Introduction to Classical Mechanics

Salah Nasri

[a]United Arab Emirates University,
Al-Ain, UAE

E-mail: snasri@uaeu.ac.ae

Abstract. These are my lecture notes on classical mechanics based on the courses of General Physics I (PHYS 105) and Analytical Mechanics (PHYS 262) that I have taught at UAEU during the Spring of 2012, Fall of 2015, and the Spring of 2016. I also included some advanced topics that can be of interest to Physics major students.

Most of the figures in these notes are produced by the author, unless indicated.

Contents

1 Newtonian Mechanics of Point-like Objects

1.1 Newton's Law

Newton's three laws of motion are stated as follows:

1. **First law:**
"A body remains at rest or in uniform motion unless acted upon by a force".

2. **Second law:**
A body acted upon by a force moves in such a manner that the time rate of change of its linear momentum equals the force.

3. **Third law:**
If two bodies exert forces on each other, these forces are equal in magnitude and opposite in direction.

The first and the third laws are pretty easy to understand. However, the second law needs more details and so I would like to elaborate a little bit more on it.

A material body can considered to consist of tiny particles and the state of each particle at an instant t is completely defined by

$$(\boldsymbol{r}(t)\ ;\ \boldsymbol{p}(t)) \tag{1.1}$$

Here $\boldsymbol{r}(t)$ represents the vector position of the particle at time t, and $\boldsymbol{p}(t)$ is its momentum at the same instant t. For a non-relativistic particle $\boldsymbol{p}(t)$ reads

$$\boldsymbol{p}(t) = m\boldsymbol{v}(t) \tag{1.2}$$

If the state of the particle is given at time $t = 0$ by $(\boldsymbol{r}_0; \boldsymbol{p}_0)$, then, according to Newton's second law, one can predicts its state at any other instant t by solving the equation

$$\boldsymbol{F}[\boldsymbol{r}(t), \dot{\boldsymbol{r}}(t)] = \frac{d\boldsymbol{p}}{dt} \tag{1.3}$$

where $\boldsymbol{F}$ is the net force that the particle experiences when is in the position $\boldsymbol{r}(t)$ at time t. In general $\boldsymbol{F}$ can depends on both position as well the velocity of the particle[1]. Note that if the mass of the particle remains unchanged, i.e. $dm/dt = 0$, then the force is proportional to the particle's acceleration:

$$\boldsymbol{F} = m\boldsymbol{a} \tag{1.4}$$

[1] For example, the force of air resistance is proportional to v^2. Another example is the force that a particle of charge q moving with velocity $\boldsymbol{v}$ in the presence of a magnetic field $\boldsymbol{B}$ is $\boldsymbol{F} = q(\boldsymbol{v} \times \boldsymbol{B})$, known as the Lorentz force, which depends on the particle's velocity.

1.2 Inertial vs gravitational mass

One of the fundamental forces in nature is the gravitational force between massive objects, deduced by Isaac Newton. It says that the force exerted on an object 1 of mass m_1 by another object of mass m_2, denoted by $\boldsymbol{F}_{12}$ is given by

$$\boldsymbol{F}_{12} = -Gm_1m_2\frac{(\boldsymbol{r}_1 - \boldsymbol{r}_2)}{|\boldsymbol{r}_1 - \boldsymbol{r}_2)|^3} \tag{1.5}$$

which is known as **Newton's universal law of gravitation**, with G is Newton's gravitational constant[1], measured by **Cavendish in 1798** (fpr more details, see Appendix 10.1). The overall minus sign in (1.5) indicates that it is an attractive force.

For example, let us consider an object of mass m near the surface of the Earth of mass $M_\oplus$, i.e. $(\boldsymbol{r}_1 - \boldsymbol{r}_2) \simeq R_\oplus\, \hat{r}$, with $\hat{\boldsymbol{r}}$ is a unit vector directed from the center of the Earth to the location of the the mass m, and $R_\oplus$ is the radius of the Earth. Then, the force that the Earth exerts on the object is

$$\boxed{\boldsymbol{F} = -m\boldsymbol{g}, \qquad \text{with} \qquad \boldsymbol{g} = \frac{GM_\oplus}{R_E^2}\,\hat{\boldsymbol{r}}} \tag{1.6}$$

This implies, according to Newton's second law, that

$$\boldsymbol{a} = -\boldsymbol{g} \tag{1.7}$$

The above relation between the particle's acceleration and the acceleration of gravity assumes that the mass that enters into the definition of of momentum (and hence in Newton's second law), and the mass that enters the force law of gravitation. Such a case is called the Weak Equivalence Principle (WEP)[2]. However, it is not clear a priori why these two kind of masses should be the same. If we denote the inertial mass by m_I, and the gravitational mass by m_g, then we have

$$\boxed{\boldsymbol{a} = -\frac{m_g}{m_I}\,\boldsymbol{g}} \tag{1.8}$$

This means that **unless** the ratio of the gravitational mass to the inertial mass, m_g/m_I, is **the same for all objects**, bodies would fall with different acceleration depending on the material that they are made out of. However, experiments showed that bodies in vacuum fall to Earth at the approximately the same rate. For instance, Galileo used a variety of materials for his test masses by rolling them down inclined tables and also using pendulum experiments[3], he showed that m_g/m_I is equal to unity with

[1] $G = 6.67 \times 10^{-11}$ N. m^2/kg^2.

[2] The WEP was crucial in the formulation of Einstein theory of general relativity.

[3] The Newton equation of motion reads $m_g g \sin\theta = m_I l\ddot{\theta}$, where l is the length of the pendulum. Then, for small amplitudes of the angle θ, the oscillation frequency reads $\omega = \sqrt{\frac{m_g}{m_I}\frac{g}{l}}$.

a precision of about 2%. Later, Newton improved on Galileo's pendulum experiments with a precision of about one per thousand.

A highly accurate method uses a sensitive torsion balance where the deviation from the WEP is described by the so-called **Eötvös parameter**

$$\eta_{A,B} = \frac{(m_g/m_I)_A - (m_g/m_I)_B}{(m_g/m_I)_A + (m_g/m_I)_A} \tag{1.9}$$

where A and B denotes the two bodies that are under the effect of the gravitational field[1]. The results of this type of experiments showed that the gravitational mass of a body is identical to its inertial mass to an accuracy of about one part in 10^{13} which is a clear evidence for the equivalence of the inertial and the gravitational mass[2]. It is expected that future experiments based on can test the equality between the inertial and gravitational mass to one part in 10^{17} using atomic interferometry (For more details see reference[3]. The assertion of the exact equality of the gravitational and the inertial mass goes under the name of the **equivalence principle**.

For the case of two extended objects, A and B, we can decompose the objects into an infinite number of point-like masses δm_a and δm_b, and use the fact that the superposition principle to write the gravitational force acting on B due A as

$$\boldsymbol{F}_{\mathrm{BA}} = -G \sum_{a,b=1}^{\infty} \frac{\delta m_a \delta m_b}{r_{ab}^2}\, \hat{\boldsymbol{r}}_{ab} \tag{1.10}$$

where r_{ab} is the distance form δm_a to the location of the mass δm_b, and $\hat{\boldsymbol{r}}_{ab}$ is a unit vector along this distance. For a continuous distribution of points with mass densities ρ_A and ρ_B for A and B, respectively, the above sum can be replaced by an integral over the volumes $\mathcal{V}_A$ and $\mathcal{V}_B$ of the two objects, i.e.[4]

$$\boldsymbol{F}_{\mathrm{BA}} = -G \int_{\mathcal{V}_A} \int_{\mathcal{V}_B} \frac{(\boldsymbol{r}'_A - \boldsymbol{r}'_B)}{|\boldsymbol{r}'_A - \boldsymbol{r}'_B|^3}\, \rho_A(\boldsymbol{r}'_A)\rho_B(\boldsymbol{r}'_B)\, d^3\hat{\boldsymbol{r}}'_A d^3\hat{\boldsymbol{r}}'_B \tag{1.11}$$

Here $d^3\hat{\boldsymbol{r}}'_A$ and $d^3\hat{\boldsymbol{r}}'_B$ denote the infinitesimal elements of volume of A and B, respectively, $\boldsymbol{r}'_A$ and $\boldsymbol{r}'_B$ are their vector position with respect to some coordinate system.

if B is a point-like object of mass m_B located at point P at the position $\boldsymbol{r}_P$, then, the (1.13) reads

$$\boldsymbol{F}_{\mathrm{BA}}(P) = -m_B\, G \int_{\mathcal{V}_A} \frac{(\boldsymbol{r}'_A - \boldsymbol{r}_P)}{|\boldsymbol{r}'_A - \boldsymbol{r}_P|^3}\, \rho_A(\boldsymbol{r}'_A)\, d^3\hat{\boldsymbol{r}}'_A \tag{1.12}$$

Note $\boldsymbol{F}_{\mathrm{BA}}$ is m_B times a vector quantity that depends only on density, size, and the position of B, but not m_B. So the point-like mass m_B can be considered as a test mass

[4]For a one or two dimensional object, the mass distribution os defined by a linear or surface density, and in this case the integral is over the length or the surface of the object, respectively.

under the in a gravitational vector field created by A at the point P where the B is located. Thus,

$$\boldsymbol{F}_{\mathrm{BA}}(P) = m_{\text{test}}\ \boldsymbol{g}_A(P) \tag{1.13}$$

with

$$\boldsymbol{g}_A(P) = -G \int_{\mathcal{V}_A} \frac{(\boldsymbol{r}'_A - \boldsymbol{r}_P)}{|\boldsymbol{r}'_A - \boldsymbol{r}_P|^3}\ \rho_A(\boldsymbol{r}'_A)\ d^3\hat{\boldsymbol{r}}'_A \tag{1.14}$$

The calculation of the gravitational field of an extended object can be cumbersome, even for systems with axial or even spherical symmetry. We will see later that one can overcome this difficulty by finding first the potential then deduce the gravitational vector field.

