tag{133}$$

Angular momentum defined by Eq. (97) in the CM frame becomes

$$\mathbf{L}' = \sum [\mathbf{r}'_i \times \mathbf{p}'_i] = \sum [(\mathbf{r}_i - \mathbf{R}) \times m_i (\mathbf{v}_i - \mathbf{V})]$$

$$= \mathbf{L} + \mathbf{R}\times\mathbf{P} - \mathbf{R}\times\sum_i m_i\mathbf{v}_i - \sum_i m_i\mathbf{r}_i\times\mathbf{V}$$

$$= \mathbf{L} + \mathbf{R}\times\mathbf{P} - \mathbf{R}\times\mathbf{P} - \mathbf{R}\times\mathbf{P} \qquad (134)$$

that finally yields

$$\mathbf{L} = \mathbf{L}^0 + \mathbf{R}\times\mathbf{P}. \qquad (135)$$

This means that the total angular momentum **L** consists of the internal angular momentum $\mathbf{L}^0$ and the angular momentum $\mathbf{R}\times\mathbf{P}$ corresponding to the motion of the system as the whole. Kinetic energy of a system of particles can be transformed as

$$E_k = \frac{1}{2}\sum_i m_i\left(\mathbf{v}_i^0+\mathbf{V}\right)^2 = \frac{1}{2}\sum_i m_i\mathbf{v}_i^{02} + \frac{M\mathbf{V}^2}{2} + \mathbf{P}^0\cdot\mathbf{V} = E_k^0 + \frac{M\mathbf{V}^2}{2} \qquad (136)$$

since $\mathbf{P}^0 = 0$. Thus also kinetic energy consists of the internal kinetic energy and the kinetic energy corresponding to the motion of the system as the whole.

An isolated system of two interacting masses can be described as one so-called *reduced mass* moving around the CM of the system. Choosing a coordinate system with the origin at the CM, one obtains the constraint on the positions of the bodies

$$m_1\mathbf{r}_1 + m_2\mathbf{r}_2 = 0. \qquad (137)$$

As the single dynamical variable one can choose the position-difference vector

$$\mathbf{r} \equiv \mathbf{r}_1 - \mathbf{r}_2. \qquad (138)$$

Using these two equations, one can express the individual positions as

$$\mathbf{r}_1 = \frac{m_2}{m_1+m_2}\mathbf{r}, \qquad \mathbf{r}_2 = -\frac{m_1}{m_1+m_2}\mathbf{r}. \qquad (139)$$

The total energy of the system

$$E = \frac{m_1\mathbf{v}_1^2}{2} + \frac{m_2\mathbf{v}_2^2}{2} + U\left(|\mathbf{r}_1-\mathbf{r}_2|\right) \qquad (140)$$

becomes

$$E = \frac{m\mathbf{v}^2}{2} + U\left(|\mathbf{r}|\right), \qquad (141)$$

where $\mathbf{v} = \dot{\mathbf{r}}$ and

$$m \equiv \frac{m_1 m_2}{m_1+m_2} \qquad (142)$$

is the reduced mass. The equation of motion of the system can be obtained from either of the two equations for the individual bodies, e.g.,

$$m_1\ddot{\mathbf{r}}_1 = m\ddot{\mathbf{r}} = -\frac{\partial U}{d\mathbf{r}_1} = -\frac{\partial U}{d\mathbf{r}}. \qquad (143)$$

The resulting equation of motion for the reduced mass

$$\partial U \qquad (144)$$

$m\ddot{\mathbf{r}} = -$

∂U ∂U

$d\mathbf{r}$

can be obtained from the equation of motion for the second body as well.

6 One-dimensional conservative motion

For one-dimensional conservative systems conservation of the total energy can be used to reduce the secondorder differential equation to a first-order differential equation that can be straightforwardly integrated. The basic form of a one-dimensional system, a particle, has the energy

$$E = \frac{mv^2}{2} + U(x). \tag{145}$$

In addition, there are systems that become one-dimensional because of a constrait. For instance, a mass on a light rod (pendulum) making a motion along a circle in the *xy* plane can be described by a single dynamical variable, the angle ϕ, that makes it effectively one-dimensional. Solving the above equation for $v = \dot{x}$, one arrives at the first-order ODE

$$\dot{x} = \sqrt{\frac{2\left[E - U(x)\right]}{m}}. \tag{146}$$

Here the integral of motion E plays the role of an integration constant. Further integration of Eq. (146) is done as follows

$$t = \int dt = \sqrt{\frac{m}{2}} \int \frac{dx}{\sqrt{E - U(x)}}. \tag{147}$$

This defines $x(t)$ implicitly. Another integration constant comes from the indefinite integral.

Motion of a mechanical system can take place only in the region $E > U$, where kinetic energy is positive. Regions $E < U$ are inaccessible for the system and are called *barrier regions*. The system hits barriers and changes direction of its motion at *turning points*, where $E = U$ and thus $v = 0$. Positions of turning points depends on the energy. If the motion of the system is limited by left and right turning ponts x_1 and x_2 being the two roots of the equation $E = U(x)$, the particle performs oscillations between these points with the period

$$T(E) = \sqrt{2m} \int_{x_1}^{x_2} \frac{dx}{\sqrt{E - U(x)}}. \tag{148}$$

In general, the period depends on the energy E. If E is close to the minimum of $U(x)$ where the latter is parabolic,

$$U(x) \cong U_0 + \frac{k}{2}\left(x - x_0\right)^2, \tag{149}$$

the system becomes a harmonic oscillator with the period independent of the energy. The motion of the harmonic oscillator has been determined above by solving its equation of motion. Within the present

approach, setting $U_0 = x_0 = 0$, one finds $x_{1,2} = \pm x_m$, where $x_m = \sqrt{2E/k}$. Then one can integrate Eq.

(147) as follows

$$t = \sqrt{\frac{m}{k}} \int \frac{dx}{\sqrt{x_m^2 - x^2}} = \frac{1}{\omega_0} \left(\arcsin \frac{x}{x_m} + \varphi_0 \right), \tag{150}$$

where $\omega_0 \equiv \sqrt{k/m}$ is the oscillator frequency and ϕ_0 is an integration constant. Resolving this formula for x, one finally obtains

$$x(t) = x_m \sin(\omega_0 t - \phi_0), \tag{151}$$

a form of Eq. (40).

Square potential energy near the minimum is a borderline. If $U(x)$ is subsquare (grows slower than x^2), the frequency of oscillations decreases with the amplitude. For a supersquare $U(x)$ oscillation frequency increases with the amplitude. For instance, for the potential energy $U(x) = A|x|^n$ turning points are $x_{1,2} = \pm x_m$, where $x_m = (E/A)^{1/n}$. The period given by Eq. (148) becomes

$$T(E) = 2\sqrt{\frac{2m}{A}} \int_0^{x_m} \frac{dx}{\sqrt{x_m^n - x^n}}. \tag{152}$$

Changing integration variable to $y \equiv x/x_m$, one obtains

$$T(E) = 2\sqrt{\frac{2m}{A}} x_m^{1-n/2} \int_0^1 \frac{dy}{\sqrt{1-y^n}} = 2\sqrt{\frac{2m}{A}} \left(\frac{E}{A}\right)^{1/n-1/2} \int_0^1 \frac{dy}{\sqrt{1-y^n}}, \tag{153}$$

where the integral is just a number. For the frequency one has

$$\omega_0(E) \propto 1/T \propto E(n-2)/(2n) \tag{154}$$

that illustrates the above statement.

Consider a particle in a washboard potential

$$U(x) = U_0\,[1 - \cos(ax)]. \tag{155}$$

Near the minima at $ax = 2\pi n$, $n = 0,\pm1,\pm2$, etc., one has $U(x) \cong kx^2/2$ with $k = U_0 a^2$ and behaves similarly to the harmonic oscillator. The whole problem is mathematically equivalent to that of pendulum. At $ax = \pi + 2\pi n$ potential energy has maxima, $U_{max} = 2U_0$. Let us calculate the period of oscillations, considering the region around the minimum at $x = 0$. Eq. (148) becomes

$$T = 2\sqrt{\frac{2m}{U_0}} \int_0^{x_m} \frac{dx}{\sqrt{\cos(ax) - \cos(ax_m)}}, \tag{156}$$

where x_m is the right turning point satisfying $E = U_0\,[1 - \cos(ax_m)]$. The integral can be transformed into the form similar to those above,

$$T = 2\sqrt{\frac{m}{U_0}} \int_0^{x_m} \frac{dx}{\sqrt{\sin^2(ax_m/2) - \sin^2(ax/2)}}. \tag{157}$$

Next, one can employ the variable change

$$\begin{aligned} \sin(ax/2) &= \sin(ax_m/2)\sin\xi \\ (a/2)\cos(ax/2)dx &= \sin(ax_m/2)\cos\xi d\xi \end{aligned} \tag{158}$$

that yields

$$T = \frac{4}{a}\sqrt{\frac{m}{U_0}} \int_0^{\pi/2} \frac{1}{\cos(ax/2)} \frac{\cos\xi d\xi}{\sqrt{1-\sin^2\xi}} = \frac{4}{a}\sqrt{\frac{m}{U_0}} \int_0^{\pi/2} \frac{d\xi}{\cos(ax/2)}. \tag{159}$$

Now, eliminating x in the integrand, one obtains

$$T = \frac{4}{a}\sqrt{\frac{m}{U_0}} K\left[\sin^2(ax_m/2)\right],$$

where $K(m)$ is the elliptic integral of the first kind,

$$K(m) = \int_0^{\pi/2} \frac{d\xi}{\sqrt{1 - m\sin^2\xi}}. \tag{160}$$

Let us consider limiting cases of this formula. In the case of small oscillations near the minimum $ax_m \ll 1$ and one can use

$$K(m) \cong \int_0^{\pi/2} d\xi \left(1 + \frac{1}{2} m \sin^2\xi\right) = \frac{\pi}{2}\left(1 + \frac{1}{4}m\right). \tag{161}$$

Then the period becomes

$$T = T_0\left[1+\frac{1}{16}\left(ax_m\right)^2+\dots\right], \qquad T_0 = \frac{2\pi}{a}\sqrt{\frac{m}{U_0}} . \qquad (162)$$

18

If turning points approach the maxima of $U(x)$, then $m = \sin^2(ax_m/2) \to 1$. At this point the elliptic integral logarithmically diverges and so does the period. This is physically understandable because the particle is spending a lot of time near turning points when the latter approach the maxima.

If the energy of the particle exceeds the maxima of potential energy, its motion becomes unbounded and unidirectional, so that there is no period any longer. Different regimes can be conveniently represented on the two-dimensional phase space of the particle (x,v). Small oscillations near minima make elliptic trajectories in the phase space. Unbounded motion makes infinite curved trajectories. In the limit of a very high energy trajectories become straight lines $v = \sqrt{2E/m}$ = const. Special trajectories are those corresponding to $E = U_{\max}$ and separate bound and unbound trajectories. They are called separatrices.

7 Systems with constraints and special coordinates

Some mechanical systems include *constraints* that are restricting motion of its parts. Simplest examples are mass on the incline, two masses connected by a light rod and a mathematical pendulum (mass on light rod, the other point of the rod fixed at the pivot point or fulcrum). Constaints that can be eliminated leading to decreasing of the number of dynamical variables of the system (its degrees of freedom) are called *holonomic*. More rare *non-holonomic* constrains cannot be eliminated and they have to be added to the equations of motion via reaction forces. A cylinder rolling on a plane withoud slipping is a system with a holonomic constraint, as the rotation angle of the cylinder can be expressed via displacement of its CM or vice versa. However, a disk or a sphere rolling on a plane without slipping can make a more complicated motion than back and forth, so that the constraint cannot be eliminated and is non-holonomic. Removing friction makes the sphere on a plane a holonomic system while the disc remains non-holonomic. Finally, any rigid body can be considered as a collection of point masses with lots of constraints, althought this point of view is not practically significant. In some important cases constrants can be resolved by choosing a special coordinate system.

7.1 Polar coordinate system; pendulum

Consider, as an example, a mathematical pendulum (simply pendulum), a mass m on a light rod of length l, moving in xy plane. It is convenient to change to polar coordinates with the center at the fulcrum

$$x = r\cos\phi, \qquad y = r\sin\phi, \qquad (163)$$

where the constraint has the simple form $r = l$. For the position vector one obtains the expression

$$\mathbf{r} = l(\mathbf{e}_x\cos\phi + \mathbf{e}_y\sin\phi), \qquad (164)$$

One can see that, under the constraint, the two-component vector $\mathbf{r} = (x,y)$ has been reduced to a single variable ϕ.

The velocity of the pendulum has the form

$$\mathbf{v} = \dot{\mathbf{r}} = l(-\mathbf{e}_x \sin\phi + \mathbf{e}_y \cos\phi)\dot{\phi}. \tag{165}$$

One can check $\mathbf{v} \cdot \mathbf{r} = 0$, that is, both vectors are perpendicular. This means that velocity is tangential with respect to the circular trajectory the pendulum is making. Kinetic enegy reads

$$E_k = \frac{m\mathbf{v}^2}{2} = \frac{ml^2\omega^2}{2}, \tag{166}$$

where

$$\omega \equiv \dot{\phi} \tag{167}$$

is the angular velocity.

Acceleration is given by

$$\dot{\mathbf{v}} = l(-\mathbf{e}_x \sin\phi + \mathbf{e}_y \cos\phi)\ddot{\phi} - l(\mathbf{e}_x \cos\phi + \mathbf{e}_y \sin\phi)\dot{\phi}^2. \tag{168}$$

One can see that the first term in $\dot{\mathbf{v}}$ is tangential and collinear with velocity, while the second term is collinear with $\mathbf{r}$ and directed toward the center (fulcrum). This is centripetal acceleration. Introducing angular velocityone can write centripetal acceleration in the form

$$\mathbf{a}_c = -\omega^2\mathbf{r}. \tag{169}$$

It is convenient to project everything on the (local) direction of the circular trajectory and the perpendicular (radial) direction. For this, one can introduce orthogonal unit vectors

$$\begin{aligned} \mathbf{e}_r &= \mathbf{r}/r = \mathbf{e}_x \cos\phi + \mathbf{e}_y \sin\phi \\ \mathbf{e}_\phi &= -\mathbf{e}_x \sin\phi + \mathbf{e}_y \cos\phi \end{aligned} \tag{170}$$

that satisfy

$$\frac{\partial \mathbf{e}_r}{\partial \varphi} = \mathbf{e}_\varphi, \qquad \frac{\partial \mathbf{e}_\varphi}{\partial \varphi} = -\mathbf{e}_r. \tag{171}$$

Note also the expression for the gradient in the polar coordinate system

$$\frac{\partial}{\partial \mathbf{r}} = \mathbf{e}_r \frac{\partial}{\partial r} + \mathbf{e}_\varphi \frac{1}{r}\frac{\partial}{\partial \varphi}. \tag{172}$$

Eqs. (170) can be used to rewrite the above formulas as

$$\begin{aligned} \mathbf{r} &= l\mathbf{e}_r \quad \mathbf{v} = \\ & l\omega\mathbf{e}_\phi \\ \dot{\mathbf{v}} &= l\dot{\omega}\mathbf{e}_\phi - l\omega^2\mathbf{e}_r, \end{aligned} \tag{173}$$

that is,

$$\begin{aligned} v_r &= 0, \qquad v_\phi = l\omega \\ \dot{v}_r &= -l\omega^2, \qquad \dot{v}_\phi = l\dot{\omega}. \end{aligned} \tag{174}$$

Note that $\dot{v}_r 6= \partial_t v_r = 0$. From Eq. (170) one obtains

Constraint causes an additional reaction force

$$\mathbf{N} = f_N \mathbf{e}_r \tag{175}$$

that helps to keep the body on the trajectory. Let us project the equation of motion

$$m\dot{\mathbf{v}} = \mathbf{F} + \mathbf{N}, \tag{176}$$

where $\mathbf{F}$ is the external force, onto $\mathbf{e}_r$ and $\mathbf{e}_\phi$. One obtains

$$\begin{aligned} m\dot{\mathbf{v}} \cdot \mathbf{e}_r &= -ml\omega^2 = (\mathbf{F} + \mathbf{N}) \cdot \mathbf{e}_r = F_r + f_N \\ m\dot{\mathbf{v}} \cdot \mathbf{e}_\phi &= ml\dot{\omega} = (\mathbf{F} + \mathbf{N}) \cdot \mathbf{e}_\phi = F_\phi. \end{aligned} \tag{177}$$

Here the second equation yields the equation of motion for the pendulum,

$$\dot{\omega} = F_\phi/(ml). \tag{178}$$

The first equation yields the reaction force,

$$f_N = -ml\omega^2 - F_r. \tag{179}$$

The latter may be important in the problem of mechanical stability of a real-life rod but is irrelevant in the idealized consideration of constraints. Thus one can just project the Newton's second law onto the direction of the trajectory enforced by the constraint, that leads to Eq. (178). If $\mathbf{F}$ is a potential force,

$$F_\varphi = \mathbf{F} \cdot \mathbf{e}_\varphi = -\frac{\partial U}{\partial \mathbf{r}} \cdot \mathbf{e}_\varphi. \tag{180}$$

Using Eq. (172), one obtains

$$F_\varphi = -\frac{1}{r}\frac{\partial U}{\partial \varphi} = -\frac{1}{l}\frac{\partial U}{\partial \varphi}. \tag{181}$$

For the gravity force acting on the pendulum,

$$\mathbf{F} = mg\mathbf{e}_x, \qquad U(\mathbf{r}) = -mg\mathbf{r} \cdot \mathbf{e}_x, \tag{182}$$

one obtains

$$F_\phi = \mathbf{F} \cdot \mathbf{e}_\phi = mg\mathbf{e}_x \cdot \mathbf{e}_\phi = -mg \sin\phi. \tag{183}$$

Potential energy expressed via ϕ has the form

$$U(\phi) = -mg\mathbf{r} \cdot \mathbf{e}_x = -mgl(\mathbf{e}_x \cos\phi + \mathbf{e}_y \sin\phi) \cdot \mathbf{e}_x = -mgl\cos\phi, \tag{184}$$

so that the total kinetic energy becomes

$$E = \frac{ml^2\omega^2}{2} - mgl\cos\varphi. \tag{185}$$

From Eq. (181) one obtains

$$F_\varphi = -\frac{1}{l}\frac{\partial U}{\partial \varphi} = -mg\sin\varphi, \tag{186}$$

same as above. This, the equation of motion for the pendulum, Eq. (178), takes the final form

$$\ddot{\varphi} + \omega_0^2 \sin\varphi = 0, \qquad \omega_0 = \sqrt{\frac{g}{l}}. \tag{187}$$

Here ω_0 is the frequency of pendulum's oscillations in the limit of small amplitude, where $\sin\phi \simeq \phi$. Note that because of the constraint the problem has become non-linear.

7.2 Spherical coordinate system; Spherical pendulum

Consider a mass m on a light rod of length a, fixed at a fulcrum at the other side, the system now being able to move on a sphere of radius a. To resolve the constraint, one can choose the spherical coordinate system

$$x = r\sin\theta\cos\phi, \qquad y = r\sin\theta\sin\phi, \qquad z = r\cos\theta, \tag{188}$$

where θ is azimuthal angle and ϕ is polar angle. In this system the constraint has the simple form $r = a$.

One can introduce the local orthogonal frame defined by

$$\begin{aligned} \mathbf{e}_r &= \mathbf{r}/r = \mathbf{e}_x \sin\theta\cos\phi + \mathbf{e}_y \sin\theta\sin\phi + \mathbf{e}_z \cos\theta \\ \mathbf{e}_\theta &= \mathbf{e}_x \cos\theta\cos\phi + \mathbf{e}_y \cos\theta\sin\phi - \mathbf{e}_z \sin\theta \\ \mathbf{e}_\phi &= -\mathbf{e}_x \sin\phi + \mathbf{e}_y \cos\phi. \end{aligned} \tag{189}$$

These vectors satisfy

$$\frac{\partial \mathbf{e}_r}{\partial \theta} = \mathbf{e}_\theta \quad \overline{\partial\varphi}$$

$$\partial\mathbf{e}_r \quad , = \quad \frac{\partial \mathbf{e}_\theta}{\partial \varphi} = \mathbf{e}_\varphi \cos\theta \quad \mathbf{e}_\phi \sin\theta \tag{190}$$

$$\frac{\partial \mathbf{e}_\theta}{\partial \theta} = -\mathbf{e}_r, \tag{191}$$

$$\frac{\partial \mathbf{e}_\phi}{\partial\theta} \qquad \frac{\partial \mathbf{e}_\phi}{\partial\phi} \qquad x \qquad y \qquad r \qquad \theta$$

as well as

$$\mathbf{e}_r \times \mathbf{e}_\theta = \mathbf{e}_\phi \tag{193}$$

plus cyclic permutations. The gradient in the spherical system has the form

$$= \;\; 0, \qquad \_\_\_ = -\mathbf{e} \cos\phi - \mathbf{e} \sin\phi = -\mathbf{e} \sin\theta - \mathbf{e} \cos\theta, \tag{192}$$

$$\frac{\partial}{\partial \mathbf{r}} = \mathbf{e}_r \frac{\partial}{\partial r} + \mathbf{e}_\theta \frac{1}{r}\frac{\partial}{\partial \theta} + \mathbf{e}_\varphi \frac{1}{r\sin\theta}\frac{\partial}{\partial\varphi}. \tag{194}$$

The velocity of the spherical pendulum can be obtained by differentiating the first line of Eq. (189) using Eq. (190),

$$\mathbf{v} = \dot{\mathbf{r}} = a\left(\frac{\partial \mathbf{e}_r}{\partial\theta}\dot\theta + \frac{\partial \mathbf{e}_r}{\partial\varphi}\dot\varphi\right) = a\mathbf{e}_\theta\dot\theta + a\mathbf{e}_\varphi \sin\theta\dot\varphi. \tag{195}$$

As was done for the polar coordinate system above, to obtain the equation of motion in the spherical system, it is sufficient to project components of Newton's second law onto the local direction of the plane, i.e., discarding centripetal terms. In vector form of this projected equation of motion reads

$$m\dot{\mathbf{v}} - \mathbf{F} - ((m\dot{\mathbf{v}} - \mathbf{F}) \cdot \mathbf{e}_r)\mathbf{e}_r = 0 \tag{196}$$

or

$$(m\dot{\mathbf{v}} - \mathbf{F}) \times \mathbf{e}_r = 0. \tag{197}$$

In terms of the spherical components of the vectors this becomes

$$m\dot{v}_\theta = F_\theta, \qquad m\dot{v}_\phi = F_\phi. \tag{198}$$

For projected acceleration one obtains

$$\begin{aligned}
\dot{\mathbf{v}}/a &= \mathbf{e}_\theta\ddot{\theta} + \mathbf{e}_\phi \sin\theta\ddot{\phi} + \mathbf{e}_\phi \cos\theta\dot{\theta}\dot{\phi} + \dot{\mathbf{e}}_\theta\dot{\theta} + \dot{\mathbf{e}}_\phi \sin\theta\dot{\phi} \\
&= \mathbf{e}_\theta\ddot{\theta} + \mathbf{e}_\varphi \sin\theta\ddot{\varphi} + \mathbf{e}_\varphi \cos\theta\dot{\theta}\dot{\varphi} + \left(\frac{\partial \mathbf{e}_\theta}{\partial\theta}\dot{\theta} + \frac{\partial \mathbf{e}_\theta}{\partial\varphi}\dot{\varphi}\right)\dot{\theta} + \left(\frac{\partial \mathbf{e}_\varphi}{\partial\theta}\dot{\theta} + \frac{\partial \mathbf{e}_\varphi}{\partial\varphi}\dot{\varphi}\right)\sin\theta\dot{\varphi} \\
&= \mathbf{e}_\theta\ddot{\theta} + \mathbf{e}_\varphi \sin\theta\ddot{\varphi} + \mathbf{e}_\varphi \cos\theta\dot{\theta}\dot{\varphi} + \mathbf{e}_\varphi \cos\theta\dot{\varphi}\dot{\theta} - \mathbf{e}_\theta \cos\theta\dot{\varphi}\sin\theta\dot{\varphi} \\
&= \mathbf{e}_\theta\ddot{\theta} + \mathbf{e}_\varphi \sin\theta\ddot{\varphi} + \mathbf{e}_\varphi 2\cos\theta\dot{\theta}\dot{\varphi} - \mathbf{e}_\theta \sin\theta\cos\theta\dot{\varphi}^2
\end{aligned} \tag{199}$$

or, finally,

$$\dot{\mathbf{v}}/a = \mathbf{e}_\theta\left(\ddot{\theta} - \sin\theta\cos\theta\dot{\varphi}^2\right) + \mathbf{e}_\varphi\left(\sin\theta\ddot{\varphi} + 2\cos\theta\dot{\theta}\dot{\varphi}\right) \tag{200}$$

with centripetal terms dropped. Thus, the equation of motion of body confined to a sphere is Eqs. (198) in the form

$$\ddot{\theta} - \sin\theta\cos\theta\dot{\varphi}^2 = \frac{1}{ma}F_\theta \tag{201}$$

$$\sin\theta\ddot{\varphi} + 2\cos\theta\dot{\theta}\dot{\varphi} = \frac{1}{ma}F_\varphi \tag{202}$$

that is strongly nonlinear.

To obtain components of the force, one has to express the potential energy via θ and ϕ and use Eq. (194). This yields

$$F_\theta = -\frac{1}{r}\frac{\partial U}{\partial\theta}, \qquad F_\varphi = -\frac{1}{r\sin\theta}\frac{\partial U}{\partial\varphi} \tag{203}$$

with $r \Rightarrow l$. If gravity force is applied,

$$U(\mathbf{r}) = -mg\mathbf{r} \cdot \mathbf{e}_z = -mga\cos\theta, \tag{204}$$

then

$$F_\theta = -mg\sin\theta, \qquad F_\phi = 0. \tag{205}$$

Consider angular momentum $\mathbf{l} = \mathbf{r} \times m\mathbf{v}$ of a body confined to a sphere. Expressing it in the spherical coordinate system and using Eq. (193), one obtains

$$\mathbf{l} = a\mathbf{e}_r \times ma\left(\mathbf{e}_\theta\dot{\theta} + \mathbf{e}_\varphi \sin\theta\dot{\varphi}\right) = ma^2\left(\mathbf{e}_\varphi\dot{\theta} - \mathbf{e}_\theta\sin\theta\dot{\varphi}\right). \tag{206}$$

The torque is defined by

$$\tau = \mathbf{r} \times \mathbf{F} = a\mathbf{e}_r \times (\mathbf{e}_\theta F_\theta + \mathbf{e}_\phi F_\phi) = a\mathbf{e}_\phi F_\theta - a\mathbf{e}_\theta F_\phi. \tag{207}$$

In the case of gravity force directed along z axis, as defined above, the component of the torque $\tau = \mathbf{r} \times \mathbf{F}$ along z axis is zero: $\tau_z = \tau \cdot \mathbf{e}_r = a\mathbf{e}_\phi F_\theta \cdot \mathbf{e}_z = 0$. This means that z component of the angular momentum is conserved. Using Eq. (189), one obtains

$$l_z = \mathbf{l} \cdot \mathbf{e}_z = -ma^2 \sin\theta\dot{\phi}\,\mathbf{e}_\theta \cdot \mathbf{e}_z = ma^2 \sin^2\theta\dot{\phi} = \text{const.} \tag{208}$$

A similar result can be obtained from Eq. (202) in the case $F_\phi = 0$ by multiplying by the integrating factor $\sin\theta$ as follows:

$$\left(\sin\theta\ddot{\varphi} + 2\cos\theta\dot{\theta}\dot{\varphi}\right)\sin\theta = \partial_t \sin^2\theta\dot{\varphi} = 0. \tag{209}$$

Integrating this, one obtains Eq. (208). Now, eliminating $\dot{\phi}$ in Eq. (201), one obtains an autonomous equation of motion for θ,

$$\ddot{\theta} = \frac{l_z^2}{(ma^2)^2}\frac{\cos\theta}{\sin^3\theta} + \frac{F_\theta}{ma}. \tag{210}$$

The total energy of a particle on a sphere can be obtained using Eq. (195),

$$E = \frac{ma^2}{2}\left(\dot{\theta}^2 + \sin^2\theta\dot{\varphi}^2\right) + U(\theta,\varphi). \tag{211}$$

In the axially-symmetric case $U = U(\theta)$ one has $l_z = \text{const}$, and the energy can be expressed as

$$E = \frac{ma^2}{2}\left(\dot{\theta}^2 + \frac{l_z^2}{(ma^2)^2}\frac{1}{\sin^2\theta}\right) + U(\theta). \tag{212}$$

This is an effectively one-dimensional motion that can be integrated after resolving for $\dot{\theta}$. Eq. (212) itself can be obtained as an integral of motion of Eq. (210) multiplying by the integrating factor $\dot{\theta}$, similarly to what was done in the main text. Moreover, one can obtain Eq. (211) by integrating Eqs. (201) and (202).

8 Motion in a central field

Consider motion of a body in a central field, $U = U(r)$, so that $\mathbf{F} = -\nabla U = -(dU/dr)\mathbf{e}_r$ is directed radially. In this case the angular momentum is conserved, $\mathbf{l} = \mathbf{r} \times m\mathbf{v} = \text{const}$, see discussion after Eq. (96). Since both $\mathbf{r}$ and $\mathbf{v}$ are perpendicular to $\mathbf{l}$, the body is moving in the plane perpendicular to $\mathbf{l}$. It is convenient to use polar coordinate system to describe this motion. Using

$$\mathbf{r} = \mathbf{e}_r r, \qquad \mathbf{v} = \mathbf{e}_\phi \dot{\phi} r + \mathbf{e}_r \dot{r} \tag{213}$$

(see Sec. 7.1), one obtains

$$l = mr^2\dot{\phi} = \text{const.} \tag{214}$$

As $r > 0$, one can see that ϕ is changing monotonically and $\dot{\phi}$ is not changing the sign. This equation can be rewritten in terms of the sectorial velocity $\dot{S}$ as

$$\dot{S} \equiv \frac{1}{2}r^2\dot{\varphi} = \frac{l}{2m} = \text{const.} \tag{215}$$

This is Kepler's second law.

Substituting Eq. (214) into the energy, one obtains

$$E = \frac{m\mathbf{v}^2}{2} + U(r) = \frac{m}{2}\left(\dot{r}^2 + r^2\dot{\varphi}^2\right) + U(r) = \frac{m}{2}\left(\dot{r}^2 + \frac{l^2}{m^2r^2}\right) + U(r) \tag{216}$$

or

$$E = \frac{m\dot{r}^2}{2} + U_{\text{eff}}(r), \tag{217}$$

where

$$U_{\text{eff}}(r) = U(r) + \frac{l^2}{2mr^2}. \tag{218}$$

The last term here is the so-called centrifugal energy. Then one can proceed as in the case of one-dimensional motion. First, one resolves the formula above for $\dot{r}$,

$$\dot{r} = \sqrt{\frac{2}{m}\left[E - U_{\text{eff}}(r)\right]}. \tag{219}$$

Integrating this, one obtains the dependence $r(t)$ implicitly,

$$t = \sqrt{\frac{m}{2}} \int \frac{dr}{\sqrt{E - U_{\text{eff}}(r)}}. \tag{220}$$

The trajectory of the body is defined in the polar system by the dependence $r(\phi)$. One can find the inverse function $\phi(r)$ by writing $\dot{\phi} = \partial_r \phi \dot{r}$ and using Eqs. (214) and (219). One obtains

$$\frac{d\phi}{dr} = \frac{\dot{\phi}}{\dot{r}} = \sqrt{\frac{m}{2}} \frac{l/(mr^2)}{\sqrt{E - U_{\text{eff}}(r)}} \tag{221}$$

and, integrating this,

$$\varphi = \int \frac{dr}{r^2} \frac{l}{\sqrt{2m\left[E - U_{\text{eff}}(r)\right]}}. \tag{222}$$

Because of the centrifugal energy, the body cannot fall into the center for typical attractive forces ($U \propto 1/r$), because the positive centrifugal energy is growing faster. As in the case of one-dimensional motion, the body is moving within the classically accessible region $U_{\text{eff}}(r) < E$. If the radial motion is between two turning points, $r_1 < r < r_2$, this is a bound state and the body is orbiting around the center. The state with only one turning point is a *scattering* state. The body is coming from infinity and then goes away after scattering on attracting or repelling center. If one chooses a turning point as the origin for ϕ, the traejectory $r(\phi)$ will be symmetric with respect to this turning point, $r(\phi) = r(-\phi)$.

The period of motion in bound states is defined as time needed for the body to go from one turning point to the other and back

$$T = \sqrt{2m} \int_{r_1}^{r_2} \frac{dr}{\sqrt{E - U_{\mathrm{eff}}(r)}}. \tag{223}$$

The rotation angle corresponding to the period of motion is given by

$$\Delta\varphi = l\sqrt{\frac{2}{m}} \int_{r_1}^{r_2} \frac{dr}{r^2} \frac{1}{\sqrt{E - U_{\mathrm{eff}}(r)}}. \tag{224}$$

Closed orbits correspond to $\Delta\phi = 2\pi m/n$, where m and n are natural numbers. Closed orbit is a special case, since for an arbitrary $U(r)$ the rotating angle $\Delta\phi$ is arbitrary. However, there are two important particular cases, in which orbits are closed. First, this is gravitation, $U \propto 1/r$. Second, this is harmonic oscillator, $U \propto r^2$.

If two turning points are close to each other, $r_1 \cong r_2 \cong R$, then the orbit is nearly circular. In this case, in the integral for $\Delta\phi$ one can approximate $1/r^2 \Rightarrow 1/R^2$. After that the imtegral becomes the same as for the period and one obtains the relation $T = mR^2\Delta\phi/l$. With $\Delta\phi = 2\pi$ and $l = mrv$ one obtains $T = 2\pi R/v$, as it should be. This is the case of rotation of the Earth around the Sun.

8.1 Kepler's problem

Consider the important case $U \propto 1/r$ that corresponds to gravitational interaction between stars and planets, as well as to the Coulomb interaction of charged particles. The effective energy of Eq. (218) has the form

$$U_{\mathrm{eff}}(r) = -\frac{\alpha}{r} + \frac{l^2}{2mr^2}, \qquad \alpha \equiv GMm, \tag{225}$$

where G is gravitational constant, M and m are masses of the two gravitating bodies. If $M \gg m$ (The Sun and the Earth), one can consider the light body m rotating around the heavy body M put in the center. This will be our main case. If the masses are comparable, they will be both moving but the problem can be reduced to a one-body problem using the reduced mass (see Sec. 5). This effective energy has the minimum Ueff(r0) = Emin with

$$r_0 = \frac{l^2}{\alpha m}, \qquad E_{\mathrm{min}} = -\frac{\alpha^2 m}{2l^2}. \tag{226}$$

For $E_{\mathrm{min}} < E < 0$ there are two turning points and the motion is bounded (orbiting). For $E > 0$ there is only one turning point and the motion is unbounded (scattering). The trajectory can be found from Eq. (222). With the new variable $u = 1/r$, $du = -dr/r^2$ one obtains

$$\begin{aligned}
\varphi &= -l \int \frac{du}{\sqrt{2m\left(E + \alpha u - \frac{l^2}{2m}u^2\right)}} = -\int \frac{ds}{\sqrt{\frac{2mE}{l^2} + \frac{2m\alpha}{l^2}u - u^2}} \\
&= -\int \frac{du}{\sqrt{\frac{2mE}{l^2} + \frac{m^2\alpha^2}{l^4} - \left(u - \frac{m\alpha}{l^2}\right)^2}} = -\arccos \frac{u - \frac{m\alpha}{l^2}}{\sqrt{\frac{2mE}{l^2} + \frac{m^2\alpha^2}{l^4}}}.
\end{aligned} \tag{227}$$

This can be rewritten as

$$\frac{r_0}{r} = 1 + \epsilon \cos \varphi, \tag{228}$$

where

$$\epsilon \equiv \sqrt{1 + \frac{E}{|E_{\min}|}} \tag{229}$$

For $E < 0$ this trajectory is an ellipse[25] (Kepler's first law) with *excentricity* and axes

$$a = \frac{r_0}{1-\epsilon^2} = \frac{\alpha}{2|E|}, \qquad b = \frac{r_0}{\sqrt{1-\epsilon^2}} = \frac{l}{\sqrt{2m|E|}}. \tag{230}$$

In the case $E = E_{\min}$ one has = 0 and the ellipse degenerates into a circle $r = r_0 = \frac{l^2}{m\alpha} = \frac{m^2 r^2 v^2}{m\alpha}$. From this one obtains $r = \alpha/(mv^2)$.

The area of the ellipse is

$$S = \pi a b = \frac{\pi \alpha l}{\sqrt{(2|E|)^3 m}}. \tag{231}$$

This area is being covered during the period of motion. With the help of Eq. (215) one obtains $S = ST^{\cdot}$ and

$$T = \frac{S}{\dot{S}} = \frac{\pi \alpha l}{\sqrt{(2|E|)^3 m}} \frac{2m}{l} = \pi \alpha \sqrt{\frac{m}{2|E|^3}}, \tag{232}$$

depending only on the energy and diverging for $E \to 0$. One can express the period via the linear size of the orbit a using Eq. (230). The result is

$$T = 2\pi \sqrt{\frac{m}{\alpha}} a^{3/2}. \tag{233}$$

The relation $T^2 \propto a^3$ is Kepler's third law.

Time dependence of the motion can be conveniently represented in parametric form. For this, one can rewrite Eq. (220) as

$$t = \sqrt{\frac{m}{2|E|}} \int \frac{r dr}{\sqrt{-r^2 + \frac{\alpha}{|E|} r - \frac{l^2}{2m|E|}}} = \sqrt{\frac{ma}{\alpha}} \int \frac{r dr}{\sqrt{-(r-a)^2 + a^2 \epsilon^2}}. \tag{234}$$

With the substitution

$$r = a(1 - \epsilon \cos \xi) \tag{235}$$

the integral can be calculated as

$$t = \sqrt{\frac{ma}{\alpha}} \int \frac{a(1-\epsilon \cos \xi) a \epsilon \sin \xi d\xi}{\sqrt{a^2 \epsilon^2 (1 - \cos^2 \xi)}} = \sqrt{\frac{ma^3}{\alpha}} \int (1 - \epsilon \cos \xi) d\xi = \sqrt{\frac{m}{\alpha}} a^{3/2} (\xi - \epsilon \sin \xi). \tag{236}$$

These two formulas provide the dependence $r(t)$ parametrically. Period of motion corresponds to changing of ξ by 2π: r returns to the same value while t increases by T given by Eq. (233).

The time dependence of the angle ϕ is now defined by Eq. (228). Moreover, one can express Descartes coordinates (x,y) parametrically, too. First, one finds

$$x = r\cos\varphi = \frac{r_0 - r}{\epsilon} = \frac{a\left(1-\epsilon^2\right) - a\left(1-\epsilon\cos\xi\right)}{\epsilon} = a\left(\cos\xi - \epsilon\right). \tag{237}$$

Then one finds

$$y = \sqrt{r^2 - x^2} = a\sqrt{1-\epsilon^2}\sin\xi. \tag{238}$$

Cases of unbound motion, $E > 0$, and repelling potential $U(r) = \alpha/r$ (unbound motion for any energy) can be considered in a similar way.

8.2 Scattering by a central field

Consider an unbound motion of a body or a particle in a central field. The particle is coming from infinity, experiences the action of the central force, and then goes to infinity again, however, having changed its direction of motion. Change of direction of a body as the result of interaction with other bodies is called *scattering*. As was mentioned above, the traejectory is symmetric, $r(\phi) = r(-\phi)$, if one chooses the turning point r_1 as the origin of ϕ. Thus the change of direction (deflection angle) χ as the result of scattering can be obtained from Eq. (222) as

$$\chi = \pi - 2\phi_0, \tag{239}$$

where

$$\varphi_0 = \int_{r_1}^{\infty} \frac{dr}{r^2} \frac{l}{\sqrt{2m\left[E - U_{\text{eff}}(r)\right]}} \tag{240}$$

is half of the angle between the initial and final parts of the trajectory. In the following it is convenient, instead of integrals of motion E and l, use the speed at infinity v_∞ and the target distance ρ. The latter is the minimal distance from the center, corresponding to the straight trajectory in the absence of the central force. Using

$$E = \frac{mv_\infty^2}{2}, \qquad l = \rho m v_\infty, \tag{241}$$

one can rewrite Eq. (240) in the form

$$\chi = \pi - 2\int_{r_1}^{\infty} \frac{dr}{r^2} \frac{\rho}{\sqrt{1 - \frac{2U_{\text{eff}}(r)}{mv_\infty^2}}}, \qquad \frac{2U_{\text{eff}}(r)}{mv_\infty^2} = \frac{\rho^2}{r^2} + \frac{2U(r)}{mv_\infty^2}. \tag{242}$$

In experiments there is usually a beam of particles with different ρ that are being scattered. It is convenient to introduce the differential scattering cross-section $d\sigma$ defined by

$$d\sigma(\chi) = dN/n, \tag{243}$$

where dN is the number of particles scattered during a unit of time within the angular interval $d\chi$ around χ and n is the number of particles crossing a unit of area of the beam during a unit of time. It is assumed that n is uniform. One can see that $d\sigma$ has the unit of area. Since $\chi = \chi(\rho)$, particles scattered within $d\chi$ are those within target distance interval $d\rho = |d\rho/d\chi|d\chi$ that corresponds to $d\chi$. The number of such particles scattered during a unit of time is $dN = 2\pi\rho d\rho n$. Thus finally one obtains

$$d\sigma(\chi) = 2\pi\rho(\chi)\left|\frac{d\rho}{d\chi}\right| d\chi, \tag{244}$$

where $\rho(\chi)$ and $d\rho/d\chi$ can be found from Eq. (242).

If the particle is being scattered on another particle that is at rest before the collision, one can change into in the center-of-mass frame and consider scattering of a particle with the reduced mass on the force center located at the CM, using the formulas above. After that one has to change back into the laboratory frame that results in redefinition of angles and transformation of the results that requires some algebra.

Consider scattering of a charged particle by the Coulomb field, $U(r) = \alpha/r$, as in celestial mechanics. Calculating the integral in Eq. (242) as was done in Eq. (227), one arrives at

$$\chi = \pi - 2\arccos\frac{1}{\sqrt{1+\rho^2/\rho_0^2}}, \tag{245}$$

where

$$\rho_0 \equiv \frac{\alpha}{mv_\infty^2} \tag{246}$$

is the characteristic distance. Resolving for ρ, one obtains

$$\rho^2 = \rho_0^2 \tan^2\left(\frac{\pi-\chi}{2}\right) = \rho_0^2 \cot^2\frac{\chi}{2}.$$

Now Eq. (244) yields

$$d\sigma = \pi\rho_0^2 \frac{\cos(\chi/2)}{\sin^3(\chi/2)} d\chi. \tag{247}$$

Instead of scattering within $d\chi$, one can consider scattering within the infinitesimal body angle

$$d\Omega = 2\pi \sin\chi d\chi. \tag{248}$$

This yields Rutherford formula

$$d\sigma = \left(\frac{\alpha}{2mv_\infty^2}\right)^2 \frac{d\Omega}{\sin^4(\chi/2)}. \tag{249}$$

In some cases one can define total scattering cross-section

$$\sigma = \int d\sigma = \int d\chi \frac{d\sigma}{d\chi}. \tag{250}$$

With the help of Eq. (244) it can be written as

$$\sigma = 2\pi \hat{}\ \rho d\rho. \tag{251}$$

The image behind this formula is a scattering center in form of a circle of some finite radius a and the area $\sigma = \pi a^2$. If the target distance ρ satisfies $\rho < a$, the particle will hit the target and be scattered. Otherwise it will not be scattered at all. For all interactions that decrease gradually at infinity, particles with *all* ρ will be scattered. Correspondingly, the integral in Eq. (250) diverges at small χ. Only if the interaction has a cut-off, the total scattering cross-section is finite. The simplest example is elastic scattering on a rigid sphere of radius a. Here one immediately finds

$$\rho = a\sin\varphi_0 = a\sin\frac{\pi-\chi}{2} = a\cos\frac{\chi}{2}. \tag{252}$$

Substitution into Eq. (250) yields the differential cross-section

$$d\sigma = \frac{\pi a^2}{2} \sin\chi d\chi. \tag{253}$$

Integrating over χ yields

$$\sigma = \int d\sigma = \frac{\pi a^2}{2} \int_0^\pi \sin\chi d\chi = \pi a^2, \tag{254}$$

as it should be.

One can also define total scattering cross-section for specific events, for instance, for falling onto the attracting center $U = -\alpha/r^2$. In this case the effective potential energy of Eq. (218) reads

$$U_{\rm eff}(r) = -\frac{\alpha}{r^2} + \frac{l^2}{2mr^2} = \left(-\alpha + \frac{l^2}{2m}\right)\frac{1}{r^2}. \tag{255}$$

Absorbed will be all particles for which the coefficient in $U_{\rm eff}(r)$ is negative,

$$\frac{l^2}{2m} = \frac{(\rho m v_\infty)^2}{2m} < \alpha, \tag{256}$$

otherwise particles will escape to infinity. In terms of the target distance this condition has the form

$$\rho^2 < \rho_0^2 \equiv \frac{2\alpha}{m v_\infty^2}. \tag{257}$$

Total scattering scoss-section then becomes

$$\sigma = \pi\rho_0^2 = \frac{2\pi\alpha}{m v_\infty^2}. \tag{258}$$

Although in this case one can find σ easily, finding the differential cross-section requires a more serious calculation.

8.3 Small-angle scattering

If target distance ρ or particle's speed are large, the trajectory is nearly a straight line, say, along x axis. Passing the force center, the particle acquires a perpendicular momentum p_y as the result. Now the small scattering angle χ can be found as

$$\chi \cong \frac{p_y}{m v_\infty} \ll 1. \tag{259}$$

The momentum p_y can be found as

$$p_y = \int_{-\infty}^{\infty} dt F_y = -\int_{-\infty}^{\infty} dt \frac{\partial U}{\partial y} = -\int_{-\infty}^{\infty} dt \frac{dU}{dr}\frac{\partial r}{\partial y} = -\int_{-\infty}^{\infty} dt \frac{dU}{dr}\frac{y}{r}. \tag{260}$$

Considering, at the lowest order in the perturbation, the motion along x as undisturbed, one can use $dt = dx/v_\infty$ and also $y = \rho$. This yields

$$p_y = -\frac{\rho}{v_\infty} \int_{-\infty}^{\infty} \frac{dx}{r} \frac{dU}{dr}. \tag{261}$$

Changing to integration over *r* with the use of

$$x = \sqrt{r^2 - \rho^2}, \qquad dx = \frac{r dr}{\sqrt{r^2 - \rho^2}}, \tag{262}$$

one obtains

$$p_y = -\frac{2\rho}{v_\infty} \int_{\rho}^{\infty} \frac{dU}{dr} \frac{dr}{\sqrt{r^2 - \rho^2}} \tag{263}$$

and, finally, the deflection angle

$$\chi = -\frac{2\rho}{m v_\infty^2} \int_{\rho}^{\infty} \frac{dU}{dr} \frac{dr}{\sqrt{r^2 - \rho^2}}. \tag{264}$$

Finding differential cross-section $d\sigma$ requires inverting the above formula that requires knowing the explicit form of $U(r)$.

The method used above is beautiful and instructive. However, the final integral formula is hardly simpler than Eqs. (239) and (246). The really new result is that for the arbitrary potential energy $U(x,y)$ (we set $z = 0$), obtained similarly,

$$\chi = -\frac{1}{m v_\infty^2} \int_{-\infty}^{\infty} dx \left. \frac{\partial U(x,y)}{\partial y} \right|_{y=\rho}. \tag{265}$$

Part II

Lagrangian Mechanics

9 Frame transformations and conservation laws

10 Constraints in the Lagrange formalism

11 Examples

Part III

Mechanics of Rigid Bodies

Part IV

Hamiltonian Mechanics

Part V Additional Topics

12 Perturbative methods

13 Parametric resonance

14 Dynamical chaos

15 Microscopic damping

Chapter 8

Electromagnetic waves

David Morin, morin@physics.harvard.edu

The waves we've dealt with so far in this book have been fairly easy to visualize. Waves involving springs/masses, strings, and air molecules are things we can apply our intuition to. But we'll now switch gears and talk about electromagnetic waves. These are harder to get a handle on, for a number of reasons. First, the things that are oscillating are electric and magnetic fields, which are much harder to see (which is an ironic statement, considering that we see with light, which is an electromagnetic wave). Second, the fields can have components in various directions, and there can be relative phases between these components (this will be important when we discuss polarization). And third, unlike all the other waves we've dealt with, electromagnetic waves don't need a medium to propagate in. They work just fine in vacuum. In the late 1800's, it was generally assumed that electromagnetic waves required a medium, and this hypothesized medium was called the "ether." However, no one was ever able to observe the ether. And for good reason, because it doesn't exist.