1.3 Inertial frames

Newton's laws of motion[2], are valid only in **inertial frames**, also called Galilean frames. These are frames that are not accelerating nor rotating.

Let us consider two inertial frames $\mathcal{S}$ and $\mathcal{S}'$, such that $\mathcal{S}'$ translates with a constant velocity, $\boldsymbol{u}$, relative to $\mathcal{S}$, as shown in the figure below. Then, the vector position of an object and the time when it is measured by an observer in $\mathcal{S}'$ are given by

$$\boldsymbol{r}' = \boldsymbol{r} - \overrightarrow{OO'}; \qquad t' = t \tag{1.15}$$

The space coordinates and time $(\boldsymbol{r}, t)$ and $(\boldsymbol{r}', t')$ are called events in the reference frame $\mathcal{S}$ and $\mathcal{S}'$, respectively. Taking the time derivatives of both sides of the first equation yields

$$[\boldsymbol{v}']_{\mathcal{S}'} = [\boldsymbol{v}']_{\mathcal{S}} - \boldsymbol{u} \tag{1.16}$$

where $[\vec{v'}]_{\mathcal{S}}$ denotes the velocity of an event in the reference frame $\mathcal{S}$, and $[\boldsymbol{v}']_{\mathcal{S}'}$ is the velocity of the same event viewed from the reference frame $\mathcal{S}'$. The above relation is called the **Galilean transformation of velocities**.

Since $\boldsymbol{u}$ is constant, differentiating (1.16) with respect to time gives

$$\left[\frac{d\boldsymbol{p}}{dt}\right]_{\mathcal{S}} = \left[\frac{d\boldsymbol{p}}{dt}\right]_{\mathcal{S}'} \tag{1.17}$$

Since the net force that is exerted on an object is the same in both reference frames, i.e. $[\boldsymbol{F}]_{\mathcal{S}} = [\boldsymbol{F}]_{\mathcal{S}'}$[3], then using (1.17) we get

$$[\boldsymbol{F}]_{\mathcal{S}} = \left[\frac{d\boldsymbol{p}}{dt}\right]_{\mathcal{S}}; \quad [\boldsymbol{F}]_{\mathcal{S}'} = \left[\frac{d\boldsymbol{p}}{dt}\right]_{\mathcal{S}'} \tag{1.18}$$

[2] In addition to the second law, there is the first law of inertia and the third law of action and reaction.

[3] If the cartesian axes of $\mathcal{S}'$ remain parallel to the ones of $\mathcal{S}$, then the numerical values of the components of the force in the two reference frames are equal.

Therefore, the Galilean transformation of the velocities leaves Newton's equation of motion unchanged in its mathematical form. In particular, if $[\boldsymbol{F}]_{\mathcal{S}} = 0$, then $[\boldsymbol{p}]_{\mathcal{S}} = [\boldsymbol{p}]_{\mathcal{S}'} = \overrightarrow{\text{Constant}}$. So with this we define an inertial frame as follows:

An inertial frame is any frame of reference in which a free particle remains at rest or moves in straight line.

Now let us ask the following question:

"What are the most general coordinate transformations (which is equivalent to a change of frame of reference) that leave the form of Newton's law of motion invariant[4]? ".

To answer this question let $\mathcal{S}$ be an inertial reference frame endowed with cartesian system of coordinates defined by an orthonormal basis $\{\boldsymbol{e}_i, i = 1, 2, 3\}$ and where a point is represented with a triplet (x^1, x^2, x^3). Newton's equations describing the motion of of a point-like object read

$$\frac{dp^i}{dt} = F^i, \quad i = 1, 2, 3 \tag{1.19}$$

where p^i and F^i are the cartesian components of the momentum and the net force that the particle experiences, respectively.
In an other frame of reference, say $\mathcal{S}'$, also endowed with a cartesian system of coordinates, the coordinates of the object in $\mathcal{S}'$ are related to the ones viewed by an observer in $\mathcal{S}$ by the transformation

$$x'^i = x'^i(\boldsymbol{r}, t); \qquad t' = t + \tau \tag{1.20}$$

Here τ is a constant with unit of time , such that $dt' = dt$. In other words, the clocks in both frame of reference tick in identical way.

We say that Newton's law of motion for an observable in $\mathcal{S}'$ remains the same if the components of the force and the rate change of the momentum take the form

$$\frac{dp'^i}{dt} = A^i{}_j \frac{dp^i}{dt} \qquad \text{and} \qquad F'^i = A^i{}_j F^j \tag{1.21}$$

For instance, if the particle is free in the reference frame $\mathcal{S}$, i.e. $F^i = 0$, then it remains

[4] Form invariant also called Covariant.

free in $\mathcal{S}'$ (since $F'^i = A^i{}_j \times 0 = 0$). So let us find which set of transformations can lead to (1.21). To compute the velocity and the acceleration of the particle we use the chain rule

$$\begin{aligned}\dot{x}'^i &= \frac{\partial x'^i}{\partial t} + \frac{\partial x'^i}{\partial x^j}\, x^j \\ \ddot{x}'^i &= \frac{\partial x'^i}{\partial t^2} + 2\frac{\partial^2 x'^i}{\partial t \partial x^j}\,\dot{x}^j + \frac{\partial^2 x'^i}{\partial x^j \partial x^k}\,\dot{x}^j\dot{x}^k + \frac{\partial x'^i}{\partial x^j}\,\ddot{x}^j\end{aligned} \tag{1.22}$$

If both observers choose to use cartesian coordinates system in their respective frames of reference, then x'^i can only be linear in $\boldsymbol{r}$:

$$x'^i(\boldsymbol{r}, t) = \mathcal{R}^i{}_j(t)\; x^j(t) + b^i(t) \tag{1.23}$$

This simplifies the expressions for the velocity and the acceleration in the reference frame $\mathcal{S}'$:

$$\begin{aligned}\dot{x}'^i &= \mathcal{R}^i{}_j\; \dot{x}^j + \frac{d\mathcal{R}^i{}_j}{dt}\; x^j + \frac{db^i}{dt} \\ \ddot{x}'^i &= \mathcal{R}^i{}_j\; \ddot{x}^j + 2\frac{d\mathcal{R}^i{}_j}{dt}\; \dot{x}^j + \frac{d^2\mathcal{R}^i{}_j}{dt^2}\; x^j + \frac{d^2 b^i}{dt^2}\end{aligned} \tag{1.24}$$

For a free particle, the observer in $\mathcal{S}$ observes that $\ddot{x}^i = 0$. If Newton's law of motion in $\mathcal{S}'$ has the same form as in $\mathcal{S}$, then the observer in $\mathcal{S}'$ must observe $\ddot{x}'^i = 0$. This can be achieved only if the following requirements are satisfied:

$$\begin{aligned}\{\mathcal{R}^i{}_j\} &= \text{ constants independent of time} \\ b^i &= c^i - v^i t\end{aligned} \tag{1.25}$$

with both v^i and c^i being constant. Now, if we choose $v^i = 0$ and $c^i = 0$, then the distance from the origin of $\mathcal{S}$ and $\mathcal{S}'$ to the particle should be the same in both reference frames. This implies that

$$\mathcal{R}^T\; \mathcal{R} = \mathbb{I} \tag{1.26}$$

Thus the matrices $\mathcal{R}$ are elements of the orthogonal group $O(3) \sim \mathbb{Z}_2 \times SO(3)$, where $\mathbb{Z}_2$ represents the reflexion transformation and $SO(3)$ is a rotation in 3-dimension. Since under the translations and/or rotation of coordinates the force vector is the same for the two observers, we have

$$F^j\; \boldsymbol{e}_j = F'^i\; \boldsymbol{e}'_i = F'^i\; \mathcal{R}_{ij}\boldsymbol{e}_j = \tag{1.27}$$

from which it follows

$$F^j = (\mathcal{R}^T)^j{}_i F'^i \quad \Rightarrow \quad F'^i = (\mathcal{R})^i{}_j F^j \tag{1.28}$$