This chapter is a bit long. The outline is as follows. In Section 8.1 we talk about waves in an extended LC circuit, which is basically what a coaxial cable is. We find that the system supports waves, and that these waves travel at the speed of light. This section serves as motivation for the fact that light is an electromagnetic wave. In Section 8.2 we show how the wave equation for electromagnetic waves follows from Maxwell's equations. Maxwell's equations govern all of electricity and magnetism, so it is no surprise that they yield the wave equation. In Section 8.3 we see how Maxwell's equations constrain the form of the waves. There is more information contained in Maxwell's equations than there is in the wave equation. In Section 8.4 we talk about the energy contained in an electromagnetic wave, and in particular the energy flow which is described by the *Poynting vetor*. In Section 8.5 we talk about the momentum of an electromagnetic wave. We saw in Section 4.4 that the waves we've discussed so far carry energy but not momentum. Electromagnetic waves carry both.[1] In Section 8.6 we discuss polarization, which deals with the relative phases of the different components of the electric (and magnetic) field. In Section 8.7 we show how an electromagnetic wave can be produced by an oscillating (and hence accelerating) charge. Finally, in Section 8.8 we discuss the reflection and transmission that occurs when

---

[1] Technically, all waves carry momentum, but this momentum is suppressed by a factor of v/c, where v is the speed of the wave and c is the speed of light. This follows from the relativity fact that energy is equivalent to mass. So a flow of energy implies a flow of mass, which in turn implies nonzero momentum. However, the factor of v/c causes the momentum to be negligible unless we're dealing with relativistic speeds.

an electromagnetic wave encounters the boundary between two different regions, such as air

and glass. We deal with both normal and non-normal angles of incidence. The latter is a bit more involved due to the effects of polarization.

8.1 Cable waves

Before getting into Maxwell's equations and the wave equation for light, let's do a warmup example and study the electromagnetic waves that propagate down a coaxial cable. This example should help convince you that light is in fact an electromagnetic wave.

To get a handle on the coaxial cable, let's first look at the idealized circuit shown in Fig. 1. All the inductors are L, and all the capacitors are C. There are no resistors in the circuit. With the charges, currents, and voltages labeled as shown, we have three facts:

1. The charge on a capacitor is $Q = CV \Longrightarrow q_n = CV_n$.
2. The voltage across an inductor is $V = L(dI/dt) \Longrightarrow V_{n-1} - V_n = L(dI_n/dt)$.
3. Conservation of charge gives $I_n - I_{n+1} = dq_n/dt$.

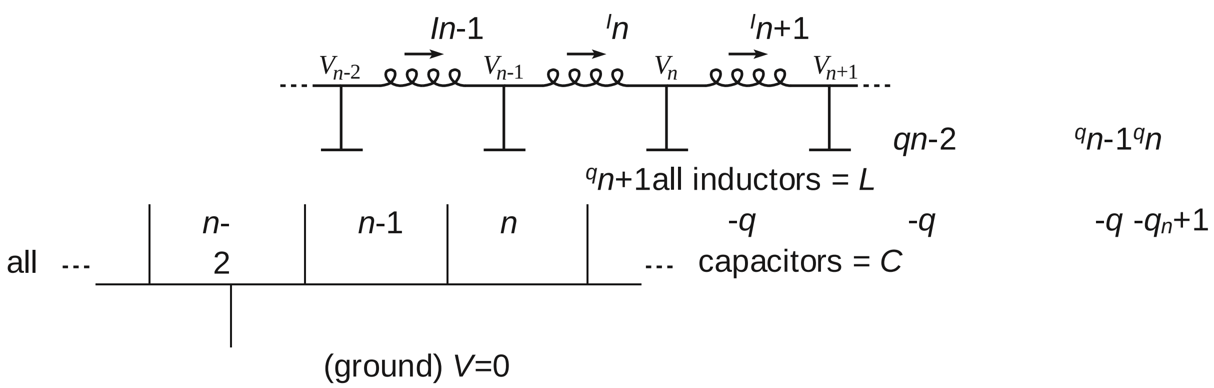


Figure 1

Our goal is to produce an equation, which will end up being a wave equation, for one of the three variables, q, I, and V (the wave equations for all of them will turn out to be the same). Let's eliminate q and I, in favor of V. We could manipulate the above equations in their present form in terms of discrete quantities, and then take the continuum limit (see Problem [to be added]). But it is much simpler to first take the continuum limit and then do the manipulation. If we let the grid size in Fig. 1 be Δx, then by using the definition of the derivative, the above three facts become

$$\begin{aligned} q &= CV, \\ -\Delta x \frac{\partial V}{\partial x} &= L\frac{\partial I}{\partial t} \\ -\Delta x \frac{\partial I}{\partial x} &= \frac{\partial q}{\partial t}\,. \end{aligned} \qquad (1)$$

,

Substituting $q = CV$ from the first equation into the third, and defining the inductance and capacitance per unit length as $L_0 \equiv L/\Delta x$ and $C_0 \equiv C/\Delta x$, the last two equations become

$$-\frac{\partial V}{\partial x} = L_0\frac{\partial I}{\partial t}, \quad \text{and} \quad -\frac{\partial I}{\partial x} = C_0\frac{\partial V}{\partial t}\,. \qquad (2)$$

If we take $\partial/\partial x$ of the first of these equations and $\partial/\partial t$ of the second, and then equate the results for $\partial^2 I/\partial x\partial t$, we obtain

$1\ \partial^2 V$ $-L$___ ___

$$\boxed{\frac{\partial^2 V(x,t)}{\partial t^2} = \frac{1}{L_0 C_0}\frac{\partial^2 V(x,t)}{\partial x^2}}$$

$\partial^2 V$ $\partial x2 = -C0$ ___

$\partial t2$ $\Longrightarrow$ (3)

0

8.2. THE WAVE EQUATION

This is the desired wave equation, and it happens to be dispersionless. We can quickly read off the speed of the waves, which is

$$v = \frac{1}{\sqrt{L_0 C_0}}\,. \qquad (4)$$

If we were to subdivide the circuit in Fig. 1 into smaller and smaller cells, L and C would depend on Δx (and would go to zero as $\Delta x \to 0$), so it makes sense to work with the quantities L_0 and C_0. This is especially true in the case of the actual cable we'll discuss below, for which the choice of Δx is arbitrary. L_0 and C_0 are the meaningful quantities that are determined by the nature of the cable.

Note that since the first fact above says that $q \propto V$, the exact same wave equation holds for q. Furthermore, if we had eliminated V instead of I in Eq. (2) by taking $\partial/\partial t$ of the first equation and $\partial/\partial x$ of the second, we would have obtained the same wave equation for I, too. So V, q, and I all satisfy the same wave equation.

Figure 2

Let's now look at an actual coaxial cable. Consider a conducting wire inside a conducting cylinder, with vacuum in the region between them, as shown in Fig.2. Assume that the wire is somehow constrained to be in the middle of the cylinder. (In reality, the inbetween region is filled with an insulator which keeps the wire in place, but let's keep things simple here with a vacuum.) The cable has an inductance L_0 per unit length, in the same way that two parallel wires have a mutual inductance per unit length. (The cylinder can be considered to be made up of a large number wires parallel to its axis.) It also has a capacitance C_0 per unit length, because a charge difference between the wire and the cylinder will create a voltage difference.

It can be shown that (see Problem [to be added], although it's perfectly fine to just accept this)

$$L_0 = \frac{\mu_0}{2\pi}\ln(r_2/r_1) \quad \text{and} \quad C_0 = \frac{2\pi\epsilon_0}{\ln(r_2/r_1)}, \tag{5}$$

where r_2 is the radius of the cylinder, and r_1 is the radius of the wire. The two physical constants in these equations are the *permeability of free space*, μ_0, and the *permittivity of free space*, ϵ_0. Their values are (μ takes on this value by definition)

$$\mu_0 = 4\pi \cdot 10^{-7}\ \text{H/m}, \quad \text{and} \quad \epsilon_0 \approx 8.85 \cdot 10^{-12}\ \text{F/m}. \tag{6}$$

H and F are the Henry and Farad units of inductance and capacitance. Using Eq. (5), the wave speed in Eq. (4) equals

$$v = \sqrt{\frac{1}{L_0 C_0}} = \sqrt{\frac{1}{\mu_0 \epsilon_0}} \approx \sqrt{\frac{1}{(4\pi \cdot 10^{-7}\ \text{H/m})(8.85 \cdot 10^{-12}\ \text{F/m})}} \approx 3 \cdot 10^{8}\ \text{m/s}. \tag{7}$$

This is the speed of light! We see that the voltage (and charge, and current) wave that travels down the cable travels at the speed of light. And because there are electric and magnetic fields in the cable (due to the capacitance and inductance), these fields also undergo wave motion. Since the waves of these fields travel with the same speed as the original voltage wave, it is a good bet that electromagnetic waves have something to do with light. The reasoning here is that there probably aren't too many things in the world that travel with the speed of light. So if we find something that travels with this speed, then it's probably light (loosely speaking, at least; it need not be in the visible range). Let's now be rigorous and show from scratch that all electromagnetic waves travel at the speed of light (in vacuum).

8.2 The wave equation

By "from scratch" we mean by starting with Maxwell's equations. We have to start somewhere, and Maxwell's equations govern all of (classical) electricity and magnetism. There are four of these equations, although when Maxwell first wrote them down, there were 22 of them. But they were gradually rewritten in a more compact form over the years. Maxwell's equations in vacuum in SI units are (in perhaps overly-general form):

Differential form	Integrated form
$\nabla \cdot \mathbf{E} = \frac{\rho_E}{\epsilon_0}$	$\int \mathbf{E} \cdot d\mathbf{A} = \frac{Q_E}{\epsilon_0}$
$\nabla \cdot \mathbf{B} = \rho_B$	$\int \mathbf{B} \cdot d\mathbf{A} = Q_B$
$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} + \mathbf{J}_B$	$\int \mathbf{E} \cdot d\mathbf{l} = -\frac{d\Phi_B}{dt} + I_B$
$\nabla \times \mathbf{B} = \mu_0\epsilon_0 \frac{\partial \mathbf{E}}{\partial t} + \mu_0 \mathbf{J}_E$	$\int \mathbf{B} \cdot d\mathbf{l} = \mu_0\epsilon_0 \frac{d\Phi_E}{dt} + \mu_0 I_E$

Table 1

If you erase the μ_0's and ${}^{2}{}_{0}$'s here (which arise from the arbitrary definitions of the various units), then these equations are symmetric in **E** and **B**, except for a couple minus signs. The ρ's are the electric and (hypothetical) magnetic charge densities, and the **J**'s are the current densities. The Q's are the charges enclosed by the surfaces that define the d**A** integrals, the Φ's are the field fluxes through the loops that define the d**l** integrals, and the I's are the currents through these loops.

No one has ever found an isolated magnetic charge (a magnetic monopole), and there are various theoretical considerations that suggest (but do not yet prove) that magnetic monopoles can't exist, at least in our universe. So we'll set ρ_B, $\mathbf{J}_B$, and I_B equal to zero from here on. This will make Maxwell's equations appear non-symmetrical, but we'll soon be setting the analogous electric quantities equal to zero too, since we'll be dealing with vacuum. So in the end, the equations for our purposes will be symmetric (except for the μ_0, the ${}^{2}{}_{0}$, and a minus sign). Maxwell's equations with no magnetic charges (or currents) are:

Differential form		Integrated form	Known as
$\nabla \cdot \mathbf{E} = \frac{\rho_E}{\epsilon_0}$	$\int \mathbf{E} \cdot d\mathbf{A} = \frac{Q}{\epsilon_0}$	E	
		Gauss' Law Z	
$\nabla \cdot \mathbf{B} = 0$		$\mathbf{B} \cdot d\mathbf{A} = 0$	No magnetic monopoles
$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$		Z $d\Phi^B$	
$\nabla \times \mathbf{B} = \mu_0\epsilon_0 \frac{\partial \mathbf{E}}{\partial t} + \mu_0$	Z $\mathbf{B} \cdot d\mathbf{l} = \mu_0{}^{2}{}_{0} \frac{d\Phi_E}{dt} + \mu_0 I$	$\mathbf{E} \cdot d\mathbf{l} = -$ dt	Faraday's Law
		$\mathbf{J}_{EE}$	Ampere's Law

Table 2

The last of these, Ampere's Law, includes the so-called "displacement current," $d\Phi_E/dt$.

Our goal is to derive the wave equation for the **E** and **B** fields in vacuum. Since there are no charges of any kind in vacuum, we'll set ρ_E and $\mathbf{J}_E = 0$ from here on. And we'll only need the differential form of the equations, which are now

$$\nabla \cdot \mathbf{E} = 0, \quad (8)$$

$$\nabla \cdot \mathbf{B} = 0, \quad (9)$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \quad (10)$$

$$\nabla \times \mathbf{B} = \mu_0\epsilon_0 \frac{\partial \mathbf{E}}{\partial t}. \quad (11)$$

These equations are symmetric in **E** and **B** except for the factor of $\mu_0{}^{2}{}_{0}$ and a minus sign. Let's eliminate **B** in favor of **E** and see what we get. If we take the curl of Eq. (10) and

8.2. THE WAVE EQUATION

then use Eq. (11) to get ride of **B**, we obtain

$$\nabla \times (\nabla \times \mathbf{E}) = -\nabla \times \frac{\partial \mathbf{B}}{}$$

$$
\begin{aligned}
& \overline{\partial t} \\
= \quad & -\frac{\partial(\nabla \times \mathbf{B})}{\partial t} \\
= \quad & -\frac{\partial}{\partial t}\left(\mu_0\epsilon_0\frac{\partial \mathbf{E}}{\partial t}\right) \\
= \quad & -\mu_0\epsilon_0\frac{\partial^2 \mathbf{E}}{\partial t^2}\,. \qquad (12)
\end{aligned}
$$

On the left side, we can use the handy “BAC-CAB” formula,

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B}). \qquad (13)$$

See Problem [to be added] for a derivation of this. This formula holds even if we have differential operators (such as ∇) instead of normal vectors, but we have to be careful to keep the ordering of the letters the same (this is evident if you go through the calculation in Problem [to be added]). Since both **A** and **B** are equal to ∇ in the present application, the ordering of **A** and **B** in the $\mathbf{B}(\mathbf{A}\cdot\mathbf{C})$ term doesn't matter. But the $\mathbf{C}(\mathbf{A}\cdot\mathbf{B})$ must correctly be written as $(\mathbf{A} \cdot \mathbf{B})\mathbf{C}$. The lefthand side of Eq. (12) then becomes

$$
\begin{aligned}
\nabla \times (\nabla \times \mathbf{E}) &= \nabla(\nabla \cdot \mathbf{E}) - (\nabla \cdot \nabla)\mathbf{E} \\
&= 0 - \nabla^2\mathbf{E}, \qquad (14)
\end{aligned}
$$

where the zero follows from Eq. (8). Plugging this into Eq. (12) finally gives

$$-\nabla^2\mathbf{E} = -\mu_0{}^2{}_0\,\frac{\partial^2\mathbf{E}}{\partial t^2} \Longrightarrow \quad \boxed{\frac{\partial^2 \mathbf{E}}{\partial t^2} = \frac{1}{\mu_0{}^2{}_0}\nabla^2\mathbf{E}} \quad (\textit{wave equation}) \qquad (15)$$

Note that we didn't need to use the second of Maxwell's equations to derive this.

In the above derivation, we could have instead eliminated **E** in favor of **B**. The same steps hold; the minus signs end up canceling again, as you should check, and the first equation is now not needed. So we end up with exactly the same wave equation for **B**:

$$\boxed{\frac{\partial^2 \mathbf{B}}{\partial t^2} = \frac{1}{\mu_0{}^2{}_0}\nabla^2\mathbf{B}} \quad (\textit{wave equation}) \qquad (16)$$

The speed of the waves (both **E** and **B**) is given by the square root of the coefficient on the righthand side of the wave equation. (This isn't completely obvious, since we're now working in three dimensions instead of one, but we'll justify this in Section 8.3.1 below.) The speed is therefore

$$c = \frac{1}{\sqrt{\mu_0\epsilon_0}} \approx 3 \cdot 10^8\ \mathrm{m/s}\,. \qquad (17)$$

This agrees with the result in Eq. (7). But we now see that we don't need a cable to support the propagation of electromagnetic waves. They can propagate just fine in vacuum! This is a fundamentally new feature, because every wave we've studied so far in this book (longitudinal spring/mass waves, transverse waves on a string, longitudinal sound waves, etc.), needs a medium to propagate in/on. But not so with electromagnetic waves.

Eq. (15), and likewise Eq. (16), is a vector equation. So it is actually shorthand for three separate equations for each of the components:

$$\frac{\partial^2 E_x}{\partial t^2} = \frac{1}{\mu_0\epsilon_0}\left(\frac{\partial^2 E_x}{\partial x^2} + \frac{\partial^2 E_x}{\partial y^2} + \frac{\partial^2 E_x}{\partial z^2}\right), \qquad (18)$$

and likewise for E_y and E_z. Each component undergoes wave motion. As far as the wave equation in Eq. (15) is concerned, the waves for the three components are completely independent. Their amplitudes, frequencies, and phases need not have anything to do with each other. *However*, there is more information contained in Maxwell's equations than in the wave equation. The latter follows from the former, but not the other way around. There is no reason why one equation that follows from four equations (or actually just three of them) should contain as much information as the original four. In fact, it is highly unlikely. And as we will see in Section 8.3, Maxwell's equations do indeed further constrain the form of the waves. In other words, although the wave equation in Eq. (15) gives us information about the electric-field wave, it doesn't give us *all* the information.

Index of refraction

In a dielectric (equivalently, an insulator), the vacuum quantities μ_0 and ϵ_0 in Maxwell's equations are replaced by new values, μ and ϵ. (We'll give some justification of this below, but see Sections 10.11 and 11.10 in Purcell's book for the full treatment.) Our derivation of the wave equation for electromagnetic waves in a dielectric proceeds in exactly the same way as for the vacuum case above, except with $\mu_0 \rightarrow \mu$ and $\epsilon_0 \rightarrow \epsilon$. We therefore end up with a wave velocity equal to

$$v = \frac{1}{\sqrt{\mu\epsilon}}. \tag{19}$$

The *index of refraction*, n, of a dielectric is defined by $v \equiv c/n$, where c is the speed of light in vacuum. We therefore have

$$v = \frac{c}{n} \quad \Longrightarrow \quad n = \frac{c}{v} = \sqrt{\frac{\mu\epsilon}{\mu_0\epsilon_0}}. \tag{20}$$

Since it happens to be the case that that $\mu \approx \mu_0$ for most dielectrics, we have the approximate result that

$$n \approx \sqrt{\frac{\epsilon}{\epsilon_0}} \qquad \text{(if } \mu \approx \mu_0\text{)}. \tag{21}$$

And since we must always have $v \le c$, this implies $n \ge 1 \Longrightarrow \epsilon \ge \epsilon_0$.

Strictly speaking, Maxwell's equations with μ_0 and ϵ_0 work in any medium. But the point is that if we don't have a vacuum, then induced charges and currents may arise. In particular there are two types of charges. There are so-called *free* charges, which are additional charges that we can plop down in a material. This is normally what we think of when we think of charge. (The term "free" is probably not the best term, because the charges need not be free to move. We can bolt them down if we wish.) But additionally, there are *bound* charges. These are the effective charges that get produced when polar molecules align themselves in certain ways to "shield" the bound charges.

For example, if we place a positive free charge q_{free} in a material, then the nearby polar molecules will align themselves so that their negative ends form a negative layer around the free charge. The net charge inside a Gaussian surface around the charge is therefore

less than q. Call it q_{net}. Maxwell's first equation is then $\nabla \cdot \mathbf{E} = \rho_{\text{net}}/{}^{2}{}_{0}$. However, it is generally much easier to deal with ρ_{free} than ρ_{net}, so let's define ^{2} by $\rho_{\text{net}}/\rho_{\text{free}} \equiv {}^{2}{}_{0}/{}^{2} < 1$. [2]
Maxwell's first equation can then be written as

$$\nabla \cdot \mathbf{E} = \frac{\rho_{\text{free}}}{\epsilon} \quad . \tag{22}$$

[2] The fact that the shielding is always proportional to q_{free} (at least in non-extreme cases) implies that there is a unique value of ^{2} that works for all values of q_{free}.

The electric field in the material around the point charge is less than what it would be in vacuum, by a factor of ϵ_0/ϵ (and ϵ is always greater than or equal to ϵ_0, because it isn't possible to have "anti-shielding"). In a dielectric, the fact that ϵ is greater than ϵ_0 is consistent with the fact that the index of refraction n in Eq. (21) is always greater than 1, which in turn is consistent with the fact that v is always less than c.

A similar occurrence happens with currents. There can be *free* currents, which are the normal ones we think about. But there can also be *bound* currents, which arise from tiny current loops of electrons spinning around within their atoms. This is a little harder to visualize than the case with the charges, but let's just accept here that the fourth of Maxwell's equations becomes $\nabla \times \mathbf{B} = \mu\epsilon\partial\mathbf{E}/\partial t + \mu\mathbf{J}_{\text{free}}$. But as mentioned above, μ is generally close to μ_0 for most dielectrics, so this distinction usually isn't so important.

To sum up, we can ignore all the details about what's going on at the atomic level by pretending that we have a vacuum with modified μ and ϵ values. Although there certainly exist *bound* charges and currents in the material, we can sweep them under the rug and consider only the *free* charges and currents, by using the modified μ and ϵ values.

The above modified expressions for Maxwell's equations are correct if we're dealing with a single medium. But if we have two or more mediums, the correct way to write the equations is to multiply the first equation by ϵ and divide the fourth equation by μ (see Problem [to be added] for an explanation of this). The collection of all four Maxwell's equations is then

$$
\begin{aligned}
\nabla \cdot \mathbf{D} &= \rho_{\text{free}}, \\
\nabla \cdot \mathbf{B} &= 0, \\
\nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t}, \\
\nabla \times \mathbf{H} &= \frac{\partial \mathbf{D}}{\partial t} + \mathbf{J}_{\text{free}},
\end{aligned}
\tag{23}
$$

where $\mathbf{D} \equiv \epsilon\mathbf{E}$ and $\mathbf{H} \equiv \mathbf{B}/\mu$. $\mathbf{D}$ is called the *electric displacement* vector, and $\mathbf{H}$ goes by various names, including simply the "magnetic field." But you can avoid confusing it with $\mathbf{B}$ if you use the letter $\mathbf{H}$ and not the name "magnetic field."

8.3 The form of the waves

8.3.1 The wavevector k

What is the dispersion relation associated with the wave equation in Eq. (15)? That is, what is the relation between the frequency and wavenumber? Or more precisely, what is the dispersion relation for each component of $\mathbf{E}$, for example the E_x that satisfies Eq. (18)? All of the components are in general functions of four coordinates: the three spatial coordinates x,y,z, and the time t. So by the same reasoning as in the two-coordinate case we discussed at the end of Section 4.1, we know that we can Fourier-decompose the function $E_x(x,y,z,t)$ into exponentials of the form,

$$Ae^{i(k_x x+k_y y+k_z z-\omega t)} \equiv Ae^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}, \quad \text{where } \mathbf{k} \equiv (k_x,k_y,k_z). \qquad (24)$$

Likewise for E_y and E_z. And likewise for the three components of **B**. These are traveling waves, although we can form combinations of them to produce standing waves.

k is known as the *wavevector*. As we'll see below, the magnitude $k \equiv |\mathbf{k}|$ plays exactly the same role that k played in the 1-D case. That is, k is the wavenumber. It equals 2π times the number of wavelengths that fit into a unit length. So $k = 2\pi/\lambda$. We'll also see below that the direction of **k** is the direction of the propagation of the wave. In the 1-D case, the wave had no choice but to propagate in the $\pm\mathbf{x}$ direction. But now it can propagate in any direction in 3-D space.

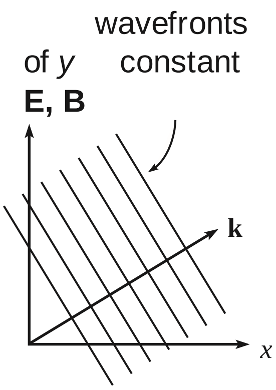


Figure 3

Plugging the exponential solution in Eq. (24) into Eq. (18) gives

$$-\omega^2 = \frac{1}{\mu_0\epsilon_0}(-k_x^2 - k_y^2 - k_z^2) \implies \omega^2 = \frac{|\mathbf{k}|^2}{\mu_0\epsilon_0} \implies \boxed{\omega = c|\mathbf{k}|} \qquad (25)$$

where $c = 1/\sqrt{\mu_0\epsilon_0}$, and where we are using the convention that ω is positive. Eq. (25) is the desired dispersion relation. It is a trivial relation, in the sense that electromagnetic waves in vacuum are dispersionless.

When we go through the same procedure for the other components of **E** and **B**, the "A" coefficient in Eq. (24) can be different for the $2 \cdot 3 = 6$ different components of the fields. And technically **k** and ω can be different for the six components too (as long as they satisfy the same dispersion relation). However, although we would have solutions to the six different waves equations, we wouldn't have solutions to Maxwell's equations. This is one of the cases where the extra information contained in Maxwell's equations is important. You can verify (see Problem [to be added]) that if you want Maxwell's equations to hold for *all* **r** and t, then **k** and ω must be the same for all six components of **E** and **B**. If we then collect the various "A" components into the two vectors $\mathbf{E}_0$ and $\mathbf{B}_0$ (which are constants, independent of **r** and t), we can write the six components of **E** and **B** in vector form as

$$\mathbf{E} = \mathbf{E}_0 e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}, \quad \text{and} \quad \mathbf{B} = \mathbf{B}_0 e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}, \qquad (26)$$

where the **k** vector and the ω frequency are the same in both fields. The vectors $\mathbf{E}_0$ and $\mathbf{B}_0$ can be complex. If they do have an imaginary part, it will produce a phase in the cosine function when we take the real part of the above exponentials. This will be important when we discuss polarization.

From Eq. (26), we see that **E** (and likewise **B**) depends on **r** through the dot product $\mathbf{k}\cdot\mathbf{r}$. So **E** has the same value everywhere on the surface defined by $\mathbf{k}\cdot\mathbf{r} = C$, where C is some constant. This surface is a plane that is perpendicular to **k**. This follows from the fact that if $\mathbf{r}_1$ and $\mathbf{r}_2$ are two points on the surface, then $\mathbf{k}\cdot(\mathbf{r}_1 - \mathbf{r}_2) = C - C = 0$. Therefore, the vector $\mathbf{r}_1 - \mathbf{r}_2$ is perpendicular to **k**. Since this holds for any $\mathbf{r}_1$ and $\mathbf{r}_2$ on the surface, the surface must be a plane perpendicular to **k**. If we suppress the z dependence of **E** and **B** for the sake of drawing a picture on a page, then for a given wavevector **k**, Fig.3 shows some "wavefronts" with common phases $\mathbf{k}\cdot\mathbf{r} - \omega t$, and hence common values of **E** and **B**. The planes perpendicular to **k** in the 3-D case become lines perpendicular to **k** in the 2-D case. Every point on a given plane is equivalent, as far as **E** and **B** are concerned.

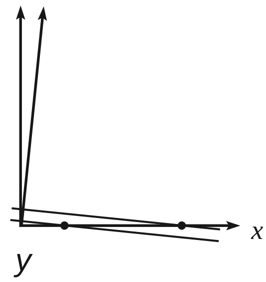
How do these wavefronts move as time goes by? Well, they must always be perpendicular to **k**, so all they can do is move in the direction of **k**. How fast do they move? The dot product $\mathbf{k}\cdot\mathbf{r}$ equals $kr\cos\theta$, where θ is the angle between

k and a given position **r**, and where $k \equiv |\mathbf{k}|$ and $r = |\mathbf{r}|$. If we group the product as $k(r\cos\theta)$, we see that it equals k

$r\cos\theta$ simply equals the x^0 value of the position. So the phase $\mathbf{k}\cdot\mathbf{r}-\omega t$ equals $kx^0 - \omega t$. We have therefore reduced the problem to a 1-D problem (at least as far as the phase is concerned), so we can carry over all of our 1-D results. In particular, the phase velocity (and group velocity too, since the wave equation

√

in Eq. (15) is dispersionless) is $v = \omega/k$, which we see from Eq. (25) equals $c = 1/\ \mu_0\epsilon_0$.



k which a point with constant phase moves in the x direction, with y and z held constant. This follows from the fact that if we let the constant y and z values be y_0 and z_0, then the phase equals $k_x x + k_y y_0 + k_z z_0 - \omega t = k_x x - \omega t + C$, where C is a constant. So we effectively have a 1-D problem for which the phase velocity is ω/k_x.

But note that this velocity can be made arbitrarily large, or even infinite if $k_x = 0$. Fig.4shows a situation where **k** points mainly in the y direction, so k_x is small. Two wavefronts are shown, and they move upward along the direction of **k**. In the time during which the lower wavefront moves to the position of the higher one, a point on the x axis with a particular constant

moves from one dot to the other. This means that it is moving very fast (much faster than the wavefronts), consistent with the fact the ω/k is very large if k is very small. In the limit where phase

Figure 4

x x

times the projection of **r** along **k**. If we rotate our coordinate system so that a new x^0 axis points in the **k** direction, then the projection

Remark: We just found that the phase velocity has magnitude

$$v = \frac{\omega}{k} \equiv \frac{\omega}{|\mathbf{k}|} = \frac{\omega}{\sqrt{k_x^2 + k_y^2 + k_z^2}}, \qquad (27)$$

and it points in the $\hat{\mathbf{k}}$ direction. You might wonder if the simpler expression ω/k_x has any meaning. And likewise for y and z. It does, but it isn't a terribly useful quantity. It is the velocity at the wavefronts are horizontal ($k_x = 0$), a point of constant phase moves infinitely fast along the x axis. The quantities ω/k_x, ω/k_y, and ω/k_z therefore *cannot* be thought of as components of the phase velocity in the phase advances by only a small amount, compared with along the **k** vector. The phase needs the entire distance between the two dots to increase by 2π along the x axis, whereas it needs only the distance between the wavefronts to increase by 2π along the **k** vector. ♣

Eq. (27). The component of a vector should be smaller than the vector itself, after all.

The vector that does correctly break up into components is the wavevector $\mathbf{k} = (k_x, k_y, k_z)$. Its magnitude $k = k_x^2 + k_y^2 + k_z^2$ represents how much the phase of the wave increases in each unit distance along the $\hat{\mathbf{k}}$ direction. (In other words, it equals 2π times the number of wavelengths that fit into a unit distance.) And k_x represents how much the phase of the wave increases in each unit distance along the x direction. This is less than k, as it should be, in view of Fig. 4. For a given distance along the x axis,

8.3.2 Further constraints due to Maxwell's equations

Fig. 3 tells us only that points along a given line have common values of **E** and **B**. It doesn't tell us what these values actually are, or if they are constrained in other ways. For all we know, **E** and **B** on a particular wavefront might look like the vectors shown in Fig. 5 (we have ignored any possible z components). But it turns out the these vectors aren't actually possible. Although they satisfy the wave equation, they don't satisfy Maxwell's equations. So let's now see how Maxwell's equations further constrain the form of the waves. Later on in Section 8.8, we'll see that the waves are even further constrained by any boundary conditions that might exist. We'll look at Maxwell's equations in order and see what each of them implies.

- Using the expression for **E** in Eq. (26), the first of Maxwell's equations, Eq. (8), gives

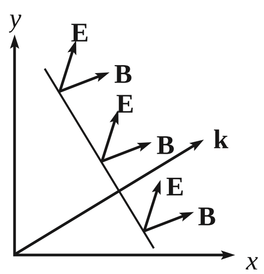
$$\nabla \cdot \mathbf{E} = 0 \implies \frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z} = 0$$
$$\implies ik_xE_x + ik_yE_y + ik_zE_z = 0$$
$$\implies \boxed{\mathbf{k} \cdot \mathbf{E} = 0} \quad (28)$$

  This says that **E** is always perpendicular to **k**. As we see from the second line here, each partial derivative simply turns into a factor if ik_x, etc.

- The second of Maxwell's equations, Eq. (9), gives the analogous result for **B**, namely,

$$\boxed{\mathbf{k} \cdot \mathbf{B} = 0} \quad (29)$$

  So **B** is also perpendicular to **k**.



(impossible **E, B** vectors)

Figure 5

- Again using the expression for **E** in Eq. (26), the third of Maxwell's equations, Eq. (10), gives

$$\nabla \times \ldots \mathbf{B} \Longrightarrow \boxed{\mathbf{k} \times \mathbf{E} = \omega\mathbf{B}} \quad (30)$$

Since the cross product of two vectors is perpendicular to each of them, this result says that **B** is perpendicular to **E**. And we already know that **B** is perpendicular to **k**, from the second of Maxwell's equations. But technically we didn't need to use that equation, because the **B** ⊥ **k** result is also contained in this **k** × **E** = ω**B** result. Note that as above with the divergences, each partial derivative in the curl simply turns into a factor if ik_x, etc.

We know from the first of Maxwell's equations that **E** is perpendicular to **k**, so the magnitude of **k**×**E** is simply |**k**||**E**| ≡ kE. The magnitude of the **k**×**E** = ω**B** relation then tells us that

$$kE = \omega B \quad \Longrightarrow \quad E = \frac{\omega}{k}B \quad \Longrightarrow \quad \boxed{E = cB} \quad (31)$$

Therefore, the magnitudes of **E** and **B** are related by a factor of the wave speed, c. Eq. (31) is very useful, but its validity is limited to a single traveling wave, because the derivation of Eq. (30) assumed a unique **k** vector. If we form the sum of two waves with different **k** vectors, then the sum doesn't satisfy Eq. (30) for any particular vector **k**. There isn't a unique **k** vector associated with the wave. Likewise for Eqs. (28) and (29).

- The fourth of Maxwell's equations, Eq. (11), can be written as ∇×**B** = $(1/c^2)\partial\mathbf{E}/\partial t$, so the same procedure as above yields

$$\boxed{\mathbf{k} \times \mathbf{B} = -\frac{\omega}{c^2}\mathbf{E}} \quad \Longrightarrow \quad B = \frac{\omega}{kc^2}E \quad \Longrightarrow \quad \boxed{B = \frac{E}{c}} \quad (32)$$

This doesn't tell us anything new, because we already know that **E**, **B**, and **k** are all mutually perpendicular, and also that $E = cB$. In retrospect, the first and third (or alternatively the second and fourth) of Maxwell's equations are sufficient to derive all of the above results, which can be summarized as

$$\boxed{\mathbf{E} \perp \mathbf{k}, \quad \mathbf{B} \perp \mathbf{k}, \quad \mathbf{E} \perp \mathbf{B}, \quad E = cB} \quad (33)$$

If three vectors are mutually perpendicular, there are two possibilities for how they are oriented. With the conventions of **E**, **B**, and **k** that we have used in Maxwell's equations and in the exponential solution in Eq. (24) (where the **k**·**r** term comes in with a plus sign), the orientation is such that **E**, **B**, and **k** form a "righthanded" triplet. That is, **E**×**B** points in the same direction as **k** (assuming, of course, that you're defining the cross product with the righthand rule!). You can show that this follows from the **k** × **E** = ω**B** relation in Eq. (30) by either simply drawing three vectors that satisfy Eq. (30), or by using the determinant definition of the cross product to show that a cyclic permutation of the vectors maintains the sign of the cross product.

A snapshot (for an arbitrary value of t) of a possible electromagnetic wave is shown in Fig. 6. We have chosen **k** to point along the z axis, and we have drawn the field only for points on the z axis. But for a given value z_0, all points of the form (x,y,z_0), which is a plane perpendicular to the z axis, have common values of **E** and **B**. **E** points in the $\pm x$ direction, and **B** points in the $\pm y$ direction. As time goes by, the whole figure simply slides along the z axis at speed c. Note that **E** and **B** reach their maximum and minimum values at the same locations. We will find below that this isn't the case for standing waves.

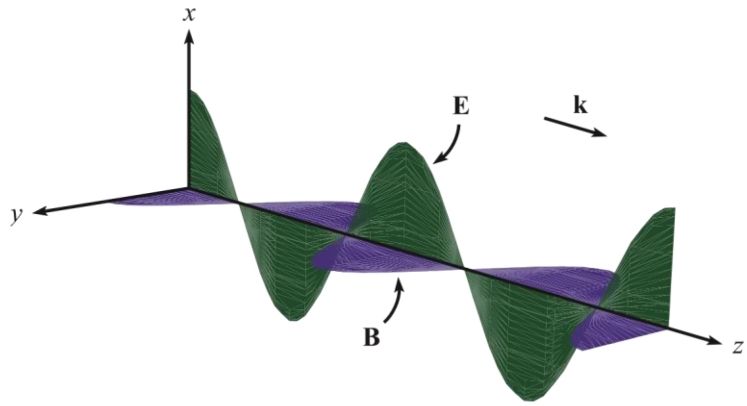


Figure 6

What are the mathematical expressions for the **E** and **B** fields in Fig. 6? We've chosen **k** to point along the z axis, so we have $\mathbf{k} = k\hat{\mathbf{z}}$, which gives $\mathbf{k} \cdot \mathbf{r} = kz$. And since **E** points in the x direction, its amplitude takes the form of $E_0 e^{i\varphi}\hat{\mathbf{x}}$. (The coefficient can be complex, and we have written it as a magnitude times a phase.) This then implies that **B** points in the y direction (as drawn), because it must be perpendicular to both **E** and **k**. So its amplitude takes the form of $B_0 e^{i\varphi}\hat{\mathbf{y}} = (E_0 e^{i\varphi}/c)\hat{\mathbf{y}}$. This is the same phase φ, due to Eq. (30) and the fact that **k** is real, at least for simple traveling waves. The desired expressions for **E** and **B** are obtained by taking the real part of Eq. (26), so we arrive at

$$\mathbf{E} = \hat{\mathbf{x}} E_0 \cos(kz - \omega t + \varphi), \quad \text{and} \quad \mathbf{B} = \hat{\mathbf{y}} \frac{E_0}{c} \cos(kz - \omega t + \phi), \qquad (34)$$

These two vectors are in phase with each other, consistent with Fig. 6. And **E**, **B**, and **k** form a righthanded triple of vectors, as required.

Remarks:

1. When we talk about polarization in Section 8.6, we will see that **E** and **B** don't have to point in specific directions, as they do in Fig. 6, where **E** points only along $\hat{\mathbf{x}}$ and **B** points only along $\hat{\mathbf{y}}$. Fig. 6 happens to show the special case of "linear polarization."
2. The **E** and **B** waves don't have to be sinusoidal, of course. Because the wave equation is linear, we can build up other solutions from sinusoidal ones. And because the wave equation is dispersionless, we know (as we saw at the end of Section 2.4) that any function of the form $f(z - vt)$, or equivalently $f(kz - \omega t)$, satisfies the wave equation. But

the restrictions placed by Maxwell's equations still hold. In particular, the **E** field determines the **B** field.

3. A static solution, where **E** and **B** are constant, can technically be thought of as a sinusoidal solution in the limit where $\omega = k = 0$. In vacuum, we can always add on a constant field to **E** or **B**, and it won't affect Maxwell's equations (and therefore the wave equation either), because all of the terms in Maxwell's equations in vacuum involve derivatives (either space or time). But we'll ignore any such fields, because they're boring for the purposes we'll be concerned with. ♣

8.3.3 Standing waves

Let's combine two waves (with equal amplitudes) traveling in opposite directions, to form a standing wave. If we add $\mathbf{E}_1 = \hat{\mathbf{x}}E_0\cos(kz - \omega t)$ and $\mathbf{E}_2 = \hat{\mathbf{x}}E_0\cos(-kz - \omega t)$, we obtain

$$\mathbf{E} = \hat{\mathbf{x}}(2E_0)\cos kz\cos\omega t. \quad (35)$$

This is indeed a standing wave, because all z values have the same phase with respect to time.

There are various ways to find the associated **B** wave. Actually, there are (at least) two right ways and one wrong way. The wrong way is to use the result in Eq. (30) to say that $\omega\mathbf{B} = \mathbf{k} \times \mathbf{E}$. This would yield the result that **B** is proportional to $\cos kz\cos\omega t$, which we will find below is incorrect. The error (as we mentioned above after Eq. (31)) is that there isn't a unique **k** vector associated with the wave in Eq. (35), because it is generated by two waves with opposite **k** vectors. If we insisted on using Eq. (30), we'd be hard pressed to decide if we wanted to use $k\hat{\mathbf{z}}$ or $-k\hat{\mathbf{z}}$ as the **k** vector.

A valid method for finding **B** is the following. We can find the traveling $\mathbf{B}_1$ and $\mathbf{B}_2$ waves associated with each of the traveling $\mathbf{E}_1$ and $\mathbf{E}_2$ waves, and then add them. You can quickly show (using $\mathbf{B} = (1/\omega)\mathbf{k} \times \mathbf{E}$ for each traveling wave separately) that $\mathbf{B}_1 = \hat{\mathbf{y}}(E_0/c)\cos(kz - \omega t)$ and $\mathbf{B}_2 = -\hat{\mathbf{y}}(E_0/c)\cos(-kz - \omega t)$. The sum of these waves give the desired associated **B** field,

$$\mathbf{B} = \hat{\mathbf{y}}(2E_0/c)\sin kz\sin\omega t. \quad (36)$$

Another method is to use the third of Maxwell's equations, Eq. (10), which says that $\nabla \times \mathbf{E} = -\partial\mathbf{B}/\partial t$. Maxwell's equations hold for *any* **E** and **B** fields. We don't have to worry about the uniqueness of **k** here. Using the **E** in Eq. (35), the cross product $\nabla \times \mathbf{E}$ can be calculated with the determinant:

$$\nabla \times \mathbf{E} = \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ E_x & 0 & 0 \end{vmatrix} = \hat{\mathbf{y}}\frac{\partial E_x}{\partial z} - \hat{\mathbf{z}}\frac{\partial E_x}{\partial y} = -\hat{\mathbf{y}}(2E_0)k\sin kz\cos\omega t - 0. \quad (37)$$

Eq. (30) tells us that this must equal $-\partial\mathbf{B}/\partial t$, so we conclude that

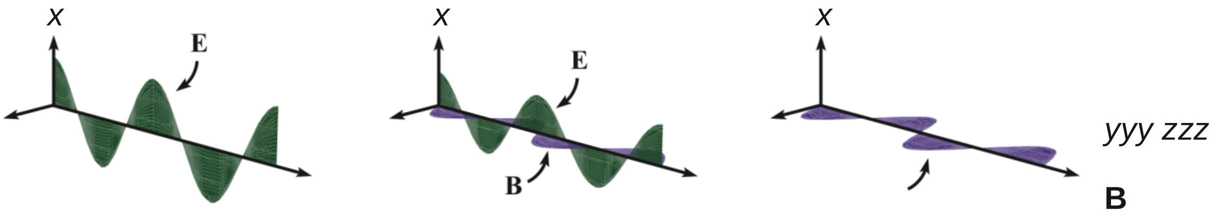
$$\mathbf{B} = \hat{\mathbf{y}}(2E_0)(k/\omega)\sin kz\sin\omega t = \hat{\mathbf{y}}(2E_0/c)\sin kz\sin\omega t. \quad (38)$$

in agreement with Eq. (36). We have ignored any possible additive constant in **B**.

Having derived the associated **B** field in two different ways, we can look at what we've found. **E** and **B** are still perpendicular to each other, which makes since, because **E** is the superposition of two vectors that point in the $\pm\hat{\mathbf{x}}$ direction, and **E** is the superposition of two vectors that point in the $\pm\hat{\mathbf{y}}$ direction. But there is a major difference between standing waves and traveling waves. In traveling waves, **E** and **B** run along in step with each other, as shown above in Fig. 6. They reach their maximum and minimum values at the same times and positions. However, in standing waves **E** is maximum *when* **B** is zero, and also *where* **B** is zero (and vice versa). **E** and **B** are 90° out of phase with each other in both time and space. That is, the **B** in Eq. (36) can be written as

$$\mathbf{B} = \hat{\mathbf{y}}\frac{2E_0}{c}\cos\left(kz - \frac{\pi}{2}\right)\cos\left(\omega t - \frac{\pi}{2}\right), \tag{39}$$

which you can compare with the **E** in Eq. (35). A few snapshots of the **E** and **B** waves are shown in Fig. 7.



(t=0) (ωt=π/4) (ωt=π/2)

Figure 7

8.4 Energy

8.4.1 The Poynting vector

The energy density of an electromagnetic field is

$$\mathcal{E} = \frac{\epsilon_0}{2}E^2 + \frac{1}{2\mu_0}B^2, \tag{40}$$

where $E \equiv |\mathbf{E}|$ and $B \equiv |\mathbf{B}|$ are the magnitudes of the fields at a given location in space and time. We have suppressed the (x,y,z,t) arguments of E, E, and B. This energy density can be derived in various ways (see Problem [to be added]), but we'll just accept it here. The goal of this section is to calculate the rate of change of E, and to then write it in a form that allows us to determine the energy flux (the flow of energy across a given surface). We will find that the energy flux is given by the so-called *Poynting vector*.

If we write E^2 and B^2 as $\mathbf{E} \cdot \mathbf{E}$ and $\mathbf{B} \cdot \mathbf{B}$, then the rate of change of E becomes

$$\frac{\partial \mathcal{E}}{\partial t} = \epsilon_0 \mathbf{E} \cdot \frac{\partial \mathbf{E}}{\partial t} + \frac{1}{\mu_0}\mathbf{B} \cdot \frac{\partial \mathbf{B}}{\partial t}. \tag{41}$$

(The product rule works here for the dot product of vectors for the same reason it works for a regular product. You can verify this by explicitly writing out the components.) The third and fourth Maxwell's equations turn this into

$$\begin{aligned} \frac{\partial \mathcal{E}}{\partial t} &= \epsilon_0 \mathbf{E} \cdot \left(\frac{1}{\mu_0 \epsilon_0}\nabla \times \mathbf{B}\right) + \frac{1}{\mu_0}\mathbf{B} \cdot (-\nabla \times \mathbf{E}) \\ &= \frac{1}{\mu_0}\Big(\mathbf{E} \cdot (\nabla \times \mathbf{B}) - \mathbf{B} \cdot (\nabla \times \mathbf{E})\Big). \end{aligned} \tag{42}$$

The righthand side of this expression conveniently has the same form as the righthand side of the vector identity (see Problem [to be added] for the derivation),

$$\nabla \cdot (\mathbf{C} \times \mathbf{D}) = \mathbf{D} \cdot (\nabla \times \mathbf{C}) - \mathbf{C} \cdot (\nabla \times \mathbf{D}). \tag{43}$$

So we now have

$$\frac{\partial \mathcal{E}}{\partial t} = \frac{1}{\mu_0}\nabla \cdot (\mathbf{B} \times \mathbf{E}). \qquad (44)$$

Now consider a given volume V in space. Integrating Eq. (44) over this volume V yields

$$\int_V \frac{\partial \mathcal{E}}{\partial t} = \frac{1}{\mu_0}\int_V \nabla \cdot (\mathbf{B} \times \mathbf{E}) \quad \Longrightarrow \quad \frac{\partial W_V}{\partial t} = \frac{1}{\mu_0}\int_A (\mathbf{B} \times \mathbf{E}) \cdot d\mathbf{A}, \qquad (45)$$

where W_V is the energy contained in the volume V (we've run out of forms of the letter E), and where we have used the divergence theorem to rewrite the volume integral as a surface integral over the area enclosing the volume. $d\mathbf{A}$ is defined to be the vector perpendicular to the surface (with the positive direction defined to be outward), with a magnitude equal to the area of a little patch.

Let's now make a slight change in notation. $d\mathbf{A}$ is defined to be an outward-pointing vector, but let's define $d\mathbf{A}_{\text{in}}$ to be the inward-pointing vector, $d\mathbf{A}_{\text{in}} \equiv -d\mathbf{A}$. Eq. (45) can then be written as (switching the order of **E** and **B**)

$$\frac{\partial W_V}{\partial t} = \frac{1}{\mu_0}\int_A (\mathbf{E} \times \mathbf{B}) \cdot d\mathbf{A}_{\text{in}}. \qquad (46)$$

We can therefore interpret the vector,

$$\boxed{\mathbf{S} \equiv \frac{1}{\mu_0}\mathbf{E} \times \mathbf{B}} \quad (\textit{energy flux} : \text{energy/(area} \cdot \text{time)}) \qquad (47)$$

as giving the flux of energy *into* a region. This vector **S** is known as the *Poynting vector*. And since $\mathbf{E}\times\mathbf{B} \propto \mathbf{k}$, the Poynting vector points in the same direction as the velocity of the wave. Integrating **S** over any surface (or rather, just the component perpendicular to the surface, due to the dot product with $d\mathbf{A}_{\text{in}}$) gives the energy flow across the surface. This result holds for any kind of wave – traveling, standing, or whatever. Comparing the units on both sides of Eq. (46), we see that the Poynting vector has units of energy per area per time. So if we multiply it (or its perpendicular component) by an area, we get the energy per time crossing the area.

The Poynting vector falls into a wonderful class of phonetically accurate theorems/results. Others are the Low energy theorem (named after S.Y. Low) dealing with low-energy photons, and the Schwarzschild radius of a black hole (kind of like a shield).