In the reference frame $\mathcal{S}'$, we have

$$m\frac{d^2x'^i}{dt^2} = \mathcal{R}^i{}_j m\frac{d^2x^i}{dt^2} = \mathcal{R}^i{}_j F^j = F'^i \tag{1.29}$$

where we used the fact that in the reference frame $\mathcal{S}$ Newton's second law holds, and the relation (1.27). The above equation shows that Newton's second law takes the same form in both reference frames. Therefore, the set of transformations that **leave the form of Newton's law invariant** are

$$\boxed{x'^i = \mathcal{R}^i{}_j\, x^j - v^i t + c^i; \qquad t' = t + \tau} \tag{1.30}$$

The above set of transformation form a group under the composition law of transformations. It is called the "**Galilean group** ", which has 10 independent parameters:

$$\mathcal{R}^i{}_j \;\rightarrow\; 3 \text{ (rotations)}; \quad v^i \;\rightarrow\; 3 \text{ (boosts)} \tag{1.31}$$
$$c^i \;\rightarrow\; 3 \text{ (translations)}; \quad \tau \;\rightarrow\; 1 \text{ (time translation)} \tag{1.32}$$

Note that for Galilean boosts, the components of the net force in the two system of references are equal, i.e. $F'^i = F^i$.

For a system of particles the invariance of the force under the boosts means that it depends upon the vector distances between the particles. For instance, suppose that there are two particles, a and b, at positions vectors $\boldsymbol{r}_a, \boldsymbol{r}_b$ and with velocities $\boldsymbol{v}_a, \boldsymbol{v}_b$. Then, the force that the particle b acts on a has the form

$$\boldsymbol{F}_{ab} = \boldsymbol{F}_{ab}(\boldsymbol{r}_a - \boldsymbol{r}_b, \boldsymbol{v}_a - \boldsymbol{v}_b, t) \tag{1.33}$$

which is invariant under the boost $\boldsymbol{r}'_{a(b)} = \boldsymbol{r}_{a(b)} - \boldsymbol{v}t$ and $\boldsymbol{v}'_{a(b)} = \boldsymbol{v}_{a(b)} - \boldsymbol{v}$, i.e. $\boldsymbol{F}'_{ab} = \boldsymbol{F}_{ab}$. Further, if the force is to obey Newton's third law, then

$$\boldsymbol{F}_{ab} = (\boldsymbol{r}_a - \boldsymbol{r}_b) g_{ab}(\boldsymbol{r}_a - \boldsymbol{r}_b, \boldsymbol{v}_{\boldsymbol{a}} - \boldsymbol{v}_b, t) \tag{1.34}$$

with g_{ab} is a scalar function, i.e. invariant under rotation. Thus, the forces act along the line joining the particles, i.e. they are central. Note that Newton's 3rd law exclude the possibility to have $\boldsymbol{F}_{ab} \propto (\boldsymbol{v}_{\boldsymbol{a}} - \boldsymbol{v}_b) h_{ab}$ in the absence of external forces to the system composed of the two particles a and b. This is because in general, $\boldsymbol{F}_{ab} \neq -\boldsymbol{F}_{ba}$ if the force is proportional to $(\boldsymbol{v}_{\boldsymbol{a}} - \boldsymbol{v}_b)$.

Remarques

Although most of the forces in nature do indeed have the property $\boldsymbol{F}_{ij} = -\boldsymbol{F}_{ji}$ and with $\boldsymbol{F}_{ij} \propto (\boldsymbol{r}_i - \boldsymbol{r}_j)$, there are systems for which this is not the case. The most famous example is the Lorentz force on two moving particles with electric charge Q:

$$\boldsymbol{F}_{12} = Q\boldsymbol{v}_i \times \boldsymbol{B}_j \tag{1.35}$$

where v_i is the velocity vector of the particle 1, and $\boldsymbol{B}_2$ is the magnetic field generated by particle 2 and it is proportional to $[\boldsymbol{v}_2 \times (\boldsymbol{r}_1 - \boldsymbol{r}_2)]$. If we suppose that the two particles cross each other as shown in the figure, then

$$\boldsymbol{F}_{12} = 0, \qquad \boldsymbol{F}_{21} \propto v_2 B_1 \; \boldsymbol{e}_z \neq 0 \tag{1.36}$$

where $\boldsymbol{e}_z$ is a unit vector coming out of the page. So, in this example $\boldsymbol{F}_{12} \neq -\boldsymbol{F}_{21}$.
Does this mean that the conservation of the total momentum is violated? The answer is **No !**. The reason is that the electromagnetic field itself carries linear and angular momentum and also energy.

1.4 Terminal Velocity of Falling Object

The equation of motion for a falling object of mass m and subject to the effect of air resistance $\boldsymbol{F}_{\text{air}} = -F_{\text{air}} \; \boldsymbol{e}_y$ (we take the y-axis to be directed downward), reads

$$m\frac{d^2y}{dt^2} = mg - F_{\text{air}}(v) \tag{1.37}$$

After some time, the force gravity will be balanced by the resistive force (the drag), and, according to newton's first law, the droplet will move with a uniform velocity, calle the **terminal velocity**, and which can be determined from the equation

$$F_{\text{air}}(v) = mg \tag{1.38}$$

At low speed the resistive force can be approximated by[5]

$$F_{\text{air}} = c_1 \; v + c_2 \; v^2 \tag{1.39}$$

[5] In general, the drag force (resistive force) on a sphere of diameter D, moving at speed v though a fluid of mass density ρ_F and viscosity μ_F has the form

$$F_{\text{drag}} = \rho_F \; R^2 \; v^2 \; \mathcal{F}(\text{Re})$$

where $\text{Re} = \left(\frac{\rho_F \; v \; D}{\mu_F}\right)$ is called the **Reynolds number**. For $\text{Re} > 10^3$, the function $\mathcal{F}(\text{Re})$ is independent of Re, i.e. it is a constant, whereas for $\text{Re} << 10^3$, it is inversely proportional to Re. For more details, see my notes on where c_1 and c_2 being constants with appropriate SI units. These constants depend on the size and the shape of the object as well on the properties of the fluid in which the object is moving.

For a spherical object of radius R (in meters), F_{air} is given by the empirical expressions

$$F_{\text{air}}^{(\text{sphere})} = 3.1 \times 10^{-4}\ R\ v + 0.87\ R^2\ v^2 \tag{1.40}$$

We see that at low speed, there are two limiting cases in which the resistive force can be linear (very small R) or quadratic (for R very large) in the speed of the object. The critical value for the radius, R_c, at which the the linear and the quadratic terms in the velocity are comparable is given by

$$R_c = 36 \times \left(\frac{10\ \text{m/s}}{v}\right)\ \text{microns} \tag{1.41}$$

We normalized the speed of the falling object in units of 10 m/s as a conservative value, although, as we will see, in the presence of air resistance it is much smaller than this value. Thus, the air resistance on a droplet of water will be linear in the v, where as it quadratic for a basket ball. Below we will study each case separately

- **Terminal velocity of a falling droplet of water**

  In this case, the terminal velocity of the droplet of water is

$$v_{\text{term}}^{(\text{droplet})} = \frac{3.2 \times 10^3}{R} \times \left(\frac{4\pi}{3} R^3 \rho_{\text{H}_2\text{O}}\right)\ g \simeq 13 \times 10^8\ R^2 \tag{1.42}$$

  where we used $\rho_{\text{H}_2\text{O}} = 10^3$ kg/m^3 for the density of water. Taking the radius of a typical droplet of water to be 10 μm, we obtain

$$\boxed{v_{\text{term}}^{(\text{droplet})} \simeq 0.013\ \text{m}/s = 46.8\ \text{meters/hr}} \tag{1.43}$$

- **Terminal velocity of a basketball**

  In this case $F_{\text{air}} \propto R^2 v^2$, and we get

$$v_{\text{term}} = \sqrt{\frac{mg}{0.87 \times R^2}} \tag{1.44}$$

  using $R = 0.12$ m, and $m = 0.60$ kg for a typical baskeball, we obtain

$$v_{\text{term}}^{(\text{basketball})} \simeq 24 \text{ m}/s \simeq 86 \text{ km/hr} \tag{1.45}$$

In general, the drag force acting on a large object moving with a speed v through a fluid has the form

$$F_D = \frac{1}{2} C_D \rho_f A v^2 \tag{1.46}$$

where C_D is a positive constant, called the drag coefficient and is measured empirically, ρ_f is density of the fluid, A is the cross-sectional area of the object. Thus, the terminal velocity reads

$$v_{\text{term}} = \sqrt{\frac{2mg}{C_D \rho_f A}} \tag{1.47}$$

Thus, lighter objects reach their terminal speed before the heavier ones. That explains the reason when stone and a feather are dropped from the same height, the stone will hit the ground first.