8.4.2 Traveling waves

Let's look at the energy density $\mathcal{E}$ and the Poynting vector **S** for a traveling wave. A traveling wave has $B = E/c$, so the energy density is (using $c^2 = 1/\mu_0\epsilon_0$ in the last step)

$$\mathcal{E} = \frac{\epsilon_0}{2}E^2 + \frac{1}{2\mu_0}B^2 = \frac{\epsilon_0}{2}E^2 + \frac{1}{2\mu_0}\frac{E^2}{c^2} \Longrightarrow \boxed{\mathcal{E} = \epsilon_0 E^2} \qquad (48)$$

We have suppressed the (x,y,z,t) arguments of $\mathcal{E}$ and E. Note that this result holds only for traveling waves. A standing wave, for example, doesn't have $B = E/c$ anywhere, so $\mathcal{E}$ doesn't take this form. We'll discuss standing waves below. The Poynting vector for a traveling wave is

$$\mathbf{S} = \frac{1}{\mu_0}\mathbf{E}\times\mathbf{B} = \frac{1}{\mu_0}E\left(\frac{E}{c}\right)\hat{\mathbf{k}}, \tag{49}$$

where we have used the facts that $\mathbf{E} \perp \mathbf{B}$ and that their cross product points in the direction of $\mathbf{k}$. Using $1/\mu_0 = c^2{}^2{}_0$, arrive at

$$\boxed{\mathbf{S} = c^2{}_0}\boxed{E}^2\boxed{\hat{\mathbf{k}}} = c\mathrm{E}\hat{\mathbf{k}}. \tag{50}$$

This last equality makes sense, because the energy density E moves along with the wave, which moves at speed c. So the energy per unit area per unit time that crosses a surface

is cE (which you can verify has the correct units). Eq. (50) is true at all points (x,y,z,t) individually, and not just in an average sense. We'll derive formulas for the averages below. In the case of a sinusoidal traveling wave of the form,

$$\mathbf{E} = \hat{\mathbf{x}}E_0\cos(kz - \omega t) \quad \text{and} \quad \mathbf{B} = \hat{\mathbf{y}}(E_0/c)\cos(kz - \omega t), \tag{51}$$

the above expressions for E and $\mathbf{S}$ yield

$$\mathrm{E} = {}^2{}_0E_0^2\cos^2(kz - \omega t) \quad \text{and} \quad \mathbf{S} = c^2{}_0E_0^2\cos^2(kz - \omega t)\hat{\mathbf{k}}. \tag{52}$$

Since the average value of $\cos^2(kz - \omega t)$ over one period (in either space or time) is 1/2, we see that the average values of E and $|\mathbf{S}|$ are

$$\mathcal{E}_{\text{avg}} = \frac{1}{2}\epsilon_0E_0^2 \quad \text{and} \quad |\mathbf{S}|_{\text{avg}} = \frac{1}{2}c\epsilon_0E_0^2. \tag{53}$$

$|\mathbf{S}|_{\text{avg}}$ is known as the *intensity* of the wave. It is the average amount of energy per unit area per unit time that passes through (or hits) a surface. For example, at the location of the earth, the radiation from the sun has an intensity of 1360 Watts/m^2. The energy comes from traveling waves with many different frequencies, and the total intensity is just the sum of the intensities of the individual waves (see Problem [to be added]).

8.4.3 Standing waves

Consider the standing wave in Eqs. (35) and (36). With $2E_0$ defined to be A, this wave becomes

$$\mathbf{E} = \hat{\mathbf{x}}A\cos kz\cos\omega t, \quad \text{and} \quad \mathbf{B} = \hat{\mathbf{y}}(A/c)\sin kz\sin\omega t. \tag{54}$$

The energy density in Eq. (48) for a traveling wave isn't valid here, because it assumed $B = E/c$. Using the original expression in Eq. (40), the energy density for the above standing wave is (using $1/c^2 = \mu_0{}^2{}_0$)

$$\mathcal{E} = \frac{\epsilon_0}{2}E^2 + \frac{1}{2\mu_0}B^2 = \frac{1}{2}\epsilon_0 A^2\big(\cos^2 kz\cos^2\omega t + \sin^2 kz\sin^2\omega t\big). \tag{55}$$

If we take the average over a full cycle in time (a full wavelength in space would work just as well), then the $\cos^2 \omega t$ and $\sin^2 \omega t$ factors turn into 1/2's, so the time average of E is

$$\mathcal{E}_{\text{avg}} = \frac{1}{2}\epsilon_0 A^2\left(\frac{1}{2}\cos^2 kz + \frac{1}{2}\sin^2 kz\right) = \frac{\epsilon_0 A^2}{4}, \tag{56}$$

which is independent of z. It makes sense that it doesn't depend on z, because a traveling wave has a uniform average energy density, and a standing wave is just the sum of two traveling waves moving in opposite directions.

The Poynting vector for our standing wave is given by Eq. (47) as (again, the travelingwave result in Eq. (50) isn't valid here):

$$\mathbf{S} = \frac{1}{\mu_0}\mathbf{E}\times\mathbf{B} = \frac{A^2}{\mu_0 c}\hat{\mathbf{k}}\cos kz\sin kz\cos\omega t\sin\omega t. \tag{57}$$

At any given value of z, the time average of this is zero (because $\cos\omega t\sin\omega t = (1/2)\sin 2\omega t$), so there is no net energy flow in a standing wave. This makes sense, because a standing wave is made up of two traveling waves moving in opposite directions which therefore have opposite energy flows (on average). Similarly, for a given value of t, the spatial average is zero. Energy sloshes back and forth between points, but there is no net flow.

Due to the fact that a standing wave is made up of two traveling waves moving in opposite directions, you might think that the Poynting vector (that is, the energy flow) should be *identically* equal to zero, for all z and t. But it isn't, because each of the two Poynting vectors depends on z and t, so only at certain discrete times and places do they anti-align and exactly cancel each other. But on average they cancel.

8.5 Momentum

Electromagnetic waves carry momentum. However, all the other waves we've studied (longitudinal spring/mass and sound waves, transverse string waves, etc.) *don't* carry momentum. (However, see Footnote 1 above.) Therefore, it is certainly not obvious that electromagnetic waves carry momentum, because it is quite possible for waves to carry energy without also carrying momentum.

A quick argument that demonstrates why an electromagnetic wave (that is, light) carries momentum is the following argument from relativity. The relativistic relation between a particle's energy, momentum, and mass is $E^2 = p^2c^2 + m^2c^4$ (we'll just accept this here). For a massless particle ($m = 0$), this yields $E^2 = p^2c^2 \Longrightarrow E = pc$. Since photons (which is what light is made of) are massless, they have a momentum given by $p = E/c$. We already know that electromagnetic waves carry energy, so this relation tells us that they must also carry momentum. In other words, a given part of an electromagnetic wave with energy E also has momentum $p = E/c$.

However, although this argument is perfectly valid, it isn't very satisfying, because (a) it invokes a result from relativity, and (b) it invokes the fact that electromagnetic waves (light) can be considered to be made up of particle-like objects called photons, which is by no means obvious. But why should the *particle* nature of light be necessary to derive the fact that an electromagnetic *wave* carries momentum? It would be nice to derive the $p = E/c$ result by working only in terms of waves and using only the results that we have developed so far in this book. Or said in a different way, it would be nice to understand how would someone living in, say, 1900 (that is, pre-relativity) would demonstrate that an electromagnetic waves carries momentum. We can do this in the following way.

Consider a particle with charge q that is free to move around in some material, and let it be under the influence of a traveling electromagnetic wave. The particle will experience forces due to the **E** and **B** fields that make up the wave. There will also be damping forces from the material. And the particle will also lose energy due to the fact that it is accelerating and hence radiating (see Section 8.7). But the exact nature of the effects of the damping and radiation won't be important for this discussion.[3]

Assume that the wave is traveling in the z direction, and let the **E** field point along the x direction. The **B** field then points along the y direction, because $\mathbf{E}\times\mathbf{B} \propto \mathbf{k}$. The complete motion of the particle will in general be quite complicated, but for the present purposes it suffices to consider the x component of the particle's velocity, that is, the component that is parallel to **E**.[4] Due to the oscillating electric field, the particle will (mainly) oscillate back and forth in the x direction. However, we don't know the phase. In general, part of the velocity will be in phase with **E**, and part will be $\pm 90^\circ$ out of phase. The latter will turn out not to matter for our purposes,[5] so we'll concentrate on the part of the velocity that is

in phase with **E**. Let's call it $\mathbf{v}_E$. We then have the pictures shown in Fig.8.

(some time t)

You can quickly verify with the righthand rule that the magnetic force $q\mathbf{v}_E \times \mathbf{B}$ points forward along **k** in *both* cases. $\mathbf{v}_E$ and **B** switch sign in phase with each other, so the two signs cancel, and there is a net force forward. The particle therefore picks up some forward momentum, and this momentum must have come from the wave. In a small time dt, the magnitude of the momentum that the wave gives to the particle is

$$|d\mathbf{p}| = |\mathbf{F}_B\,dt| = |q\mathbf{v}_E \times \mathbf{B}|dt = qv_E B\,dt = \frac{qv_E E\,dt}{c}. \quad (58)$$

What is the energy that the wave gives to the particle? That is, what is the work that the wave does on the particle? (In the steady state, this work is balanced, on average, by the energy that the particle loses to damping and radiation.) Only the electric field does work on the particle. And since the electric force is qE, the amount of work done on the particle in time dt is

[3] If the particle is floating in outer space, then there is no damping, so only the radiation will extract energy from the particle.

[4] If we assume that the velocity v of the particle satisfies v ¿c (which is generally a good approximation), then the magnetic force, $q\mathbf{v}\times\mathbf{B}$ is small compared with the electric force, $q\mathbf{E}$. This is true because $B = E/c$, so the magnetic force is suppressed by a factor of v/c (or more, depending on the angle between **v** and **B**) compared with the electric force. The force on the particle is therefore due mainly to the electric field.

[5] We'll be concerned with the work done by the electric field, and this part of the velocity will lead to

$$dW = \mathbf{F}_E \cdot d\mathbf{x} = (qE)(v_E\, dt) = qv_E E\, dt. \tag{59}$$

(The part of the velocity that is ±90° out of phase with **E** will lead to zero net work, on average; see Problem [to be added].) Comparing this result with Eq. (58), we see that

$$|d\mathbf{p}| = \frac{dW}{c}. \tag{60}$$

In other words, the amount of momentum the particle gains from the wave equals $1/c$ times the amount of energy it gains from the wave. This holds for any extended time interval Δt, because any interval can be built up from infinitesimal times dt.

Since Eq. (60) holds whenever any electromagnetic wave encounters a particle, we conclude that the wave actually carries this amount of momentum. Even if we didn't have a particle in the setup, we could imagine putting one there, in which case it would acquire the momentum given by Eq. (60). This momentum must therefore be an intrinsic property of the wave.

Another way of writing Eq. (60) is

$$\frac{1}{A}\left|\frac{d\mathbf{p}}{dt}\right| = \frac{1}{c}\cdot\frac{1}{A}\frac{dW}{dt}, \tag{61}$$

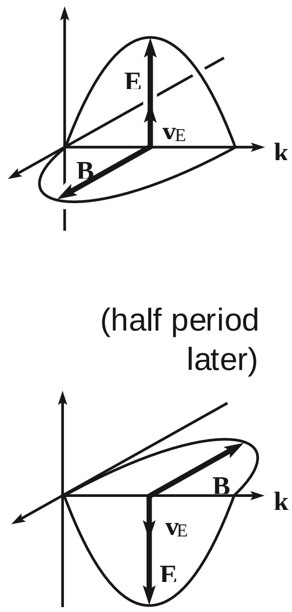
where A is the cross-sectional area of the wave under consideration. The lefthand side is the force per area (in other words, the pressure) that the wave applies to a material. And from Eqs. (46) and (47), the righthand side is $|\mathbf{S}|/c$, where **S** is the Poynting vector. The pressure from an electromagnetic wave (usually called the *radiation pressure*) is therefore

$$\text{Radiation pressure} = \frac{|\mathbf{S}|}{c} = \frac{|\mathbf{E}\times\mathbf{B}|}{\mu_0 c} = \frac{E^2}{\mu_0 c^2}. \tag{62}$$

You can show (see Problem [to be added]) that the total force from the radiation pressure from sunlight hitting the earth is roughly $6\cdot 10^8$ kgm/s^2 (treating the earth like a flat coin and ignoring reflection, but these won't affect the order of magnitude). This force is negligible compared with the attractive gravitational force, which is about $3.6\cdot 10^{22}$ kgm/s^2. But for a small enough sphere, these two forces are comparable (see Problem [to be added]).

Electromagnetic waves also carry angular momentum *if* they are polarized (see Problem [to be added]).

zero net work being done.



$q\mathbf{v}\times\mathbf{B}$ points along **k** in both cases

Figure 8

FORCES AND MOTION

Overview

In Grade 7, you described an object's motion in terms of displacement, speed or velocity, and acceleration. You performed activities wherein you interpreted or created visual representations of the motion of objects such as tape charts and motion graphs. The concepts were arrived at by studying examples of uniform motion or objects moving in straight line at constant speed. Then you were also introduced to non-uniform motion where the object covers unequal distances or displacements at equal intervals of time. When a jeepney starts moving, it speeds up. When a jeepney nears a stop sign, it slows down. The jeepney is covering different displacements at equal time intervals and hence it is not moving at a uniform velocity. In other words, the jeepney is accelerating.

Most of the motions we come across in our daily life are non-uniform and the primary cause of changes in motion is FORCE. In this module, you will learn about the effects of force on motion. Newton's Three Laws of Motion – the central organizing principle of classical mechanics – will be presented and applied to real-life situations.

At the end of Module 1, you will be able to answer the following key questions:

Do forces always result in motion?

What are the conditions for an object to stay at rest, to keep moving at constant velocity, or to move with increasing velocity?

How is force related to acceleration?

In the lower grades, you learned that an object can be moved by pushing or pulling. In physics, this push and pull is referred to as *force* (F). Consider a ball on top of a table as shown in Figure 1. If someone pushes the ball, it will move or roll across the surface of the table (Figure 1a). And when it is again pushed in the direction of its motion, it moves farther and even faster (Figure 1b). But when you push it on the other side instead, opposite to the direction of its motion, the ball may slow down and eventually stop (Figure 1c). Lastly, when you push it in a direction different from its original direction of motion, the ball also changes its direction (Figure 1d). Force therefore can make objects move, move faster, stop, or change their direction of motion. But is this always the case? Can force always bring about change in the state of motion of an object?

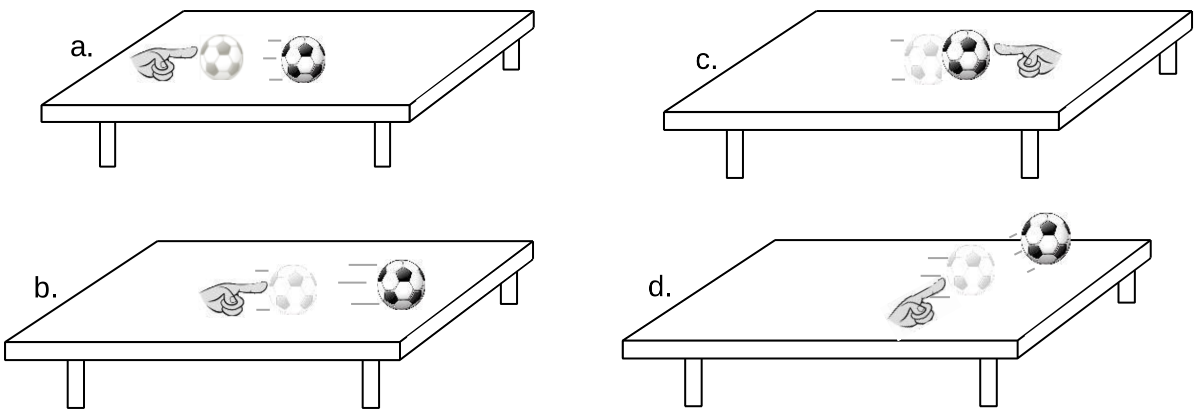


Figure 1. Effect of force on an object

Balanced and Unbalanced Forces

An object may be acted upon by several forces. For example, an object may be pushed and pulled in different directions at the same time. To identify which of these forces would be able to cause change in the motion of the object, it is important to identify all the forces acting on it.

To accurately describe the forces acting on an object, it is important for you to be familiar first with the following terms: *magnitude, direction, point of application,* and *line of action*. Forces are described in terms of these properties. **Magnitude** refers to the size or strength of the force. It is commonly expressed in Newton (N). Consider the diagram in Figure 2 showing a force, represented by the arrow, acting on a ball. The direction of the arrow indicates the **direction** of the force while the length of the arrow represents the relative magnitude of the force. If the force applied on the ball is doubled, the length of the arrow is increased two times. The **line of action** is the straight line passing through the **point of application** and is parallel to the direction of the force.

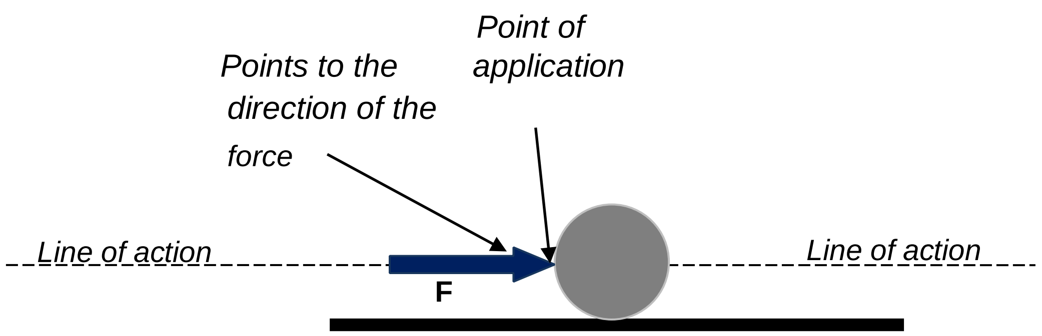


Figure 2. Force acting on a ball

Activity 1
Forces on objects at rest

Objectives:

After performing this activity, you should be able to identify the forces acting on an object at rest.

Materials:

pen pair of scissors string book

Procedure

Situation 1

1. Hang a pen by a piece of string as shown in Figure 3a.

Q1. Is the pen at rest or in motion?

Figure 3a. Hanging pen

Q2. Are there forces acting on the pen? If yes, draw the forces. You may use arrows to represent these forces.

2. Cut the string with a pair of scissors.

Q3. What happens to the pen? What could have caused the pen's motion?

Situation 2

1. Place a book on top of a table as shown in Figure 3b.

Q4. Is the book at rest or in motion?

Q5. Are there forces acting on the book? If yes, draw the forces acting on the book.

2. Let one member of your group push the book in one direction and another member push it in the opposite direction at the same time with the same amount of push (force).

Figure 3b. Book on a table

Q6. Did the book move? How will you make the book move?

In the situations above, both the pen and the book are at rest. But this does not mean that there are no forces acting on them. So what causes them to stay in place? Consider the next activity.

Activity 2
Balance of forces

Objectives:

After performing this activity, you should be able to:

1. examine the conditions when two forces balance, and

2. explain the effect of balanced forces on the state of motion of an object.

Materials:

4 sets spring balance
1 piece of sturdy cardboard
threads

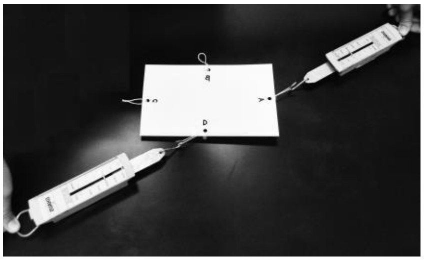


Figure 4

Procedure:

1. Bore four holes around the cardboard as shown. Label the holes A, B, C, and D.

2. Attach threads to the holes.

3. Attach a spring balance to thread A and another one to thread D. Hold the cardboard to keep it still. Pull the balances along the same line such that when released, the cardboard remains at rest.

4. When the cardboard is at rest, examine the magnitudes and directions of the two forces by reading the spring balance.

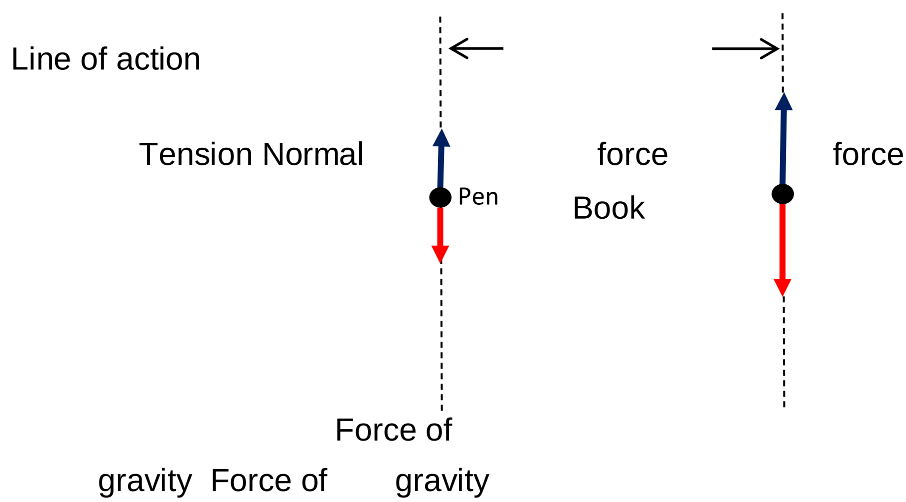
5. Draw the line of action of the forces acting on the cardboard. Extend the lines until they intersect. Mark the point of intersection and draw arrows starting at this point to represent the forces acting on the cardboard.

6. Repeat steps 3 to 5 for pair B and C.

Q7. When the cardboard is at rest, how do the magnitudes and directions of the pair of forces acting on it compare?

7. Now here is a challenge. Find out the directions of all the forces such that when all the threads were pulled with the same amount, the cardboard will not move or rotate when released.

Q8. If you draw the lines of action of all the forces acting on the board and extend the lines, what will you get?



(a) Pen (b) Book

Figure 5: Force diagram

The diagram in Figure 5 shows the forces acting on the a) pen and b) book in Activity 1. You learned in lower grades that all objects fall down because gravity pulls on them towards the center of the earth. But what makes the pen and the book stay at rest? The pen stays in place because of another force that acts on it that is supplied by the string which we refer to in physics as tension force (T). The book, on the other hand, stays at rest because of the upward push exerted on it by the table which we refer to as normal force (F_n). Both the tension force and normal force counteract the pull of gravity (F_g) that acts on the objects. Study the diagram. How do the lengths of the arrows in each case compare? How do the magnitudes and directions of the pair of forces compare?

In both cases, we can infer that the objects remained at rest because the forces acting on them are equal in magnitude and in opposite directions and they lie along the same line of action (Figure 5). The forces are **balanced**. This was also demonstrated in Activity 2. Also, if you try out step 7 in Activity 2, you will find that the lines of action of the four forces intersect through a single point. This also explains why the body does not move or rotate.

Unbalanced Forces

If you cut the string connected to the pen, the pen will fall. Or if you push the book on one side across the table, the book will move but will not continue moving if you don't continuously push it. The pen falls down because there is no more force acting on it to counteract the pull of gravity. The book moves because of the push that you applied. In other words, the forces acting on these objects are no longer balanced. If an object initially at rest is under an unbalanced force, it moves in the direction of the unbalanced force.

How about if the object is already in motion, how will the unbalanced force affect its motion?

Place a ball on the desk then push it gently to one side. Observe the motion of the ball as it rolls down the desk. What makes the ball stop rolling after sometime? Again, you need to identify the forces acting on the ball. You can see in Fig. 6 that the force of gravity and the normal force are again acting on the ball. But these forces are balanced, and so the ball stays on top of the desk. However, there is another force that acts on the ball along the horizontal line or along the force that set the ball in motion. Do you still remember your lesson on friction in the lower grades? You learned that friction is a force that acts between surfaces that are in contact with one another. Friction in general acts opposite the direction of motion. In the case of the rolling ball, the frictional force acts between the surfaces of the ball and the desk and slows down the motion of the ball.

As the ball rolls to the right as shown in Figure 6, friction acts to the left to retard its motion. Since you did not push the ball continuously there is no force present to balance the force of friction. So the ball slowed down and eventually stopped.

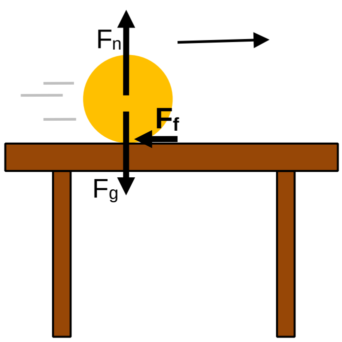


Figure 6. Forces acting on a rolling ball

Again, due to the unbalanced force, the object changes its state of motion hence we say that it accelerates. Note that acceleration is not just an increase in velocity**,** but also a decrease in velocity**.**

Combining Forces

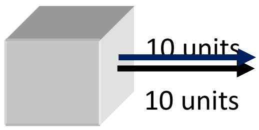
When we combine or add forces to determine the net or unbalanced force, we will limit our discussion to those forces which act along the same line of action. The algebraic signs + and – are used to indicate the direction of forces. Unlike signs are used for forces acting in opposite directions, like in the case of the book lying on the table. The force of gravity (F_g) and normal force (F_n) are assigned opposite signs - F_n is given a positive (+) sign while F_g is given a negative (-) sign. If both F_g and F_n are given a magnitude value of 3 units, then the net force along this line (vertical) will be:

$$
\begin{aligned}
F_{net} &= F_n + F_g \\
&= 3 \text{ units} + (-3 \text{ units}) \\
&= 0
\end{aligned}
$$

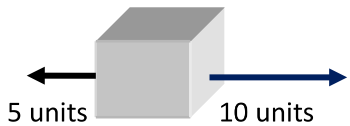
If the sum of the forces equate to zero, they are considered **balanced**. If the algebraic sum is not equal to zero, the forces are not balanced. The non-zero sum is the net or unbalanced force. This unbalanced or net force would cause a change in a body's state of motion.

Concept check:

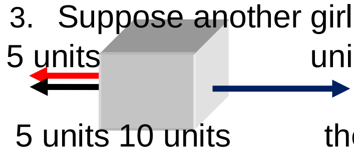
Study the illustrations and answer the questions that follow.



1. A boy and a girl are pulling a heavy crate at the same time with 10 units of force each. What is the net force acting on the object?



2. What if the boy and the girl pull the heavy crate at the same time in opposite directions with 10 units and 5 units of force respectively, what will be the net force on the object? Will the object move? To what direction will it move?



3. Suppose another girl pulls the heavy crate in with 5 units of force in the same direction as the girl, what will be the net force that will act on the object? Will the object move?

Newton's Three Laws of Motion

The principles behind Newton's laws of motion are very significant in understanding the motion of objects in our universe. Their applications are all around us. Understanding these laws therefore helps us understand why the things around us move or behave the way they do.

Newton's First Law of Motion: Law of Inertia

You learned that if the forces acting on an object at rest are balanced or if their algebraic sum equate to zero, the object stays at rest. This illustrates Newton's First Law of Motion, a principle that was primarily based on the works of Galileo. The following examples will help you understand this principle better.

Activity 3
Investigating inertia

Objective:

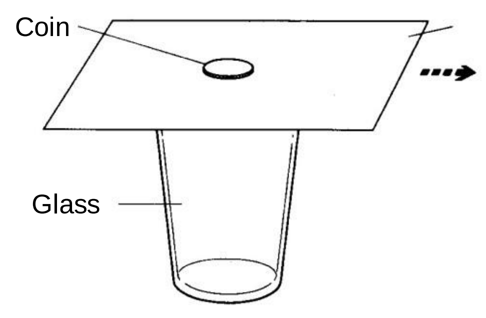
At the end of this activity, you should be able to demonstrate Newton's first law of motion.

Materials:

empty glass	5-peso coins (5 pcs or more)
cardboard	plastic ruler
1 peso coin	

Procedure

Coin Drop



1. Arrange the setup as shown in Figure 7.

2. Slowly pull the cardboard with your hand and observe what happens.

3. Arrange again the setup as shown. This time, quickly flick the cardboard with your finger. Observe again what happens.

Figure 7. Cardboard and coin

Q9. What happens when you slowly pulled the cardboard? Explain.

Q10. What happens when you flicked the cardboard? Explain.

Stack of Coins

4. Stack the coins on a flat level surface.

5. Quickly hit the coin at the bottom with the edge of the ruler.

Q11. What happens when you hit the coin at the bottom? Why is this so?

The examples above demonstrate the property of an object to resist any change in its state of motion. In physics, this property is known as **inertia**. The coin dropped into the glass because it was trying to remain in its state of rest. How about in the second example? How will you explain the behavior of the coins when one of them was hit with an edge of a ruler?

Measure of Inertia

All objects have the tendency to resist changes in their state of motion or keep doing what they are doing. However, changing a body's state of motion depends on its inertia. A more massive object which has more inertia is more difficult to move from rest, slow down, speed up, or change its direction.

Newton's first law states that *an object at rest will stay at rest or an object in motion will stay in motion and travel in straight line, as long as no external net force acts on it*. The object will change its state of motion only if there is unbalanced or net force acting upon it.

Law of Inertia

A body will remain at rest or move at constant velocity unless acted upon by an external net or unbalanced force.

Newton's Second Law of Motion: Law of Acceleration

You learned that when the velocity of a moving body changes, we describe the motion as one with acceleration. Is there any relationship between acceleration and any unbalanced force that acts on the body? Find out in the next activity.

Activity 4
Force and acceleration

Objective:

After this activity, you should be able to describe how the net force acting on an object affects its acceleration.

Procedure:

Consider this situation below:

A group of students conducted an experiment to determine the relationship between the force acting on the object and its acceleration. They used identical rubber bands to pull the cart as shown in Figure 8. They varied the number of rubber bands to vary the force acting on the cart. They started with 1 rubber band,

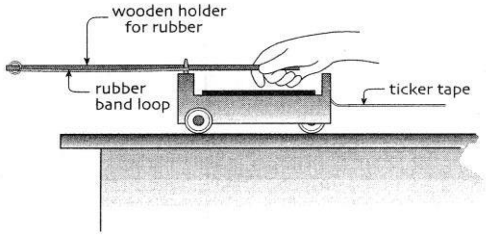


Figure 8. Cart pulled by rubber bands

then with 2, 3, and 4 rubber bands, making sure that they stretched the rubber bands to the same length every time they pull the cart. They used a ticker tape timer to determine the acceleration of the cart. A ticker tape was connected to the cart such that when the cart was pulled, the paper tape will be pulled through the timer. And as the paper tape was pulled through the timer, small dots are formed on the tape.

Starting with the tape for 1 rubber band, they marked the first clear dot and every 6th dot thereafter and cut the tape along these points (Figure 9). Then they pasted the strips side by side in order on a graphing paper to produce the tape chart for F=1 unit. They did the same for the other tapes to produce tape charts for F=2 units, F=3 units, and F=4 units.

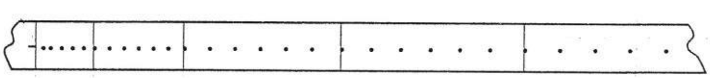


Figure 9: Sample tape

A. Tape chart analysis

1. Obtain from your teacher the copies of the tape charts produced by the students for the 4 runs.

Q12. Compare the charts. What similarities and differences have you noticed among them?

The length of strip in each chart represents the total distance travelled by the cart over a time interval of 0.10 seconds. Recall that the total distance travelled over a unit time gives the average velocity of the moving body, or speed when travelling in straight line. Hence, *each strip represents the average velocity of the cart over a time interval* of 0.10 seconds.

2. Examine the tape chart for F=1 unit.

Q13. What does the increase in the lengths of the strips suggest? What can you say about the motion of the cart - is it moving in uniform motion or is it accelerating? Is this also true with the other runs?

Q14. How do you compare the increase in length of the strips in F= 1 unit? What does this tell you about the change in the velocity of the cart? Is this also true with the other tape charts?

Q15. How do you compare the increase in length of the strips among the four tape charts? Which tape chart shows the greatest increase in the length of the strips? Which tape chart shows the least increase in the length of the strips?

3. Draw a line that passes through all the dots at the ends of the strips in F=1 unit. Do the same for the other tape charts.

Q16. Describe the line formed. Does the same pattern exist for the other tape charts?

B. Quantitative analysis

You can also use the tape chart to compute for the average velocity (v_{ave}), change in velocity (Δv), and acceleration (a) of the cart for each run. Work only on the tape chart assigned to your group. Other groups will be working on the other charts. You may follow the simple instruction below.

4. Label each strip 1,2,3,4, and 5 as shown in Figure 10.

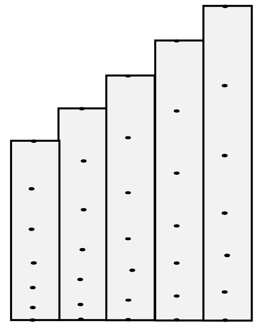


Figure10: Sample tape chart

5. Compute for the average velocity of the cart over each time interval by measuring the length of the strip and dividing it by the time covered to travel such distance. Example, if the length of the strip is equal to 2.5 cm, then the average velocity during that time interval will be

$$v_{ave} = 2.5 \text{ cm} / 0.10\text{sec}$$
$$= 25 \text{ cm/s}$$

Q17. How do the values of v_{ave} compare? What does this tell you about the motion of the cart?

6. Next, determine the difference in the average velocities (Δv) of the cart between two successive time intervals. Example, you can get the difference in the average velocities between strips 1 & 2, between strips 2 & 3, and so on.

Q18. How do the computed values of Δv compare? What does this tell you about the motion of the cart?

7. Recall that acceleration is defined as the change in velocity per unit of time. To get the acceleration of the cart, divide your computed values of Δv in step 6 by 0.10 seconds, the unit of time. Have at least three computed values of acceleration.

Q19. How do your computed values of acceleration compare?

8. Compute for the average acceleration a_{ave}.

9. Ask from the other groups the values of a_{ave} for the other tape charts. Record them all in Table 1 below.

Table 1. Computed values of a_{ave}

Tape chart	# of rubber bands	Computed a_{ave}
F = 1 unit	1	
F = 2 units	2	
F = 3 units	3	
F = 4 units	4	

Q20. In this activity, the number of rubber bands represents the magnitude or amount of the force acting on the cart. How is acceleration of the cart related to the amount of force acting on it?

If the net force acting on an object is constant, its velocity changes at a constant rate over time. Hence, it is considered to be moving with constant acceleration. In the tape chart, this is indicated by the uniform increase in length of the strips over time. But if the force acting on the object is changed, its acceleration will also change. In your previous activity, you noticed that as the number of rubber bands increases, the acceleration of the cart also increases. When the net force is doubled, acceleration is also doubled. When it is tripled, acceleration is also tripled. We can therefore say that at constant mass, the acceleration of an object is directly proportional to the magnitude of the unbalanced force F acting on it. This relationship can be mathematically expressed as:

$a = kF$ where k = mass

What if the mass of the object is changed and the force is kept constant? Acceleration also varies with the mass of the object. As the mass of the object increases, with the same amount of force applied, its acceleration decreases. This relationship can also be expressed as:

$a = k\,(1/m)$ where k = net force

If you combine these two relationships, you would come up with this relationship:

Law of Acceleration

“The acceleration of an object is directly proportional to the magnitude of the net force acting on it and is inversely proportional to its mass.”

This statement actually pertains to Newton’s second law of motion or Law of Acceleration, because it is concerned with the relation of acceleration to mass and force. This can be expressed in equation form as:

Acceleration = Net force / Mass

$$a = F_{net} / m$$

This is often rearranged as: $F_{net} = ma$

Like any other quantity, force has a unit and is expressed in Newton (N). One Newton is defined as the amount of force required to give a 1-kg mass an acceleration of 1 m/s/s, or

$$1\text{Newton (N)} = 1\text{kg/ms}^2$$

Sample mathematical problem:

Suppose a ball of mass 0.60 kg is hit with a force of 12 N. Its acceleration will be:

$$a = \frac{F_{Net}}{m}$$

$$= \frac{12N}{0.60kg}$$

$$a = 20m/s^2$$

If the force is increased to 24 N for the same ball then,

$$a = \frac{24N}{0.6kg} = 40m/s^2$$

Free Fall and Newton's Second Law of Motion

Suppose you drop two books of different masses from the same height, which will hit the ground first?

Think about this: If we use the law of acceleration, the heavier book must be the one to hit the ground first because gravity pulls on it with more force. But if we use the law of inertia, the lighter book must be the one to hit the ground first because of its lesser inertia. But if you actually try it out, you would find that they will both reach the floor at the same time. How come?

Gravity acts on all objects on the earth's surface and causes them to accelerate when released. This acceleration, known as the acceleration due to gravity g, is the same for all objects on earth and is equal to 9.8 m/s^2. This means that when objects fall, their velocities increase by 9.8 m/s every 1 second.

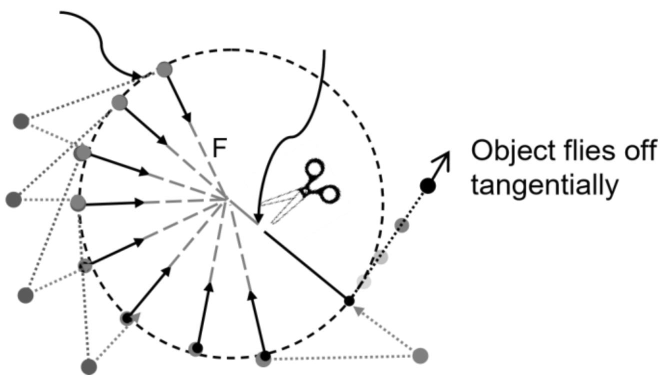
The books in the example above fall to the ground at the same rate (acceleration) even if they differ in mass. And since they were released from the same height at the same time, they will reach the ground at the same time.

Circular Motion and Newton's Second Law of Motion

Newton's Second Law was arrived at by studying straight line motion. Does this law apply to circular motion as well?

Try to whirl an object tied to a string horizontally above your head. Then observe what happens if you release the object. How does it travel after release?

You learned in Grade 7 that acceleration does not only refer to change in speed. It also refers to change in direction. In the case of circular motion, the whirling object accelerates not due to the change in its speed but to the change in the direction of its velocity. By Newton's second law of motion, a net force must be acting on accelerating objects. So where is this net force coming from? *For the stone to move in a horizontal circle, what must you do?* You have to pull the stone inward towards the center of the circular path, right? So the force comes from the string that pulls the object towards the center of its circular path (Figure 11). If you remove this force by either cutting or releasing the string, you will observe that the object will continue to move straight and fly off tangential to the path. This is the natural tendency of the object if there is no net force acting on it, according to the First Law of Motion. But because of the net force from the string, instead of going straight, the object accelerates inwards thereby covering a circular path. The object is said to be in uniform circular motion.



(Circular) path

Figure 11. Object in circular motion

<u>Think about this!</u>

If the object in uniform circular motion is accelerating towards the center of the circle, why does it maintain a circular path at a constant radius and never get closer to the center of the circle?

Newton's Third Law of Motion: Law of Interaction

Activity 5
Action-reaction

Objective:

In this activity, you should be able to compare two interacting forces in terms of magnitude and direction.

Materials:

2 spring balances string

Procedure:

1. Connect 2 spring balances with their hooks. Ask your partner to hold one end of the balance while you hold the other end horizontally. Pull the spring balance while your partner just holds the other end. Record the reading on each balance.

Q21. What is the reading on your balance and that of your partner? What do these values represent?

Q22. How do you compare the direction of your partner's and your force?

2. Pull the spring balance harder. Be careful not to exceed the maximum reading on the spring balance.

Q23. What is the reading on your balance and that of your partner?

Q24. How do you explain your observation?

3. Attach one end of your spring balance to the wall, while the other end is connected to the second spring balance. Ask your partner to pull the spring balance. Observe the reading on each balance.

Q25. What is the reading in each balance?

Q26. Compare the direction of the forces exerted on the two ends of the connected spring balance.

In the simplest sense, a force is a push or a pull. However, Newton realized that a force is not a thing in itself but part of mutual action, an interaction, between one thing and another.

For example, consider the interaction between a hammer and a nail. A hammer exerts a force on the nail and drives it into a board. But this is not the only force present for there must also be a force exerted on the hammer to stop it in the process. What exerts this force? The nail does. Newton reasoned that while the hammer exerts a force on the nail, the nail exerts a force on the hammer. So, in the interaction between the hammer and the nail, there is a pair of forces, one acting on the nail and the other acting on the hammer. Such observations led Newton to his third law: the law of interaction.

In Activity 5, you observed the similarities and differences between the interacting forces in terms of magnitude and direction. This relationship is stated in Newton's Third Law of Motion – Law of Interaction.

Law of Interaction (Action-Reaction)

"For every action, there is an equal and opposite reaction."

Because the forces are equal in magnitude and opposite in direction, do you think they will cancel each other? In this case, no addition of forces will take place because these forces are acting on different bodies. The spring balances act on each other.

The difference between the forces related to Law of Interaction and forces in a balanced state are as follows:

Action-Reaction Forces	Balanced Forces
- Two forces are equal in size. - Two forces are opposite to each other in terms of direction. - Two forces have the same line of action. - Action acts on one object, while reaction acts on another object.	• Two forces are equal in size. • Two forces are opposite to each other in terms of direction. • Two forces act along the same line. • Two forces act upon the same object.

References and Links

Herr, N. (2008). *The sourcebook for teaching science: Strategies, activities, and instructional resources, grades 6-12*. San Francisco, CA: Jossey-Bass.

Department of Education, Culture and Sports. (DECS). (n.d.). *Physics teachers resource manual*. Pasig City. Author.

Department of Education. (DepEd). (2002). *Lesson Plans for the 2002 Basic Education Curriculum: Science IV*. Pasig City. Author.

The Physics Classroom (1996-2013). *Newton's laws.* Retrieved from www.physicsclassroom.com

University of the Philippines National Institute for Science and Mathematics Education Development. (2002). *Practical work on high school physics: Sourcebook for teachers.* Quezon City: Author.

Suggested time allotme

MODULEUnit 1 **2**

WORK AND ENERGY

Overview

In Module 1, you utilized Newton's Laws to analyze the motion of objects. You investigated the motion of an object in relation to force, mass and accel eration.

In this module, motion will be investigated from the perspective of work and energy. The concept of force, which you have taken up in Module 1, will be related to the concepts of work and energy.

At the end of this module, you should be able to answer the following questions:

What is work?
What is energy?
How are work, energy and power related?

What is *Work*?

What comes to your mind when you hear the word 'work'? The word *work* has many meanings. When people ask, "W*hat is your work?"* They refer to a job or employment. When people say, "*I'll meet you after work."* They refer to the part of a day devoted to an occupation or undertaking. When your teacher asks, "*Have you done your homework*?" They refer to the task or activity needed to be accomplished.

In Physics, *work* is an abstract idea related to energy. When work is done it is accompanied by a change in energy. When work is done by an object it loses energy and when work is done on an object it gains energy.

In Module 1, you learned that force can change the state of motion of an object. If an object is at rest, it can be moved by exerting force on it. If an object is moving, it can be made to move faster or stopped by applying force on it. In order to say that work is done on an object, there must be force applied to it and the object moves in the direction of the applied force.

<u>Work is done</u> if the object you push moves a distance in the direction towards which you are pushing it.

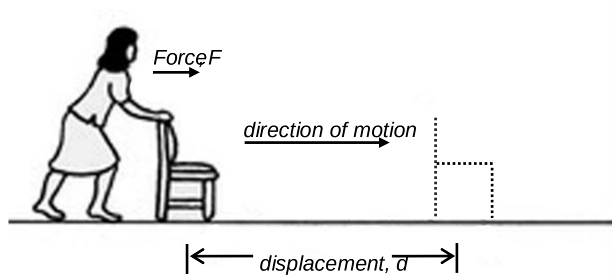


Figure 1. A girl pushing a chair

<u>No work is done</u> if the force you exert does not make the object move.

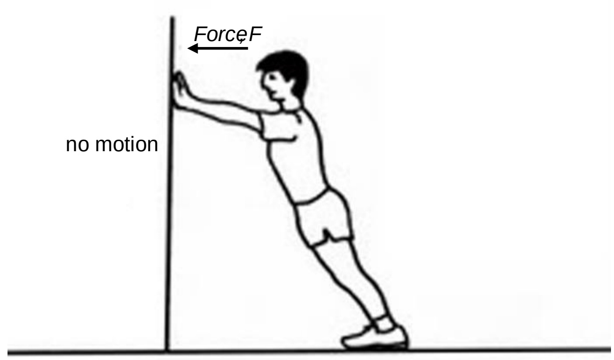


Figure 2. A boy pushing a wall

<u>No work is done</u> if the force you exert does not make the object move in the same direction as the force you exerted.

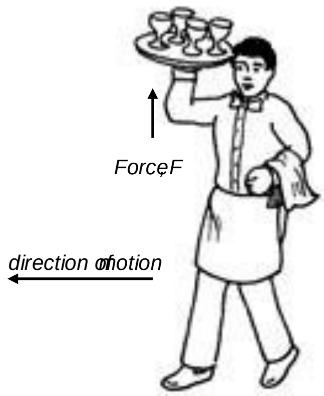


Figure 3. A waiter carrying a tray

Do activity 1 to see how well you understood ‘work’.

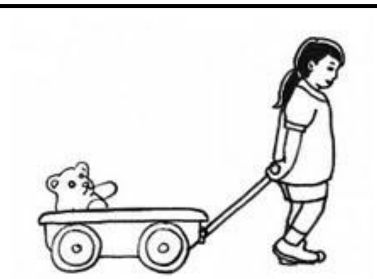
Activity 1
Is there work done?

Objective:

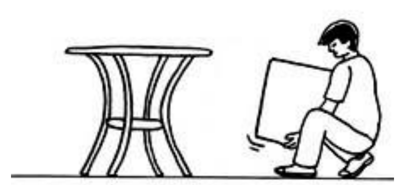
After performing this activity, you should be able to explain if work is done in situations represented.

Procedure:

Tell whether the situations shown below represent examples of work. Identify the one doing the work and on which object the work is done. Write in your notebook your answers and explanations.

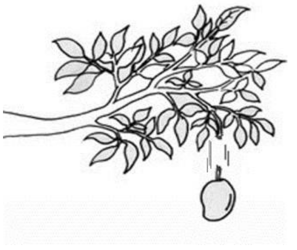


1. A girl pulling her cart.



2. A man lifting a box to be placed on a table.

3. A girl carrying a bag walking down a street.



4. A mango fruit falling from a branch.

Calculating Work

Work is done when the force (**F**) applied to the object causes the object to have a displacement (**d**) in the same direction as the force applied. The symbol for work is a capital **W**. The work done by a force can be calculated as

$$W = Fd$$

As you have learned in Chapter 1, the unit of force is

$$unit\ of\ force = kg\frac{m}{s^2} \text{ or } newton, N$$

Hence, the unit for *Work*, *W*

$$W = Fd$$

$$unit\ of\ work = unit\ of\ force\ x\ unit\ of\ displacement$$

$$unit\ of\ work = N\ m$$

$$unit\ of\ work = Nm \text{ or } joules, J$$

The unit, joule (J) is named after the English Physicist James Prescott Joule. This is also a unit of energy. One (1) Joule is equal to the work done or energy expended in applying a force of one Newton through a distance of one meter.

Sample problem:

Suppose a woman is pushing a grocery cart with a 500 Newton force along the 7 meters aisle, how much work is done in pushing the cart from one end of the aisle to the other?

$$W = Fd$$

$$W = 500\ N\ (7\ m)$$

$$W = 3500\ Nm$$

$$W = 3500\ J$$

Try solving this:

A book of mass 1 kg is on the floor. If the book is lifted from the floor to the top shelf which is 2 meters from the floor, how much work is done on the book?

Work is a Method of Transferring Energy

In Grade 7, you learned that there are different ways by which energy can be transferred from one place to another. Sound and light are transferred by waves; electrical energy is transferred by moving electrical charges through a complete circuit; and heat is transferred either by randomly moving particles, or by electromagnetic waves. Work is also a means of transferring energy from one object to another.

Do this!
Play a bowling game. Roll a plastic or rubber ball along the floor to hit an empty plastic bottle.

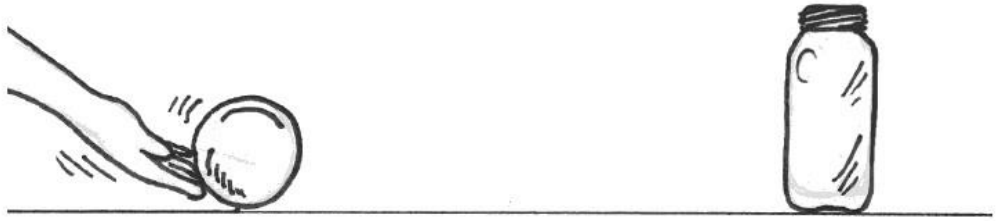


Figure 4. A ball and a plastic bottle

Is there work done on the ball?
What can a moving ball do?

You have done work on the ball. The force you exerted in pushing the ball is in the same direction as the motion of the ball. But then you did not continuously push the ball until it hits the empty bottle. You just gave it a nudge and then it rolled The force exerted on the ball changed the ball's motion. *'Something'* was transferred to the ball causing it to move continuously. That *'something'* is called energy. The energy became energy of motion of the ball.

A rolling ball can do work on the plastic bottle. When the ball hits the plastic bottle, it can push it through a distance. Thus, a moving object can do work on anything it hits because of its motion energy. Hence, energy is oftentimes defined as the ability or capacity to do work.

Since work is done on the ball, it gains energy while the person that does work on it loses energy. In the same manner, the rolling ball that does work on the empty plastic bottle loses energy while the bottle gains energy. This shows that when work is done, energy is transferred.