If the falling object is spherical in shape, the terminal speed reads[6]

$$v^{(\text{sphere})} = \sqrt{\frac{4gD}{3C_D} \frac{\rho_s}{\rho_f}} \tag{1.48}$$

where D is the diameter of the spherical object, ρ_s its density, and ρ_f is the density of the fluid in which the object is moving. Thus, in this case, the terminal velocity goes like the square-root of its diameter, and so the bigger the object is the bigger is its terminal velocity.

Objects for which their is not proportional to $A^{3/2}$, the larger A is, the smaller is its terminal velocity. This is why using parachutes reduces the speed of the jumper by about %90 at the impact on the ground. For instance, the terminal velocity of a free fall person with arms stretched out (skydiver) is[7]

[6] If the buoyancy effect is not negligible, the expression of the terminal velocity of the spherical object in the fluid reads

$$v^{(\text{sphere})} = \sqrt{\frac{4gD}{3C_D} \left(\frac{\rho_s - \rho_f}{\rho_f} \right)}$$

[7] The drag coefficient is approximately $C_D^{(\text{skydiver})} \sim 02$.

$$\boxed{v_{\rm term}^{\rm (Skydiver)} \simeq \; 55 \text{ m}/s \simeq \; 200 \text{ km}/hr} \tag{1.49}$$

Of course hitting the ground at this speed will kill you[8]. However, deploying a parachute increases the surface area substantially, and in just couple of seconds the speed gets reduced to a new terminal velocity of about 5 m/s $\simeq$ 20 km/hr, which is allows you to land safely[9].

The speed of the falling object under the effect of the drag force (1.46) at any instant can be obtained by solving the differential equation (obtained from (1.37))

$$\frac{dv}{dt} = g\left(1 - \frac{v^2}{v_{\rm term}^2}\right) \tag{1.50}$$

Assuming that initially the object was at rest, we find

$$v(t) = v_{\rm term} \tanh\left(\frac{t}{\tau}\right) = v_{\rm term}\left(\frac{e^{2t/\tau} - 1}{e^{2t/\tau} + 1}\right), \qquad \tau_* = \left(\frac{v_{\rm term}}{g}\right) \tag{1.51}$$

We see that after $t = \text{few} \times \tau_*$, the speed of the object approaches its terminal velocity.

Sometimes it is useful to give the speed as function of vertical coordinate y. We choose the y-axis to be directed downward with its origin to be the point at which the object starts falling. We write

$$\frac{dv}{dt} = \frac{dv}{dy}\frac{dy}{dt} = \frac{1}{2}\frac{dv^2}{dy} \tag{1.52}$$

Using the expression of dv/dt in Eq (1.50), we obtain a differential equation for the speed with y as the independent variable, given by

$$\frac{dv^2}{dy} = 2g\left(1 - \frac{v^2}{v_{\rm term}^2}\right) \tag{1.53}$$

[8]If you skydive in a configuration such that you arms are not stretched out, you can reach a terminal velocity of about 270 km/hr.

[9] Military parachutes are designed so that you land with a speed of about 40 km/hr, which is the equivalent to jumping from a two story building which can be very harmful. That is why they have a special procedure where they hit and roll right after the impact.

The solution of the above equation is[10]

$$\upsilon(y) = \upsilon_{\rm term}^2 \left(1 - e^{-y/y_*}\right), \quad y_* = \frac{\upsilon_{\rm term}^2}{2g} \tag{1.54}$$

where we assumed that the object started falling from rest.

Example 1.1: Free fall of the moon toward the Earth

As an example of a central force we consider the gravitational force between the moon and the Earth with masses M_m and $M_\oplus$, respectively. Applying Newton's second law to each one of them gives

$$\begin{aligned} M_m \ddot{\boldsymbol{r}}_1 &= -G M_m M_\oplus \frac{(\boldsymbol{r}_1 - \boldsymbol{r}_2)}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|^3} \\ M_\oplus \ddot{\boldsymbol{r}}_2 &= -G M_\oplus M_m \frac{(\boldsymbol{r}_2 - \boldsymbol{r}_2)}{|\boldsymbol{r}_2 - \boldsymbol{r}_2|^3} \end{aligned} \tag{1.55}$$

Suppose that the moon were caused to suddenly stop rotating around the Earth, so that at a certain instant, the moon and the Earth free fall into one another under the influence of their mutual gravitational force. **How long will it take for the moon to collide with the Earth's surface?**

Adding the two equations in (1.55) yields

$$\frac{d^2}{dt^2}\left(M_m \boldsymbol{r}_1 + M_\oplus \boldsymbol{r}_2\right) = 0 \tag{1.56}$$

which implies that the point whose position is

$$\boldsymbol{r}_{CM} = \frac{M_m \boldsymbol{r}_1 + M_\oplus \boldsymbol{r}_2}{M_m + M_\oplus} \tag{1.57}$$

[10] The differential equation (1.53) can be solved by making the change of variable $u = 1 - \upsilon^2/\upsilon_{\rm term}^2$, so that the equation is written as

$$\frac{du}{dy} = -\left(\frac{2g}{\upsilon_{\rm term}^2}\right) u \ \to u(y) = u(y=0) e^{-2gy/\upsilon_{\rm term}^2} = \left(1 - \upsilon_0^2/\upsilon_{\rm term}^2\right) e^{-2gy/\upsilon_{\rm term}^2}$$

where υ_0 is the initial speed of the falling object. Thus, we find

$$\upsilon(y) = \upsilon_{\rm term}^2 \left(1 - e^{-2gy/\upsilon_{\rm term}^2}\right) + +\upsilon_0^2 \ e^{-2gy/\upsilon_{\rm term}^2}$$

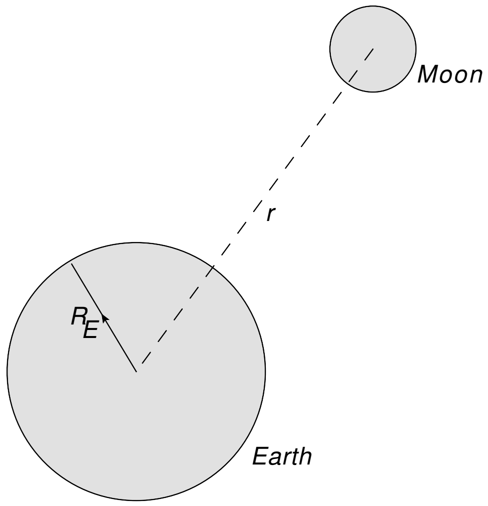


Figure 1: Free fall of the moon under the effect of Earth gravity.

moves with a constant velocity, as would a free body. This point is called "**the center of mass**" of the system Moon-Earth[6]. Now multiplying the first equation in (1.55) by $M_\oplus$ and the second one by M_m, and then subtracting the first from the second gives

$$\frac{d^2\boldsymbol{r}}{dt^2} = -G_N \frac{M}{r^2}\ \hat{r} \tag{1.58}$$

where $\boldsymbol{r} = \boldsymbol{r}_2 - \vec{r}_1$ is the relative vector position of the moon with respect to the Earth; $M = M_\oplus + M_m$ is the total mass of the Moon-Earth system , and $\hat{r} = \boldsymbol{r}/r$ is a unit vector along the distance vector from M_m to $M_\oplus$. Since there is no rotation (we have assumed that the Moon stopped turning), we have

$$\frac{d^2 r}{dt^2} = -G_N \frac{M}{r^2} \implies \frac{d}{dt}\left(\frac{1}{2}\dot{r}^2\right) = -G_N \frac{M}{r^2}\ \dot{r} \tag{1.59}$$

By integrating the above equation and assuming that for an observer on Earth the initial velocity of the Moon is zero, we find

$$\upsilon(r) = -\left[2G_N M\left(\frac{1}{r} - \frac{1}{r_0}\right)\right]^{1/2} \tag{1.60}$$

where we have assumed the minus sign reflects the fact that $dr/dt < 0$ since r is decreasing in time. Here $r_0 = r(0)$ is the Moon-Earth distance at the instant they started their free fall. Hence, the time it takes the Moon to be at a distance r from the center of the Earth is

$$t(r) = \int_{r_0}^{r} \frac{dr'}{\upsilon(r')} = -\sqrt{\frac{r_0}{2G_N M}} \int_{r_0}^{r} dr' \sqrt{\frac{r'}{r_0 - r'}} \tag{1.61}$$

[6]Note that the center of mass of a system is just a point in space, it is not itself a material body. Later in the lectures I will return to the concept of center of mass when I discuss system of particles.

Making the change of variable $r' = r_0 \sin^2\theta$, we obtain

$$t(r) = \sqrt{\frac{r_0^3}{2G_N M}}\left[\frac{\pi}{2} + \sqrt{\frac{r}{r_0}\left(\frac{r}{r_0} - 1\right)} - \sin^{-1}\left(\sqrt{\frac{r}{r_0}}\right)\right] \tag{1.62}$$

Assuming that the Earth is point-like, then at the collision $r(t_{\rm coll}) = 0$, and we obtain

$$t_{\rm coll} = \frac{\pi}{2}\left(\frac{r_0^3}{2G_N M}\right)^{1/2} \tag{1.63}$$

If we consider the orbit of the Moon just before it stopped rotating to be circular, then its period $T_{\rm Moon}$ is related to r_0 by the relation

$$T_{\rm Moon} = \frac{2\pi r_0}{\upsilon_{JB}} = 2\sqrt{2}\pi\left(\frac{r_0^3}{2G_N M}\right)^{1/2} \tag{1.64}$$

where $\upsilon_{JB} = (G_N M/r_0)^{1/2}$ is the speed of the Moon just before it stopped circulating. Thus, the time it takes the Moon to collide with the Earth is given in terms of the orbital period of the Moon[7] as

$$t_{\rm coll} = \frac{T_{\rm Moon}}{4\sqrt{2}} \simeq 5 \text{ days} \tag{1.65}$$

However, the speed of the Moon when it crashes on Earth is $\upsilon(r=0) = \infty$, which, of course, is unphysical. The reason this is because we assumed that the Earth is point like. To get a realistic estimate we take $r(t_{\rm coll}) = R_\oplus$, and in this case we find that[11]

$$\upsilon_{\rm coll} \simeq \sqrt{\frac{2G_N M}{R_\oplus}} \simeq \sqrt{\frac{2G_N M_\oplus}{R_\oplus}} \tag{1.66}$$

Using $M_\oplus \simeq 6\times10^{24}$ kg, $R \simeq 6.4\times10^3$ m, and $G_N = 6.7\times10^{-11}$ N. m^2/kg^2, we obtain

$$\upsilon_{\rm coll} \simeq 3.5\times10^5 \text{ m}/s \simeq \mathbf{10^{-3}} \times \textbf{speed of light} \tag{1.67}$$

Example 1.2 Distance from Earth to the Lagrange point L_1

In 1772, Jospeh Lagrange, a French Mathematician and Physicist, discovered that for there 5 points where the gravitational fields of the Earth and the Sun, together with the centrifugal force, cancel exactly (see Fig.**??**). Hence it would be ideal to place a satellite at a Lagrange point guaranteeing its stability with minimal energy input.

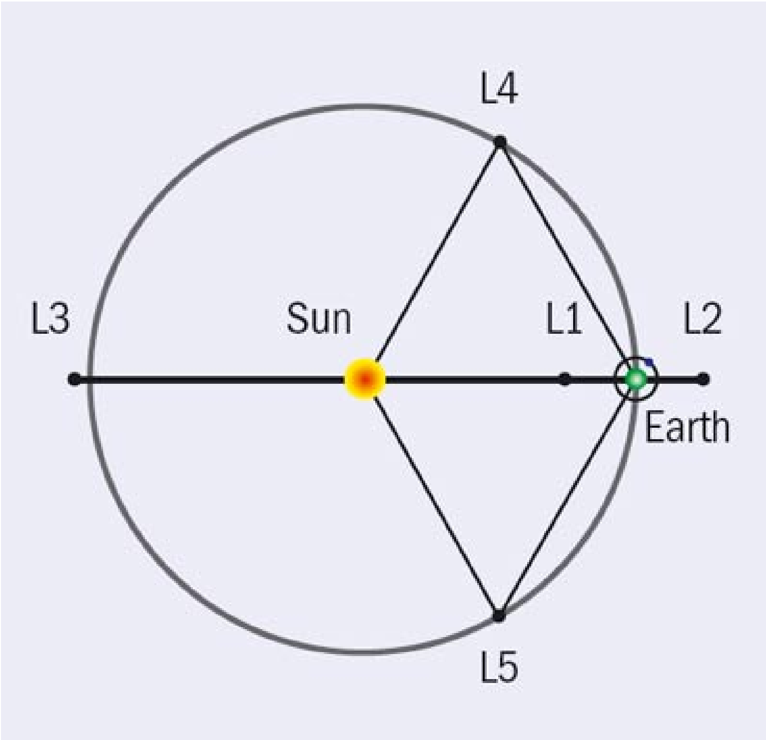


Figure 2: A schematic of the five Lagrange points for the Earth-Sun system [4].

One of these point, L_1, lies between the Earth and the Sun. In this example, we will determine the distance from this Lagrange point to the Earth.

In progress .....

Example 1.3: The sliding chain on a table

Consider a chain of length L with uniform mass density λ where $(L - z)$ of its length on table top and z of its length hanging over the edge. We would like to determine how long it takes the whole chain to go over the edge.

We will consider two cases:

- **There is no friction between the table and the chain**: Let $\mathbf{T}_{12}$ the force of tension that the segment of the chain hanging over the edge exerts on the one that is on the top table at the connecting point of the two segments, and $\mathbf{T}_{21}$ is the corresponding reaction force acting on chain hanging over the edge. But according to Newton's third law, these two forces are opposite in direction and

[7] The orbital period of the Moon is $T_{\text{Moon}} = 28$ days.

[11] The collision time is, to a very good approximation (up to order $R_\oplus/r_0$) is equal to the same value we obtained by taking $r(t_{\text{coll}}) = 0$.

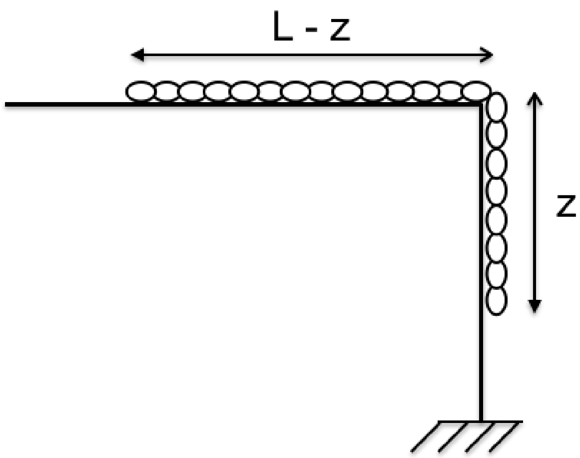


Figure 3: Sliding chain on a table.

equal in magnitude, which will denote by T_c. Then the rate of change of the momentum for each chain segment reads

$$\text{chain on the top table}: \frac{d}{dt}\left[\lambda(L-z)dz/dt\right] = T_c \tag{1.68}$$

$$\text{chain hanging over the edge}: \frac{d}{dt}\left[\lambda z dz/dt\right] = \lambda z g - T_c \tag{1.69}$$

or, equivalently

$$\text{chain on the top table}: \lambda(L-z)\frac{d^2z}{dt^2} = T_c + \lambda v^2 \equiv T(z) \tag{1.70}$$

$$\text{chain hanging over the edge}: \frac{d}{dt}\lambda z\frac{d^2z}{dt^2} = \lambda z g - T_c - \lambda v^2 \equiv \lambda z g - T(z) \tag{1.71}$$

The above system of equations is the same as the one describing the motion two objects of masses, connected by a massless string going over a smooth where one of the masses is on the plane of the table and the other other hanging over the edge. Thus, the motion of falling chain in Fig. 3 can be modeled by such a system where $m_1 = \lambda(L-z)$ and $m_2 = \lambda z$ connected by a massless rope (see Fig.4).

By adding the equations (1.70) and (1.71) we obtain the differential equation[12]

[12]We could have wrote this equation from the start using Newton's second law, $F = Ma$, where M is the total mass of the chain, i.e. λL, and $F = \lambda z g$ is the force gravity due to the weigh of the chain hanging over the edge, since the resultant of the other forces vanish (the weight of the chain on the top table and the normal force acting on it by the table cancel each other, whereas the sum of the tensions $\mathbf{T}_{12}$ and $\mathbf{T}_{21}$ vanishes by Newton's third law.)