Kinetic Energy

The energy of a moving object is called energy of motion or kinetic energy (KE). The word kinetic comes from the Greek word *kinetikos* which means moving. Kinetic energy quantifies the amount of work the object can do because of its motion.

The plastic or rubber ball you pushed to hit an empty plastic bottle earlier has kinetic energy. The force applied caused the ball to accelerate from rest to a certain **velocity, *v*.** In Module 1, you learn that acceleration is the rate of change in velocity. In the equation,

$$a = \frac{v - v_i}{t}$$

where ***v*** is the final velocity, $\mathbf{v_i}$ is the initial velocity and **t** is the time.

Since the ball started from rest, the initial velocity is zero. Thus, the acceleration is

$$a = \frac{v}{t}$$

Substituting this in Newton's second law

$$F = ma$$

$$F = m\frac{v}{t}$$

The equation in finding the average velocity of the ball is

$$\bar{v} = \frac{v_i + v_f}{2}$$

Since the initial velocity is zero, the average velocity, $\bar{\boldsymbol{v}}$ is

$$\bar{v} = \frac{v_f}{2}$$

or

$$\bar{v} = \frac{v}{2}$$

The distance travelled by the ball before it hits the empty plastic bottle is given by the equation

$$d = v\bar{t}$$

where $\bar{v}$ refers to the average velocity

$$d = \frac{v}{2}t$$

Let's put the equations together. Since $\boldsymbol{W} = \boldsymbol{Fd}$ and $\boldsymbol{F} = \frac{\boldsymbol{mv}}{\boldsymbol{t}}$**,** we get

$$W = \frac{mv}{t}d$$

$$W = \frac{mv}{t}(\frac{1}{2}vt)$$

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

This shows that the work done in accelerating an object is equal to the kinetic energy gained by the object.

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

From the equation, you can see that the kinetic energy of an object depends on its mass and velocity. What will happen to the KE of an object if its mass is doubled but the velocity remains the same? How about if the velocity is doubled but the mass remains the same?

As you have learned in Module 1, the unit for *mass* is *kg* while for *velocity* it is *meter per second*.

Hence, the unit for *Kinetic Energy*, *KE* is

$$unit\ of\ KE = unit\ of\ mass\ x\ unit\ of\ velocity$$

$$unit\ of\ KE = kg\left(\frac{m}{s}\right)^2$$

$$unit\ of\ KE = kg\frac{m^2}{s^2}$$

But,

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

$$unit\ of\ KE = Nm\ or\ joules, J$$

Try solving this:

A 1000 kg car has a velocity of 17 m/s. What is the car’s kinetic energy?

Potential Energy

In activity 1 you were asked if the illustration of a man lifting a box demonstrates work.

Figure 5. A man lifting a box

Which/who is doing work in the illustration? Is it the table, the box, or the man? Yes you are correct if you answer “The man is doing work on the box.” What is the direction of the force exerted by the man on the box? Yes, it is upward. What is the direction of the motion of the box? Yes, it is upward. Then we can say, work is done by the man on the box.

As discussed previously, work is a way of transferring energy. Since the work is done by the man, he loses energy. The work is done on the box, hence the box gains energy.

In Grade 6, you learned about the force of gravity. It is the force that the earth exerts on all objects on its surface. It is always directed downward or towards the center of the earth. Hence, when an object is lifted from the ground, the work done is against the force of gravity. An object gains energy when raised from the ground and loses energy when made to fall. The energy gained or lost by the object is called *gravitational potential energy* or simply potential energy (PE).

For example when a 1.0 kg book is lifted 0.5 m from the table, the force exerted in lifting the book is equal to its weight.

$$F = Weight = mg$$

The acceleration due to gravity, ***g*** is equal to 9.8 meters per second squared. The work done

in lifting the book is

$$W = Fd$$

where the displacement (***d***) is the height (***h***) to which the object is lifted.

$$W = mgh$$

This shows that the work done in lifting an object is equal to the potential energy gained by the object.

$$PE = mgh$$

The potential energy of the book lifted at 0.5 m relative to the table is:

$$PE = 1\,kg\ x\ 9.8\ \frac{m}{s^2}\ x\ 0.5\ m$$

$$PE = 4.9\ J$$

If the book is lifted higher than 0.5 m from the table, what would happen to its potential energy?

The potential energy gained and lost by an object is dependent on the reference level. Consider a table and a chair shown in Figure 6. If the same 1.0 kg book is held 1 m above the table, the potential energy gained by it is 9.8 J with the table as the reference level; it is 14.7 J if the reference level were the chair; and 19.6 J if the reference level were the floor. If the book is released from a height of 2 m, the potential energy lost when it reaches the level of the table top is 9.8 J; 14.7 J when it reaches the level of the chair; and 19.6 J when it reaches the floor.

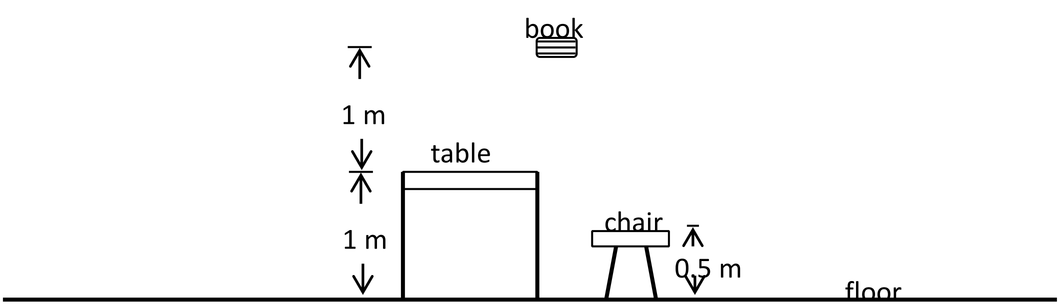


Figure 6. A table and a chair

Try solving this:

If the same 1.0 kg book is lifted to 0.5 m above the table, but the table top is 1.0 m above the floor, what would be the potential energy of the book if the reference level were the floor?

The energy of an object above the ground is called potential energy because it is a 'stored' energy. It has the potential to do work once released. Think of water held in a dam. It has potential energy. Once released, the water has the potential to move objects along its way. The potential energy of the water is transformed into kinetic energy.

The gravitational potential energy is just one type of potential energy. Another type is the elastic potential energy. Springs and rubber bands are called elastics. When elastics are stretched and then let go, they will return to their original form if they were not stretched beyond their elastic limit.

The force needed to stretch or compress elastics depends on the elasticity of the object and the change in elongation. The relationship between the force and the change in elongation (Δl) was first observed by Robert Hooke, hence, the name Hooke's Law expressed as:

$$F \propto \Delta l$$

$$F = k\Delta l$$

The proportionality holds true as long as the elastic limit of the elastics has not been reached. The proportionality or force constant *k* is a measure of the elasticity of the material.

Consider a spring. Since the force exerted in stretching a spring causes a change in length, then work is done on the spring. When work is done, energy is transferred. Thus, the stretched spring gains potential energy. The work done to stretch the spring a distance *x* (the symbol *x* is used instead of *d*) is equal to its potential energy. In equation;

$$W = PE = \frac{1}{2}kx^2$$

The elastic potential energy depends on how much the elastic object is stretched or compressed and the elasticity of the material.

What are the games you play using rubber bands? What do you do with the rubber bands in the games? Do Activity 2 to see how a rubber band 'stores' potential energy.

Activity 2
Rolling toy

Objective:

After performing this activity, you should be able to explain how a twisted rubber band can do work and relate the work done to potential energy.

Materials Needed:

1 clear plastic container with cover
1 rubber band
1 pc 3-cm round barbecue sticks 1 pc barbecue stick
with sharp part cut masking tape

Procedure:

1. Make a hole at the center of the cover and at the bottom of the plastic container.

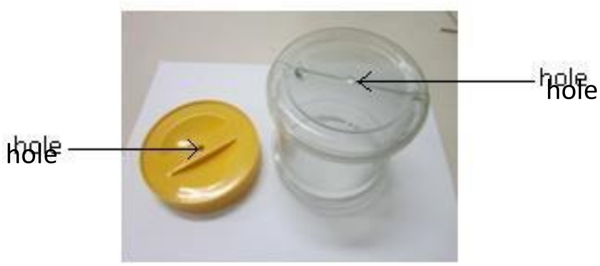


Figure 7. A plastic container with holes

2. Insert the rubber band into the hole at the bottom of the container. Insert in between the rubber band the 3-cm barbecue stick. Tape the barbecue stick to keep it in place.

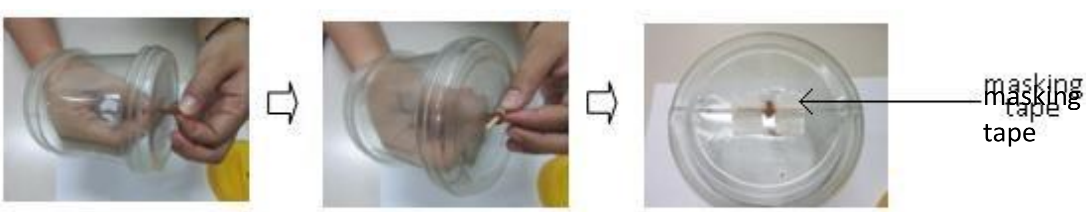


Figure 8. Steps in inserting the 3-cm barbecue stick

3. Insert the other end of the rubber band into the hole in the cover. Insert a bead or a washer to the rubber band before inserting the long barbecue stick.

Figure 9. Steps in inserting the bead and the long barbecue stick

4. You just made a toy. Twist the rubber band by rotating the long barbecue stick.

Figure 10. Rotating the long barbecue stick

5. Lay the toy on the floor. Observe it.

Figure 11. Finished toy

Q1. What happens to the toy?

Q2. What kind of energy is 'stored' in the rubber band?

Q3. What kind of energy does a rolling toy have?

Q4. What transformation of energy happens in a rolling toy?

Work, Energy, and Power

So far, we have discussed the relationship between work and energy. Work is a way of transferring energy. Energy is the capacity to do work. When work is done by an object it loses energy and when work is done on an object it gains energy. Another concept related to work and energy is power.

Power is the rate of doing work or the rate of using energy. In equation,

$$P = \frac{Work}{time} = \frac{Energy}{time}$$

The unit for power is joules per second. But maybe, you are more familiar with *watts* which is commonly used to measure power consumption of electrical devices. The unit *watt* is named after James Watt who was a Scottish inventor and mechanical engineer known for his improvements on steam engine technology. The conversion of unit from joules per second to watts is:

$$1\ watt = \frac{1\ joule}{1\ second}$$

Do Activity 3 to see your power output in walking or running up a flight of stairs.

Activity 3
How POWER-ful am I?

Objective:

After performing this activity, you should be able to compute for your power output in walking or running up a flight of stairs.

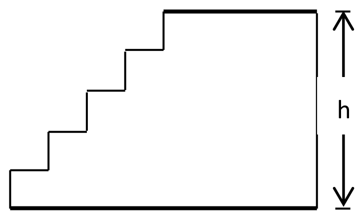
Materials Needed:

meterstick timer

Procedure:

1. Write the group members' names in the first column of Table 1.
2. Enter each member's weight in column 2. To solve for the weight, multiply the mass (in kg) by acceleration due to gravity (g=9.8 m/s^2).

3. Measure the height of the flight of stairs that you will climb. Record it on the table.



4. Each member will walk or run up the flight of stairs. Use a stopwatch or any watch to get the time it takes for each member to climb the stairs. Record the time in the 4th column.

5. Solve for the energy expended by each member. Record them in the 5th column of the table.

6. Compute for the power output of each member.

Table 1

Name	**Weight (N)**	**Height of stairs (m)**	**Time taken to climb the stairs (s)**	**Energy expended (J)**	**Power (J/s)**

Q1. Who among the group members had the highest power output?

Q2. What is the highest power output?

Q3. Who among the group members had the lowest power output?

Q4. What is the lowest power output?

Q5. What can you say about the work done by each member of the group? Did each member perform the same amount of work in climbing the stairs?

Q6. What factor/s determined the highest/lowest power output?

These are the concepts that you need to remember about work and energy:

- Work is done on an object when the force applied to it covers a distance in the direction of the applied force.
- Work is a way of transferring energy.
- When work is done by an object it loses energy and when work is done on an object it gains energy.
- The energy of an object enables it to do work.
- A moving object has energy called energy of motion or kinetic energy.
- An object above a specified level has energy due to its position called potential energy.
- An elastic object that is stretched or compressed or twisted has energy called potential energy.
- Power is the rate of doing work or the rate of using energy.

References and Links

Henderson, Tom. (2013). *Work and energy*. Retrieved from https://itunes.apple.com/us/book/work-and-energy/id611940649?mt=13

Hewitt, P.G. (2002). *Conceptual physics*. USA: Prentice-Hall, Inc. Saddle River, New Jersey.

Kirkpatrick, L.D., & Wheeler, G.F. (1998). *Physics: A world view*. Forth Worth: Saunders College.

Ostdiek, V.J., & Bord, D.J. (1987). *Inquiry into physics*. New York: West Publishing.

DepEd. *Science and Technology IV*. SEDP Series. (1992). Philippines: Book Media Press, Inc.

Suggested time allotme

MODULE 1 Unit **3**

SOUNDS

Overview

"*Hey I just met you and this is crazy. So here's my number so call me maybe..*." This is the popular song of Carly Rae Jepsen. I bet you know this song. Can you sing the other lines? Is this the ring tone of your mobile? What about your ring back tone? Would you want that of Maroon 5's payphone? "^{Cadd9}*I'm at the payphone trying to*^{G} *call home.* ^{Em}*All of my change I've spent*^{Dsus4} *on you..."* These are cool, lovely tunes, and nice sounds.

The Science of Sound has gone all the way from a mere transfer of energy to the creation of tunes and music for entertainment. Most of our gadgets are sound embedded to amuse us. In the field of geology and oceanography, sound is used to determine depths. The health sciences are also using sound for medical purposes. Some animals are dependent on sound for movement. The newest focus of sound science is on ecology where ecological patterns and phenomena are predicted based on sounds released by the different components of the ecosystem. So, are you ready to have fun with sounds?

In this module, you will learn sound propagation. While you learn about sound, wave description and characteristics will also be introduced to you. Among the characteristics, you will focus on the speed of sound. You will find out through simple activities through which medium sound travels fastest. You will also find out how the temperature of the medium affects the speed of sound. In the quest to explore more about sound science, you will be acquainted with the properties of waves, specifically reflection and refraction.

Through which medium does sound travel fastest- solid, liquid, or gas?

How does the temperature of the medium affect the speed of sound?

How are reflection and refraction manifested in sound?

Propagation and Characteristics of Sound

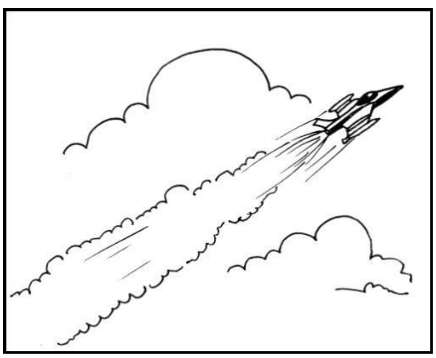


Figure 1. Supersonic

Figure 2. Hearing Sounds

Have you experienced hearing a sonic boom? Figure 1 shows a whitish cloud at the tail end of the aircraft. This usually happens when the aircraft travels at a speed faster than the speed of sound, i.e., the aircraft travels at supersonic speed producing a sonic boom.

A sonic boom happens when the aircraft or any vehicle breaks the sound barrier while it accelerates and outruns the speed of sound. A loud explosive sound is heard on the ground and is called a sonic boom. The aircraft that does this is usually called supersonic. There are more amazing occurrences or phenomena related to sound. Read on and find out.

Sound Propagation

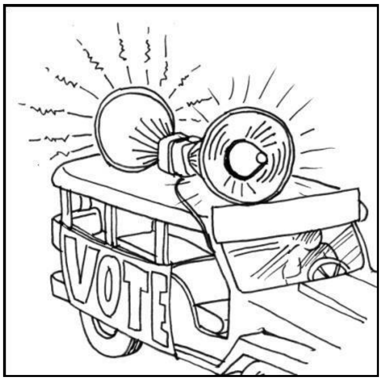


Figure 3. Propagating Sound

Sound consists of waves of air particles. Generally, sound propagates and travels through air. It can also be propagated through other media. Since it needs a medium to propagate, it is considered a mechanical wave. In propagating sound, the waves are characterized as longitudinal waves. These are waves that travel parallel to the motion of the particles. Do all these terms and concepts seem confusing? Let's try the succeeding activities to get a clearer picture of what sound waves are.

Activity 1
The dancing salt and the moving beads!

Objectives:

At the end of the activity, you will be able to infer that:

1. sound consists of vibrations that travel through the air; and
2. sound is transmitted in air through vibrations of air particles

Materials:

1 rubber band
1 piece of plastic sheet
1 empty large can of powdered milk - 800 g
1 wooden ruler
1 empty small can of evaporated milk - 400 mL rock salt
1 dowel or 1 wooden rod
1 blue bead
4 colored beads
3 inches of tape
2 large books
scissors

5 pieces of string
paper slinky spring transistor
radio

Procedure:

Part A: Vibrations produce sound

1. Prepare all the materials needed for the activity. Make sure that you find a work area far enough from other groups.

2. Put the plastic tightly over the open end of the large can and hold it while your partner puts the rubber band over it.

3. Sprinkle some rock salt on top of the plastic.

4. Hold the small can close to the salt and tap the side of the small can with the ruler as shown in Figure 4.

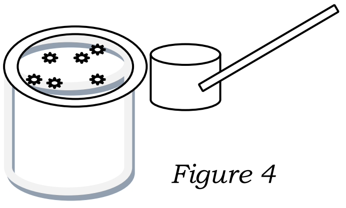


Figure 4

Q1. What happens to the salt?

5. Try tapping the small can in different spots or holding it in different directions. Find out how you should hold and tap the can to get the salt to move and dance the most.

Q2. How were you able to make the salt move and dance the most?

Q3. What was produced when you tapped the small can? Did you observe the salt bounce or dance on top of the plastic while you tapped the small can?

Q4. What made the salt bounce up and down?

Q5. From your observations, how would you define sound?

6. Switch on the transistor radio and position the speaker near the large can. Observe the rock salt.

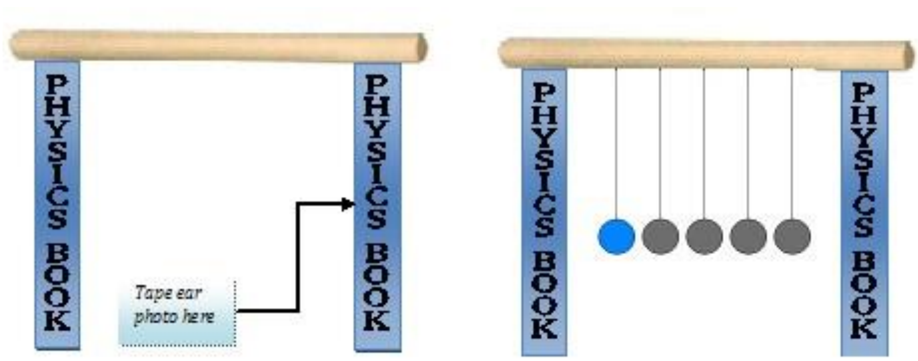
7. Increase the volume of the radio while it is still positioned near the large can. Observe the rock salt again.

Q6. What happened to the rock salt as the loudness is increased?

Q7. Which wave characteristic is affected by the loudness or the intensity of sound?

Part B: Transmitting sound

8. Let 2 books stand up as shown in Figure 5. Place the dowel on top of the 2 books.



Tape ear photo here

Figure 5. Set up for Activity 1B

9. Cut out an image of a human ear from a magazine and tape it to one of the books.

10. Start with the blue bead. Tape the string to the mark on the dowel that is farthest away from the ear.

11. Then tape the 4 colored beads to the other 4 marks. Make sure that all the beads hang in a straight line.

12. The colored beads represent air particles. Create vibrations (sound) in the air by tapping the blue bead toward the colored beads.

Q8. What happens to the other colored beads when the blue bead is tapped?

13. Create more vibrations by continuously tapping the blue bead and observe the other beads.

Q9. Are there occasion when the beads converge then expand?

14. If the beads represent air particles, what do the converging and expanding of the beads represent?

15. Connect one end of the slinky to a fixed point. Hold the other end then push and pull the slinky continuously. Record your observations.

Q10. Are there converging and expanding parts of the slinky?

Q11. How then is sound classified as a wave?

16. This time shake the other end of the slinky while the other end is still connected to the fixed point. Record your observations.

Were you able to get good sets of data from the activity? Did you enjoy watching the salt dance and the beads move? The salt and the beads represent particles of air when disturbed. The disturbance encountered by the salt and the beads causes the salt to bounce up and down and the beads to move together and spread alternately. In grade 7, you discussed that energy is transferred or transmitted from one object to another. Bouncing salt is also a manifestation of energy transmission. When sound is created by tapping the small can, the wave (sound) is transmitted by air to the larger can causing the plastic cover of the larger can to vibrate transferring energy to the rock salt. And voila!—dancing rock salt!

What about the beads? Did you observe the alternating converging and spreading of the beads? Compare this to your observations in the slinky spring. The converging portions of the beads match the compressions in the slinky while the spreading portions are the rarefactions of the slinky. With the compressions and rarefactions, what you were able to produce is called a longitudinal wave. Longitudinal waves are waves that are usually created by pulling and pushing the material or medium just like in the slinky (Figure 6). Alternating compressions and rarefactions are observed. These compressions and rarefactions move along with the direction of the pushing and pulling activity of the material or medium. Thus, the wave moves parallel to the motion of material or the particles of the medium. This is known as a **longitudinal wave**.

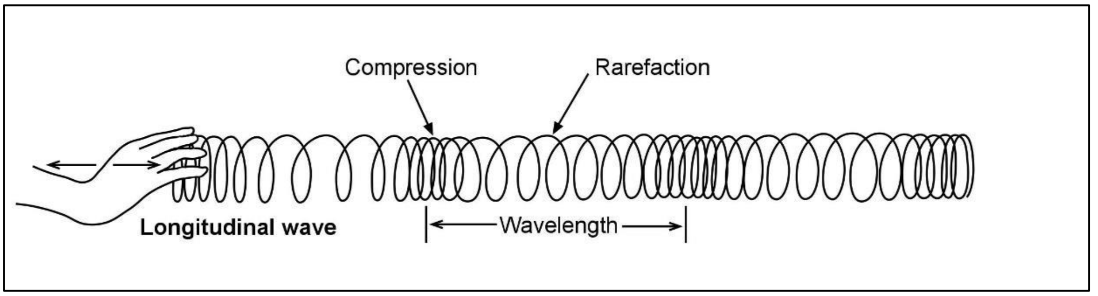


Figure 6. Longitudinal wave

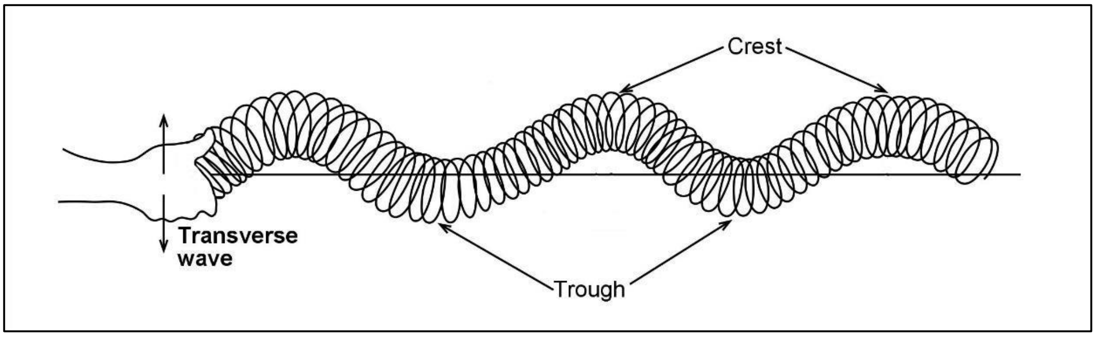


Figure 7. Transverse wave

Let us compare the longitudinal wave with the other kind of wave known as a transverse wave in Figure 7. The compressions resemble the trough while the rarefactions are the crests. Do you still remember these characteristics of waves? *The trough is the lowest part of a transverse wave while the crest is the highest portion*. The distance from one compression to the next or between two successive compressions in a longitudinal wave equals the **wavelength**. If you count the number of compressions passing by a certain point in 1 second, you are able to determine the **frequency** of the longitudinal wave. If you multiply the measured wavelength and the computed frequency you will be able to determine the speed of the wave. In equation,

$$v = f\lambda$$

There are other variations in the equation for the speed of the wave. The **period** of the longitudinal wave is the reciprocal of its frequency $\left[T = \frac{1}{f}\right]$. This means that the speed of the wave can be expressed as the ratio of the wavelength and the period,

$$v = \frac{\lambda}{T}$$

Let us try to compare the characteristics of longitudinal wave with that of the transverse wave in Activity 2.