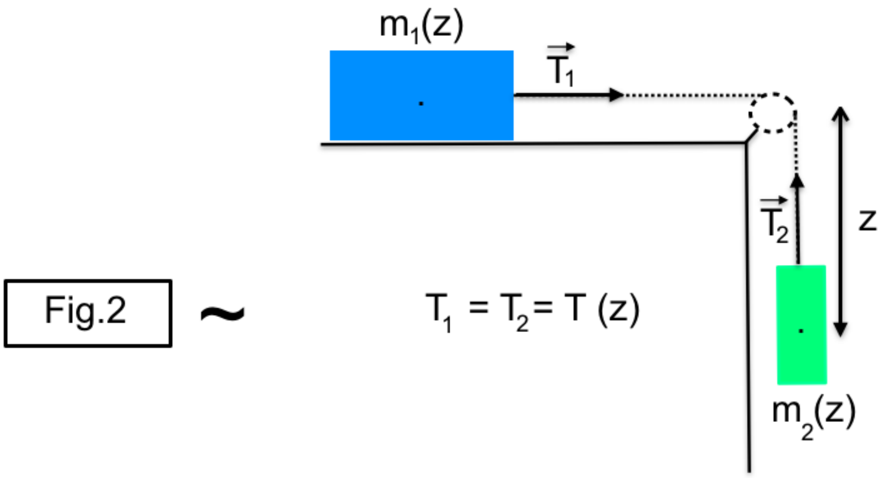


Figure 4: A model for a sliding chain on a table.

$$\frac{d^2z}{dt^2} - \frac{g}{L}z = 0 \tag{1.72}$$

which has a general solution of the form :

$$z(t) = c_1 \exp\left(\sqrt{\frac{g}{L}}t\right) + c_2 \exp\left(-\sqrt{\frac{g}{L}}t\right) \tag{1.73}$$

where c_1 and c_2 are constant of integration to be determined from the initial constant. For instant, let us assume that initially the length of the hanging chain is z_0 and that it was released from rest. Then, using these initial conditions we obtain that $c_1 = c_2 = z_0/2$. This implies[13]

$$z(t) = \frac{x_0}{2}\left[\exp\left(\sqrt{\frac{g}{L}}t\right) + \exp\left(-\sqrt{\frac{g}{L}}t\right)\right] = \frac{z_0}{2}\cosh\left(\sqrt{\frac{g}{L}}t\right) \tag{1.74}$$

To get the time t_f it takes the whole chain to go over the edge of the table, and solve $z(t_f) = L$ for the variable t_f. We obtain

$$t_f = \sqrt{\frac{L}{g}}\cosh^{-1}\left(\frac{L}{z_0}\right) \tag{1.75}$$

with $\cosh^{-1}$ is inverse function of cosh. We can write the above expression in more useful form as[14]

[13]Recall that $\cosh(\alpha) = (e^{\alpha} + e^{-\alpha})/2$.

[14]We use the identity

$$\cosh^{-1}(\alpha) = \ln\left(\alpha + \sqrt{\alpha^2 - 1}\right)$$

$$t_f = \sqrt{\frac{L}{g}} \ln\left(\frac{L}{z_0} + \sqrt{\frac{L^2}{z_0^2} - 1}\right) \tag{1.76}$$

- **There is friction between the table and the chain:**

  If κ the coefficient of friction, . then, we should add the friction force term $f_\kappa = -\kappa\lambda\,(L-x)\,g$ to the right hand side of Eq (1.70), i.e.

$$\lambda(L-z)a = T - -\kappa\lambda\,(L-z)\,g \tag{1.77}$$

  Adding the above equation to Eq (1.71), we obtain

$$\frac{d^2z}{dt^2} - \frac{g}{L}z + \kappa\frac{(L-z)}{L}g \tag{1.78}$$

  or, equivalently,

$$\frac{d^2z}{dt^2} - \frac{g}{L}\,(1+\kappa)\left[z - \frac{\kappa}{1+\kappa}L\right] \tag{1.79}$$

  By making the change variable $z' = z - \frac{\kappa}{1+\kappa}L$, the above differential equation can be brought to form

$$\frac{d^2z'}{dt^2} - \frac{g}{L}\,(1+\kappa)\,z' = 0 \tag{1.80}$$

  which has the solution

$$z'(t) = z_0' \cosh\left(\sqrt{\frac{g\,(1+\kappa)}{L}}t\right) \tag{1.81}$$

  Hence,

$$z(t) - \frac{\kappa}{1+\kappa}L = \left(z_0 + \frac{\kappa}{1+\kappa}L\right)\cosh\left(\sqrt{\frac{g\,(1+\kappa)}{L}}t\right) \tag{1.82}$$

  Setting $z(t_f) = L$ into the above equation and solving for t_f, we get

$$t_f = \sqrt{\frac{L}{g\,(1+\kappa)}}\cosh^{-1}\left[\frac{L - \frac{\kappa}{1+\kappa}L}{z_0 - \frac{\kappa}{1+\kappa}L}\right] \tag{1.83}$$

  or, equivalently, in terms of the logarithmic function as

$$t_f = \sqrt{\frac{L}{g(1+\kappa)}} \ln\left[\frac{L-\frac{\kappa}{1+\kappa}L}{z_0-\frac{\kappa}{1+\kappa}L} + \sqrt{\left(\frac{L-\frac{\kappa}{1+\kappa}L}{z_0-\frac{\kappa}{1+\kappa}L}\right)^2 - 1}\right] \tag{1.84}$$

Example 1.4: The motion of a space rocket[15]

Consider a rocket containing a combustible material that burns at some rate (see figure 1). The gases produced by the burning combustible are exhausted from the bottom of the rocket, usually at fixed rate relative to the rocket. Let $\boldsymbol{v}(t)$ be the velocity of the rocket relative to an observer on Earth, and $\boldsymbol{u}$ the velocity of the ejected gases relative to the rocket which we assume to be constant. At some instant t, the rocket together with the combustible material that remains unburned has a momentum

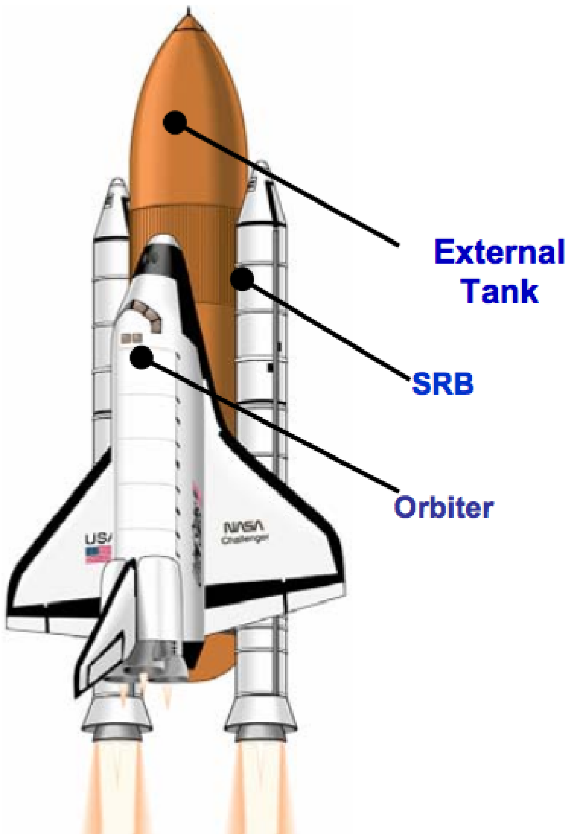


Figure 5: Typical space-rocket. Taken from from [5].

$$\boldsymbol{p}(t) = m(t)\ \boldsymbol{v}(t) \tag{1.85}$$

Then, after some time δ, an amount "$-\delta m$" of the combustible has been burned[8] and exits the rocket with a velocity $\boldsymbol{v}_g = \boldsymbol{v}(t) + \boldsymbol{u}$ while the " rocket + remaining fuel " of

[15] Was first derived in 1897 by the **Konstantin Tsiolkovsky**, a Russian rocket scientist.
[8] Note that δm is negative.

mass $(m + \delta m)$ travels upward with a velocity $(\boldsymbol{v}(t) + \boldsymbol{\delta v})$. So at the instant $(t + \delta t)$ the momentum of the "rocket + remaining fuel" is

$$\begin{aligned} \boldsymbol{p}(t+\delta t) &= (m+\delta m)(\boldsymbol{v}(t)+\boldsymbol{\delta v}) - \delta m \boldsymbol{v_g} \\ &\simeq m\boldsymbol{v}(t) + m\boldsymbol{\delta v} + \delta m(\boldsymbol{v} - \boldsymbol{v_g}) \\ &= \boldsymbol{p}(t) + m\boldsymbol{\delta v} - \delta m \boldsymbol{u} \end{aligned} \tag{1.86}$$

where we neglected the term $\delta\boldsymbol{\delta v}$. The above equation can be rewritten as

$$\delta\boldsymbol{p}(t) := \boldsymbol{p}(t+\delta t) - \boldsymbol{p}(t) = m\frac{\delta \boldsymbol{v}}{\delta t} - \frac{\delta m}{\delta t}\boldsymbol{u} \tag{1.87}$$

or, equivalently,

$$m\frac{d\boldsymbol{v}}{dt} = \frac{dm}{dt}\boldsymbol{u} + \boldsymbol{F}^{(\mathrm{ext})} \tag{1.88}$$

where by definition $\boldsymbol{F}^{\mathrm{ext}} \equiv \frac{d\boldsymbol{p}}{dt}$. Note that even in the absence of external forces there is an effective force acting on the rocket:

$$\boldsymbol{F}^{(\mathrm{eff})} = \frac{dm}{dt}\,\boldsymbol{u} \;\equiv\; \textbf{Rocket thrust} \tag{1.89}$$