Activity 2
Characteristics of waves: Comparing longitudinal and transverse waves
Objectives:

At the end of the activity, you will be able to:

1. distinguish the different characteristics of waves;
2. determine their frequency and wavelength; and
3. compute the wave speed based on the frequency and wavelength

Materials:

Pentel pen or permanent marker stopwatch or mobile phone
meterstick
old calendar (big poster calendar) or old newspaper metal slinky

Procedure:

1. Place the old calendar or old newspaper on the floor. Make sure that the newspaper or old calendar is long enough to accommodate the full length of the slinky spring.

2. Put the slinky on top of the old newspaper or old calendar. Ask one of your groupmates to hold one end of the slinky at the one end of the newspaper. This will serve as the *fixed end*.

3. Another groupmate will hold the other end of the slinky. This is the *movable end*.

4. The other members of the group should be along the sides so they can mark the corresponding crests. Identify a reference point **(point A)** along the slinky from which you are going to base your frequency count.

5. Shake the movable end. Apply just enough force to create large wave pulses. Make sure, however, that the *crest* and *trough* parts will still be formed within the newspaper area.

6. Another groupmate should count the number of pulses passing through point A in a minute. This is the *frequency* in waves per minute. You can convert this later to waves per second.

7. While your classmate is creating *transverse waves* by shaking the slinky, note by marking on the newspaper the crest and the trough of the created wave pulses.

8. Trace the wave form then measure the *wavelength* of the wave pulses. Record all your data on the answer sheet provided.

9. Repeat steps 5 to 8 for two more trials. Compute for the wave speed in each of the 3 trials. Determine also the average speed of the wave in the slinky.

10. For the second set up, repeat the whole procedure (steps 1 to 9) but this time instead of shaking the slinky, pull and push the slinky to create a longitudinal wave.

11. Note and mark the areas/regions in the newspaper where the slinky forms compressions and rarefactions.

12. Count the number of compressions passing through point A in a minute. This is the *frequency of the longitudinal wave in waves per minute*.

13. Measure the length between 2 compressions. This is the *wavelength of the longitudinal wave.*

14. Do this for three more trials, and then compute for the wave speed and the average speed of the wave in the slinky.

Q12. When there are more waves passing through the reference point in a period of time, which wave characteristic also increases?

Q13. When there are more waves passing through the reference point in a period of time, what happens to the wavelength of the waves?

As you have observed in Activity 2, there are many characteristics common to both transverse wave and longitudinal wave. The difference is in the motion of particles with respect to the direction of travel of the wave. Again, in a transverse wave, the movement of particles is perpendicular to the direction of wave travel. In a longitudinal wave, on the other hand, travel is parallel to the movement of the particles (Figure 8). In longitudinal waves, compressions are created when a push is applied on air. When air is pushed, there is a force applied on a unit area of air. From your science in the lower grades, the force applied per unit area is called **pressure**. This means that longitudinal waves are created by pressure and are also called *pressure waves*. Basically, sound as you have observed it is a longitudinal wave and a pressure wave. Just like the transverse waves, it has wave characteristics. Its movement is parallel to the particle motion. But do the particles in a way affect the movement of sound? What factors affect sound speed? Let us try finding this out in the next activities.

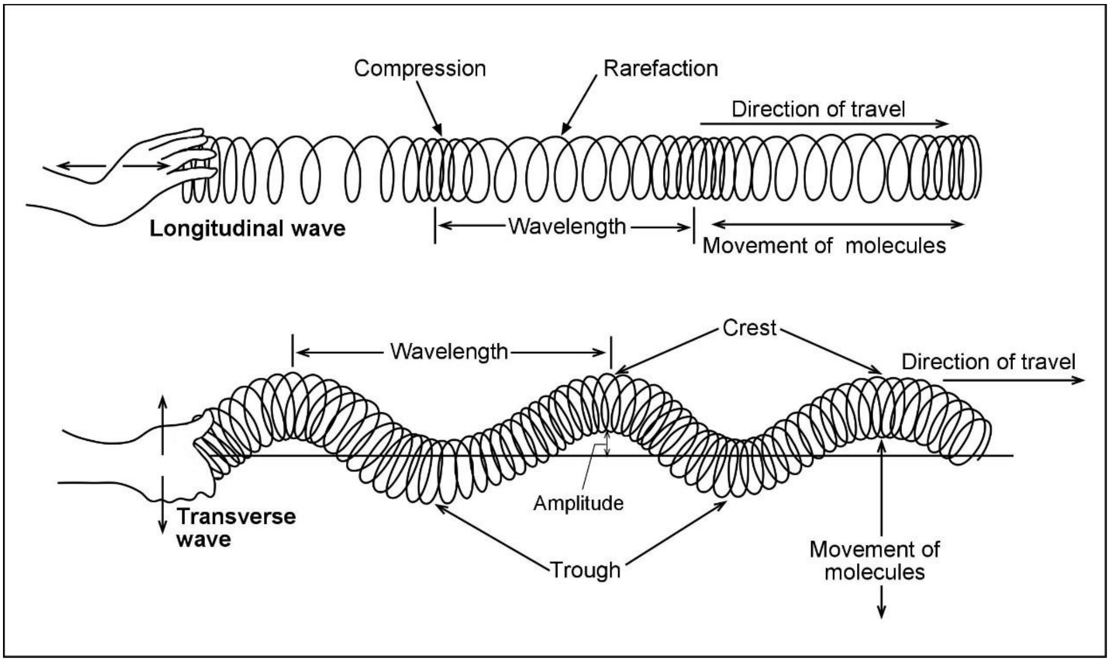


Figure 8. Transverse and longitudinal waves

Activity 3
Sound race...Where does sound travel fastest?

Objective:

At the end of the activity, you will be able to distinguish which material transmits sound the best.

Materials:

watch/clock that ticks mobile phone wooden dowel
80-100 cm long metal rod 80-100 cm long
string (1 meter) metal spoon
3 pieces zip lock bag (3x3) or waterproof mobile phone carrying case

Procedure:

1. Hold a ticking watch/clock as far away from your body as you can. Observe whether or not you can hear the ticking.

2. Press one end of the wooden dowel against the back part of the watch and the other end beside your ear. Listen very well to the ticking sound. Record your observations.

3. Repeat step #2 using a metal rod instead of the wooden dowel. Record your observations.

Q14. Did you hear the watch tick when you held it at arm's length? When you held it against the wooden dowel? When you held it against the metal rod?

4. Repeat steps #1 to #3 using a vibrating mobile phone instead. Record your observations.

Q15. Did you hear the mobile phone vibrate when you held it at arm's length? When you held it against the wooden dowel? When you held it against the metal rod?

5. Place the mobile phone in the waterproof carrying case and dip it in a basin of water while it vibrates.

Q16. Based on your observations, which is a better carrier of sound? Air or wood? Air or water? Air or metal? Water or metal?

6. At the center of the meter long string, tie the handle of the metal spoon. Hold the string at each end and knock the spoon against the table to make it ring or to create a sound. Listen to the ringing sound for a few seconds then press the ends of the strings against your ears. Observe and record the difference in sound with and without the string pressed against your ear.

7. Knock the spoon against the table. When you can no longer hear the sound of the ringing spoon, press the ends of the string against your ears. Record whether or not you could hear the ringing of the spoon again.

Q17. How did the sound of the spoon change when the string was held against your ears?

Q18. When the ringing of the spoon was too quiet to be heard through the air, could it be heard through the string?

Q19. Is the string a better carrier of sound than air?

So, through which material does sound travel fastest? Through which material did sound travel the slowest? Why does sound travel fastest in solids and slowest in air? Do you have any idea what makes sound move fast in solids?

Figure 9 shows a model for the three states of matter. Identify which is solid, liquid or gas. Now, do you have any hint why sound moves fastest in a solid medium? To give us a better picture of the differences of the three states of matter, consider worksheet 1. Then with the aid of Activity

No.4 entitled *Chimes... Chimes... Chimes...* you will be able to determine what makes solid the best transmitter of sound.

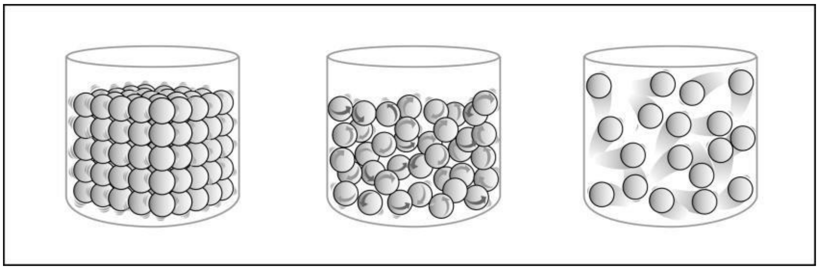


Figure 9. A model for the three states of matter

Worksheet 1: Solids, Liquids, & Gases

Direction: Using several resources and references, compare the different characteristics of solids, liquids, and gases by completing the table below.

Comparing Solids, Liquids, and Gases

Characteristics	**Solid**	**Liquid**	**Gas**
Intermolecular spacing			
Volume			
Ability to flow			
Compressibility			
Density			

Activity 4
Chimes… Chimes… Chimes…

Objective:

At the end of the activity, you will be able to infer using improvised chimes that closely spaced materials are the best transmitters of sound.

Materials:

materials for chime nylon string or thread
plastic lid or wood about 1 ½ foot long small electric
fan scissors nail and hammer beads paint iron stand

Procedure:

Improvised Chime

1. Go on a treasure hunt and look for items that will create a lovely sound when they collide, such as seashells, bells, beads, spoons, forks, and stones.

2. If the items are thin enough, poke a hole through them with a nail. Then pull a piece of string or nylon thread through each hole, and tie a knot.

3. For heavier objects, such as stones, spoons, or forks; wrap the string around the object a few times, and rub non-toxic liquid glue over the string to hold it in place.

4. Next, find a colorful plastic lid or a nice looking pieces of wood to serve as the top of the wind chime.

5. Tie at least 6 of these stringed objects on the plastic lid or on the wood. Make sure that the strands are evenly spaced and are not too far apart from each other.

6. Finally, tie another string at the two ends of the plastic lid or on the wood for hanging the chime.

Sounding the Chimes

1. Hang your chime in an iron stand where there is no wind source except your handy fan.

2. With the 6 stringed objects hanging on the wooden or plastic lid, switch on the fan and observe. This is your **CHIME 1**. Listen to the sound created by your chime. Ask one of your groupmates to move away from the chime until the sound is not heard anymore. Measure this distance from the chime to your groupmate and record your results.

3. Repeat step #2 but add 4 more stringed objects on the chimes creating chime with 10 stringed objects. Make sure that you tie the additional stringed objects in between the original ones. This is your **CHIME 2**.

Q20. With which chime did you record a longer distance?

Q21. Which chime had more stringed objects? Which chime had more closely spaced stringed objects given the same wooden lid?

4. Repeat step #2 but add 4 more stringed objects on the chime creating a chime with 14 stringed objects. This is your **CHIME 3.**

Q22. With which chime did you record the longest distance?

Q23. Which chime has the most stringed objects? Which chime has the most closely spaced stringed objects given the same wooden lid?

Q24. How would you relate the measured distance reached by the sound created by the chime and the spacing of the stringed objects in each of the 3 chimes?

Q25. Which chime is capable of transmitting sound the best?

Q26. How would you relate the distance of the stringed objects in the chime and the capability of the chime to transmit sound?

The speed of sound may differ for different types of solids, liquids, and gases. For one, the elastic properties are different for different materials. This property (elastic property) is the tendency of a material to maintain its shape and not deform when a force is applied to the object or medium. Steel for example will experience a smaller deformation than rubber when a force is applied to the materials. Steel is a rigid material while rubber can easily deform and is known as a flexible material.

At the molecular level, a rigid material is distinguished by atoms and/or particles with strong forces of attraction for each other. Particles that quickly return to their rest position can vibrate at higher speeds. Thus, sound can travel faster in mediums with higher elastic properties (like steel) than it can through solids like rubber, which have lower elastic properties.

Does the phase of matter affect the speed of sound? It actually has a large impact upon the elastic properties of a medium. Generally, the bond strength between particles is strongest in solid materials and is weakest in gases. Thus, sound waves travel faster in solids than in liquids, and faster in liquids than in gases. While the density of a medium also affects the speed of sound, the elastic properties have a greater influence on wave speed. Among solids, the most rigid would transmit sound faster. Just like the case of wood and metal in Activity 3.

What other factors may affect the speed of sound in a medium? What about temperature? Can the temperature of the medium affect how sound moves? Find out in the next activity.

Activity 5
Faster sound… In hotter or cooler?

Objective:

At the end of the activity, you will be able to determine how temperature affects the speed of sound.

Materials:

3 pieces 1000 mL graduated cylinders or tall containers thermometer
bucket of ice
electric heater or alcohol lamp tuning fork

Procedure:

1. Label the 3 graduated cylinders with HOT, ROOM TEMP, COLD respectively.

2. Half-fill the ROOM TEMP graduated cylinder with tap water.

3. Sound the tuning fork by striking it on the sole of your rubber shoes and hold it on top of the graduated cylinder.

4. When no loud sound is produced increase the amount of water up to a level where loud sound is produced when the vibrating tuning fork is placed on top. Note this level of water.

5. Fill the HOT graduated cylinder with hot water (about 70°C) to the same level as that of the ROOM TEMP cylinder.

6. Fill the COLD graduated cylinder with COLD water (about 5°C) at the same level as that of the ROOM TEMP cylinder.

7. Determine the temperature of the water in each of the cylinders just before sounding the tuning fork.

8. Sound the tuning fork in each of the cylinders and note the sound produced by each cylinder. Record all your observations.

9. Do this for three trials focusing on the differences in the pitch of the sound each cylinder creates. Record all your observations.

Q27. Which cylinder gave the loudest sound?

Q28. Which cylinder gave the highest pitched sound?

Q29. If pitch is directly dependent on frequency, then, which cylinder gives the highest frequency sound?

Q30. Since wave speed is directly dependent on frequency, then, which cylinder gives the fastest sound?

Q31. How would you relate the temperature of the medium with the speed of sound?

Now you know that the speed of sound is directly affected by the temperature of the medium. The hotter the medium the faster the sound travels. Heat, just like sound, is a form of kinetic energy. At higher temperatures, particles have more energy (kinetic) and thus, vibrate faster. And when particles vibrate faster, there will be more collisions per unit time. With more collisions per unit time, energy is transferred more efficiently resulting in sound traveling quickly.

Sound travels at about 331 $\frac{m}{s}$ in dry air at 0° C. The speed of sound is dependent on temperature of the medium where an increase is observed with an increase in temperature. This means that at temperatures greater than $0^{o}C$ speed of sound is greater than 331 $\frac{m}{s}$ by an amount 0.6 $\frac{m/s}{C}$ of the temperature of the medium. In equation,

$$v = 331\frac{m}{s} + 0.6\frac{m/s}{C}(T)$$

where T is the temperature of air in Celsius degree and 0.6 $\frac{m/s}{C}$ is a constant factor of temperature. Let's try it out at a room temperature of 25^{o}Celsius.

Sample Problem

roblem: What is the speed of sou
5

olution:

°Celsius?

Given: T = 25°Celsius Equa
$v = 331\frac{m}{s} + 0.6\frac{m/s}{C}(T)$

Solution:

$$v = 331\frac{m}{s} + 0$$
$$v = 331\frac{m}{s} + 1$$
$$v = 346\frac{m}{s}$$

Properties of Sound

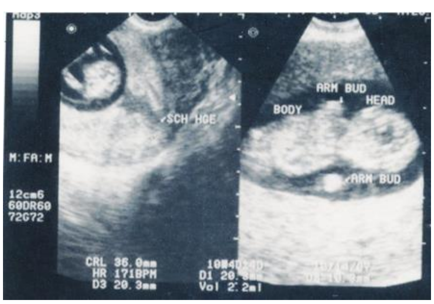


Figure 10. Ultrasound image


Figure 11. Live concert

Figures 10 and 11 are the amazing contribution of sound to other fields such as health, wellness and the arts particularly the music industry. We can experience or observe these as consequences of what are commonly called properties of sound waves. Ultrasound works on the principle of reflection of sound waves while concerts in open field benefit from refraction of sound. Want to know more about these amazing sound treats?

- **Reflection of Sound**

Figure 12. Bathroom singing

A lot of people love to sing inside the bathroom because of privacy. A study conducted noted that people would open their mouths wide when they sing in private places like the baths. Another reason is the hard wall surfaces of the bathroom usually made of wood or tiles brings about multiple reflection of sound. These hard walls or surfaces and the small dimension of the bathroom typically create an aurally pleasing acoustic environment with many echoes and reverberations contributing to the fullness and depth of voice. Well, this may not be the effect in the outside world though.
Look at Figure 12 and try it yourself.

Just like any other wave, sound also exhibits reflection. Reflection is usually described as the turning back of a wave as it hits a barrier. **Echo** is an example of a *reflected sound*. **Reverberation** on the other hand refers to the *multiple reflections or echoes in a certain place*. A reverberation often occurs in a small room with height, width, and length dimensions of approximately 17 meters or less. This best fits the bathroom which enhances the voice.

In theaters and movie houses, there are also reverberations and echoes. But these are not pleasing to the ears during a play or a movie. To lessen these, designers use curtains and cloth cover for the chairs and carpets. Check out the different movie houses and look for features inside that decreases reverberations and echoes.

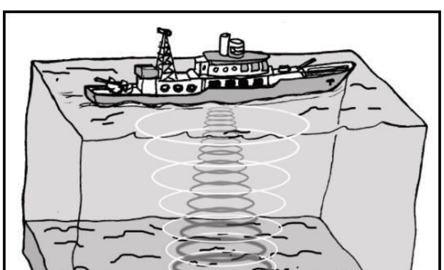
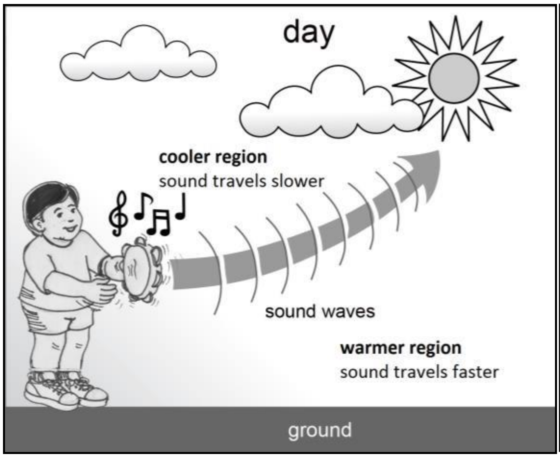


Figure 13. SONAR

Echo sounding is another application of sound reflection. This is used by scientists to map the sea floor and to determine the depth of the ocean or sea. This is just the same as how bats use sound to detect distances. What about you, can you identify other applications of sound reflection?

- **Refraction of Sound**

Have you ever wondered why open field concerts are usually held during nighttime? Having concert at night gives a chance for everyone to see and enjoy the live show because there is no work and no school. Sound also contributes to this scheduling of concerts. Usually, sound is heard better in far areas during nighttime than during daytime. This happens due to what is known as refraction. Refraction is described as the change in speed of sound when it encounters a medium of different density. As what you had earlier in this module, sound travels faster in hotter media. This change in speed of sound during refraction is also manifested as sort of “bending” of sound waves.



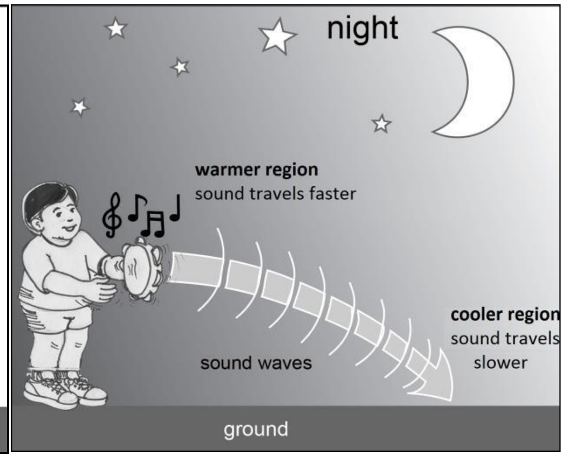


Figure14. Sound refraction at day time Figure15. Sound refraction at night time

When sound propagates in air, where the temperature changes with altitude, sound bends towards the hotter region. In this case, refraction happens. The refraction is due to the different refractive indices of air because of the difference in temperature. At daytime, when the sun is shining, the air near Earth’s surface is cooler than the air above. From what you encountered in Activity 5, you learned that sound travels faster in hotter medium. Since Earth’s surface is cooler than air above during daytime, then sound would move from the cooler region (Earth surface) towards the hotter air above. Thus, sound waves will be refracted to the sky (Figure 14). At night time, the air near the Earth’s surface is heated by the heat emitted by the ground, making it hotter than the air above which is cooler due to the absence of the sun during nighttime. This makes sound move from the cooler air above towards the hotter air near the earth’s surface. Thus, sound waves are refracted to the Earth’s surface (Figure15). This makes open field concerts better done during nighttime as sound waves are refracted from the stage towards the audience. This gives a clearer and more audible music to enjoy.

Now on a more concrete sense let us try to observe how longitudinal waves reflect and refract. In Activity No. 6, you will be able to observe how reflection and refraction are exhibited by longitudinal waves using our metal slinky.

Activity 6
Reflecting and refracting sound

Objective:

At the end of the activity, you will be able to observe how longitudinal waves reflect and refract.

Materials:

metal slinky (large coil) metal slinky (small coil)

Procedure:

Sound Reflection

1. Connect the *fixed end* to a wall or post. Make or create longitudinal waves by pushing and pulling the *movable end* part.

2. Observe the longitudinal waves as the waves hit the wall or post. Record your observations.

3. Note the positions of the compressions before they reach the post. Note also the locations or positions of the compressions after hitting the wall of the post.

4. Do this for 3 trials.

Q32. What happens to the compressions or rarefactions when they hit the wall or a fixed end?

Q33. Are the compressions found on the same location in the slinky before and after hitting the wall?

Q34. What happens to sound waves when they hit a fixed end or the wall?

Sound Refraction

1. Connect the *fixed end* of the metal slinky (small coil) to a wall or post. Then connect another slinky (large coil) to the other end of the small coil. Make or create longitudinal waves by pushing and pulling the *movable end* of the metal slinky (large coil).

2. Observe the longitudinal waves as the waves move from the large coil-metal slinky to the small coil metal slinky. Record your observations.

3. Observe the frequency, amplitude, and speed of the longitudinal waves as the waves move from the large coil metal slinky to the small coil metal slinky.

4. Do this for 3 trials.

Q35. What happens to the frequency of the longitudinal waves as the waves move from the large coil slinky to the small coil slinky?

Q36. What would be an observable change in sound when the frequency changes?

Q37. What happens to the amplitude of the longitudinal waves as the waves move from the large coil slinky to the small coil slinky?

Q38. What happens to sound when the amplitude of the sound changes?

Q39. What happens to the speed of the longitudinal waves as the waves move from the large coil slinky to the small coil slinky?

Summary

Sound waves are examples of longitudinal waves. They also exhibit characteristic features such as frequency, amplitude, wavelength, period and wave speed. The crest and the trough, however, are synonymous to compressions and rarefactions. These compressions and rarefactions are created when the particles of the medium are alternately pushed and pulled. The alternate pushing and pulling mechanically exerts force on unit areas of air particles and thus creating pressure waves. Compressions form when air particles or molecules of the medium are pushed creating lesser distance between particles, while rarefactions occur when the particles are somewhat pulled away from other particles creating a wider distance between particles. This alternating compressions and rarefaction make up the longitudinal waves like sound waves.

Just like other waves, the speed of a sound wave is determined by taking the product of the frequency and the wavelength. Speed of sound however is dependent on factors such as density and elasticity of the medium and temperature. The more elastic the medium is the faster the sound travels. Likewise, a direct relation is observed between temperature and sound speed.

Properties of waves such as reflection and refraction are also evident in sound waves. Reflected sound is known as an echo while repeated echo in a small dimension space or room is called reverberation. Change in speed resulting to bending of sound or refraction are usually observed with changes in temperature at certain altitude. What about transverse waves like light? Can we also observe these properties? Let's find out in the next module!

Links

Cheung Kai-chung (Translation by Yip Ying-kin). (n.d.). *Why do sound waves transmit farther at night? Is it because it is quieter at night?* Retrieved from http://www.hk-phy.org/iq/sound_night/sound_night_e.html

Gibbs, K. (2013). *The refraction of sound in hot and cold air*. Retrieved from http://www.schoolphysics.co.uk/age11-14/Sound/text/Refraction_of_sound/index.html

Suggested time allotme