Since the fuel is being consumed, the rocket continuously loosing mass, i.e. $dm/dt < 0$, and hence $\boldsymbol{F}^{(\mathrm{eff})}$ is directly opposite to the velocity of the gases relative to the rocket, which cause the rocket to rise. If we approximate $\boldsymbol{F}^{(\mathrm{ext})}$ to constant gravitational force of the Earth on the rocket, $m\boldsymbol{g}$, then by integrating both sides of the rocket equation (1.88), we obtain

$$\boldsymbol{v}(t) = \boldsymbol{v}_0 - \log\left(\frac{m_0}{m(t)}\right)\boldsymbol{u} + \boldsymbol{g}(t - t_0) \tag{1.90}$$

where $\boldsymbol{v}_0$ and m_0 are the velocity and the mass of the rocket at the initial time t_0[10]. Choosing $t_0 = 0$, and the y-axis oriented up-war, Eq (1.90) reads

$$\boxed{v(t) = v_0 + \log\left(\frac{m_0}{m}\right) u - gt} \tag{1.91}$$

The above formula was first derived in 1897 by the Konstantin Tsiolkovsky, a Russian rocket scientist, and it is known as the **Tsiolkovsky rocket equation**.

[10] A net zero acceleration is possible only if

$$\frac{dm}{dt}\boldsymbol{u} = -m(t)\boldsymbol{g} \;\longrightarrow\; m(t) = m_0\ \exp\left(-\frac{g}{u}(t - t_0)\right)$$

For instance, the **Saturn V** rocket used by NASA for the moon landings[11] (Appolo 11 mission), has exhaust speed $u = 2.91\ km/s$, the stage 1 burn-time $t_1 = 120\ s$, total mass $m_0 = 28 \times 10^5\ kg$, stage 1 propellant mass $m_P = 20 \times 10^5\ kg$. Thus, in the first stage the Saturn V rocket has

$$\begin{aligned} \Delta \upsilon^{\text{thrust}} &= u \log\left(\frac{m_0}{m_0 - m_P}\right) \simeq 3.65\ km/s \\ \Delta \upsilon^{\text{drag}} &= -g t_1 = -1.18\ km/s \end{aligned} \tag{1.92}$$

So in the first stage, the rocket speed is

$$\upsilon^{\text{stage 1}} = (5.65 - 1.18)\ km/s = \mathbf{2.47}\ \text{km}/s \tag{1.93}$$

which is smaller than the Earth escape speed which is about **11.2** km/s (see Example 1.2 for details) or even the orbital speed needed for a satellite to maintain a stable low Earth orbit (LEO) [16] This is the main reason why **multi-stage rocket** is necessary[17].

1.5 Work and kinetic energy

The work done by a net force $\boldsymbol{F}$ on a particle of mass m, moving along a trajectory $\mathcal{C}$ connecting the points $\boldsymbol{r}(t_1)$ and $\boldsymbol{r}(t_2)$ is:

$$W^{\text{net}}_{1\to 2} = \int_{\mathcal{C}} \boldsymbol{F}.d\boldsymbol{l} = m \int_{\upsilon(t_1)}^{\upsilon(t_2)} d\boldsymbol{\upsilon}.\ \boldsymbol{\upsilon}\ dt \tag{1.94}$$

which can be expressed as

$$W^{\text{net}}_{1\to 2} = T(\upsilon_2) - T(\upsilon_1) \tag{1.95}$$

where

$$\boxed{T(\upsilon) := \tfrac{1}{2} m\ \upsilon^2} \tag{1.96}$$

is called the kinetic energy of the particle. The expression (1.94) is called the **work-kinetic energy theorem**, which states that the change in the kinetic energy of a particle is equal to the work done by the net force acting upon it.

We define the quantity

$$\boxed{\mathbf{p} := \tfrac{\partial T}{\partial \upsilon_i}\ \hat{e}_i = m\ \boldsymbol{\upsilon}} \tag{1.97}$$

which is called the momentum of a particle of mass m and velocity υ.

[11] It was launched on July, 1969, from Kennedy Space Center in Merritt Island, Florida.
[16] The orbital speed of a satellite in LEO (i.e. at $500 - 1500$ km/s above earth's surface) is about km/s.
[17] For more details see my lecture notes on "**The motion of a Rocket**" posted on the link http://faculty.uaeu.ac.ae/snasri.

1.6 Conservative force

In many situations the work done by a force $\boldsymbol{F}$ between two points A and B does not depend on the path taken. That is:

$$\int_{A;\mathcal{C}_1}^{B} \boldsymbol{F}.\, d\boldsymbol{l} = \int_{A;\mathcal{C}_2}^{B} \boldsymbol{F}.\, d\boldsymbol{l} \tag{1.98}$$

where $\mathcal{C}_1$ and $\mathcal{C}_2$ are arbitrary paths joining the points A and B. If the above equation is true, then we can write

$$\oint_{\mathcal{C}} \boldsymbol{F}.\, d\boldsymbol{l} = 0, \qquad \forall \text{ closed curve } \mathcal{C} \tag{1.99}$$

In this case we say that $\boldsymbol{F}$ is a **conservative force**. Now by applying Stokes's theorem, Eq (1.99) reads

$$\oint_{\mathcal{C}_1} \boldsymbol{F}.\, d\boldsymbol{l} = \int_{\mathcal{S}} (\boldsymbol{\nabla} \times \boldsymbol{F}).\, d\overrightarrow{\mathcal{S}} = 0 \tag{1.100}$$

where $\mathcal{S}$ is the surface bounded by $\mathcal{C}$. Since $\mathcal{C}$ is arbitrary, so is $\mathcal{S}$ then it follows that

$$\boldsymbol{\nabla} \times \boldsymbol{F} = 0 \tag{1.101}$$

Hence, $\boldsymbol{F}$ is conservative if and only if it is **rotation free**[18]. In this case we can express the vector force as a gradient of a scalar function, i.e.

$$\boxed{\boldsymbol{F}^{(\text{cons})} = -\boldsymbol{\nabla}\, V(\boldsymbol{r})} \tag{1.102}$$

and $V(\boldsymbol{r})$ is called the **potential energy function** corresponding to the force $\boldsymbol{F}$. The minus sign in front of the gradient in the Eq (2.5) is conventional, where the force is directed in the direction of decreasing potentials. If $\boldsymbol{F}(\boldsymbol{r})$ is known then one can obtain, up to a constant, the associated potential energy from the equation

$$V(\boldsymbol{r}) = -\int_{\boldsymbol{r}_*}^{\boldsymbol{r}} \boldsymbol{F}.\, d\boldsymbol{r} \tag{1.103}$$

Here $\boldsymbol{r}_*$ represents some fixed point, a "reference point", of our choosing where the potential energy vanishes.

[18] In particular, in one dimension, if a force acting on an object is a function of position, it is a conservative force.

1.7 Total Mechanical Energy and Power

Now, using the work-kinetic energy theorem we can write

$$W_{1\to 2} = V(\boldsymbol{r}_1) - V(\boldsymbol{r}_2) = T(\boldsymbol{r}_2) - T(\boldsymbol{r}_1) \tag{1.104}$$

which implies that

$$V(\boldsymbol{r}_1) + T(\boldsymbol{r}_1) = V(\boldsymbol{r}_2) + T(\boldsymbol{r}_2) \tag{1.105}$$

This means that the quantity

$$E = T + V = \frac{1}{2}mv^2 + U(\boldsymbol{r}) \tag{1.106}$$

which is called the **total mechanical energy of the particle** is constant during the motion, i.e. it is conserved:

$$\Delta E = E(t_2) - E(t_1) = 0, \qquad \forall t_{1,2} \tag{1.107}$$

For a general potential $V(\boldsymbol{r}, t)$, taking the time derivative of E in (1.106) we have

$$\begin{aligned}
\frac{dE}{dt} &= m\boldsymbol{v}.\, \frac{d\boldsymbol{v}}{dt} + \frac{dU}{dt} \\
&= m\boldsymbol{v}.\, \frac{d\boldsymbol{v}}{dt} + \boldsymbol{v}.\, \boldsymbol{\nabla} V + \frac{\partial V}{\partial t} \\
&= \boldsymbol{v}.\, (m\frac{d\boldsymbol{v}}{dt} + \boldsymbol{\nabla} V) + \frac{\partial V}{\partial t} \\
&= \boldsymbol{v}.\, (m\frac{d\boldsymbol{v}}{dt} - \boldsymbol{F}) + \frac{\partial V}{\partial t} && (1.108) \\
& && (1.109)
\end{aligned}$$

which upon using Newton's second law gives

$$\boxed{\frac{dE}{dt} = \frac{\partial V}{\partial t}} \tag{1.110}$$

Thus, for conservative forces, the potential V can not depend explicitly on time. When a force does a work, we define its instantaneous power by

$$\boxed{P := \frac{dW}{dt}} \tag{1.111}$$

which is measured in units of Watt[19], denoted by W, with $1\ W = 1J/s$. If we average over a certain period of time Δt we obtain the average power

$$P_{\text{av}} = \frac{1}{\Delta t}\int_{t_0}^{t_0+\Delta t} \frac{dW}{dt}\, dt = \frac{\Delta W}{\Delta t} \tag{1.112}$$

[19]In the British system of units, the power is measured in **horse power** (HP), where 1 HP $\simeq 746\ W$.

where ΔW represents the amount of work donee during the time period Δt.
If an object moves with a velocity $\boldsymbol{v}$ under the effect of a force $\boldsymbol{F}$, then the work done during an infinitesimal time dt is

$$dW = \boldsymbol{F}.\, d\boldsymbol{l} = \boldsymbol{F}.\, d\boldsymbol{v}\; dt \tag{1.113}$$

which implies that

$$\boxed{P = \boldsymbol{F}.\, \boldsymbol{v}} \tag{1.114}$$

Taking the time derivative of the kinetic energy for a body of mass m, yields

$$\frac{dK}{dt} = 2\frac{1}{2}m\frac{d\boldsymbol{v}}{dt}.\, \boldsymbol{v} \tag{1.115}$$

or, equivalently,

$$\boxed{\frac{dK}{dt} = \boldsymbol{F}_{\text{net}}.\, \boldsymbol{v} = P_{\text{net}}} \tag{1.116}$$

where we used Newton's second law $m\frac{d\boldsymbol{v}}{dt} = \boldsymbol{F}_{\text{net}}$, with $\boldsymbol{F}_{\text{net}}$ being the net force. Hence, the time rate of change of the kinetic energy is equal to the power of the net force applied to the object.

Example 1.5 Two Important Conservative Forces

Here are we discuss two examples of conservative systems:

1. **Simple harmonic oscillator**
   The restorative force exerted by a spring, $\boldsymbol{F} = -k\boldsymbol{r}$, with k a constant which has unit of force per unit length, is a conservative force since

$$\boldsymbol{\nabla} \times \boldsymbol{F} = -k\; \boldsymbol{\nabla} \times \boldsymbol{r} = 0 \tag{1.117}$$

   Hence,

$$V(\boldsymbol{r}) = -\int_{\boldsymbol{r}_*}^{\boldsymbol{r}} \boldsymbol{F}.\, d\boldsymbol{r} = \frac{1}{2}kr^2 \tag{1.118}$$

   where we have taken the reference point $\boldsymbol{r}_*$ for the potential energy. So, the energy of a simple harmonic oscillator (SHO) is

$$E_{SHO} = \frac{1}{2}mv^2 + \frac{1}{2}kr^2 = \text{ constant} \tag{1.119}$$

   In one dimension, we can solve for for the speed v and get

$$v = \sqrt{v_0^2 - \omega^2 x^2}, \quad v_0 = \frac{2E}{m}, \quad \omega = \sqrt{\frac{k}{m}} \tag{1.120}$$

   where v_0 is the speed of the particle at $x = 0$.

2. **Free fall in gravitational field**
   Consider two particles: one of mass M fixed at the origin[14] and the other one, of mass m, situated at position $\boldsymbol{r}$. Then the gravitational force acting on m due to M is

$$\boldsymbol{F}_g = -G\frac{Mm}{r^2}\,\hat{r}, \qquad \hat{r} = \frac{\boldsymbol{r}}{r} \tag{1.121}$$

   It is straightforward to check that $\boldsymbol{\nabla}\times\boldsymbol{F} = 0$, which means that the gravitational force is a conservative force. The potential energy of m_2 due to the gravitational interaction with m_1 is

$$V(\boldsymbol{r}) = -\int_{\boldsymbol{r}_*}^{r} \boldsymbol{F}.\,d\boldsymbol{r} = GMm\int^{r}_{\boldsymbol{r}_*}\frac{d\boldsymbol{r}}{r^2}.\,\hat{r} = GMm\int^{r}_{\boldsymbol{r}_*}\frac{dr}{r^2} \tag{1.122}$$

   A convenient reference point would be a point at infinity, i.e. the situation when the mass m is infinitely separated from M. In this case we have

$$V(\boldsymbol{r}) = -G\frac{Mm}{r} \tag{1.123}$$

   The total energy of the mass m is then given by

$$E = \frac{1}{2}mv^2 - G\frac{Mm}{r} = \text{ constant} \tag{1.124}$$

Example 1.6 **The Escape Velocity from the Surface of Earth**

The **escape velocity** is the minimum speed v_e that an object needs to have to break free from the surface of a large body such as moon, planet or a star. Suppose that an object of mass m leaves the surface of stationary object of mass $M >> m$, with initial speed v_0. If, for simplicity, we assume to be spherical of radius R, then the energy conservation $E(r=R) = E(r), \forall r > R$, implies that

$$\frac{1}{2}mv_0^2 - G\frac{Mm}{R} = \frac{1}{2}mv_\infty^2 \tag{1.125}$$

where v_∞ is the speed of the particle at $r=\infty$, i.e. at the point where the mass m has completely escaped the effect of the gravity due to the mass M. Since the right hand side of the above equation can not be negative, the escape velocity correspond to the value of v_0 for which $v_\infty = 0$, i.e.

$$v_e = \sqrt{2\frac{G_N M}{R}} \tag{1.126}$$

which is independent of the mass of the escaping object[20]. Applying the above formula for the escape velocity to the Earth, we obtain

[14] In general we can not assume that particle 1 is fixed to the origin. This could be a good approximation when $M >> m$.

[20] It is important to note that the above calculation neglects the effect of air resistance.

Object	$M/M_\oplus$	$R/R_\oplus$	$\upsilon_e\ (km/s)$
Sun	333×10^3	109	**617.5**
Mercury	5×10^{-2}	0.4	4.3
Venus	0.8	0.9	10.3
Earth	1	1	**11.2**
Mars	0.5	0.1	5
Jupiter	318	11	59.6
Saturn	95	9	35.6
Uranus	14.5	4	21.3
Neptune	17	4	23.8
Moon	10^{-2}	0.3	**2.4**
Titan	2×10^{-3}	0.4	2.6
Europa	8×-3	0.3	2

Table 1: The values of escape velocity of some of the astronomical objects.

$$\upsilon_e^{(\oplus)} = 6.9\ \mathrm{m}/s = 11.2\ \mathrm{km}/s \tag{1.127}$$

For completeness, in table .we give the values of the escape velocity for different planets in the solar system and some other astronomical objects[21].

Example 1.7: Tunnel through the Earth[22]

Consider a straight tunnel drilled through the Earth with length $d \leq D_\oplus$, with $D_\oplus$ is the Earth's diameter[23]. We assume that the density of the Earth is uniform and that a particle inside the tunnel can slide without friction or air resistance.
We would like to estimate the time it takes a particle of mass m to reach the other end of the tunnel. There are two ways to do that:

- **Using Newton's 2nd law:**
  The mass m is subject to a Newton universal force of gravity force F_G due to the mass $M(r)$ contained inside the sphere of radius r; that is

$$\boldsymbol{F}_g = -G\frac{M(r)m}{r^2}\ \hat{\boldsymbol{r}} = G_N\frac{M_\oplus mr}{R_\oplus^3}\ \hat{\boldsymbol{r}} \tag{1.128}$$

  where in the second equality we used the **assumption that the mass of the Earth $M_\oplus$ is uniformly distributed over its volume**. So, Newton's second law reads

$$\frac{d^2\boldsymbol{r}}{dt^2} = G\frac{M_\oplus r}{R_\oplus^3}\ \hat{\boldsymbol{r}} \tag{1.129}$$

[21] **Titan** is the Saturn's largest moon, and **Europa** is one of Jupiter's moon.
[22] The idea of the gravity tunnel was proposed by Paul Cooper in 1966 in an article entitled "Through the Earth in Forty Minutes, published in the American Journal of physics.
[23] Note that the tunnel does not necessarily passes by the Earth's center.

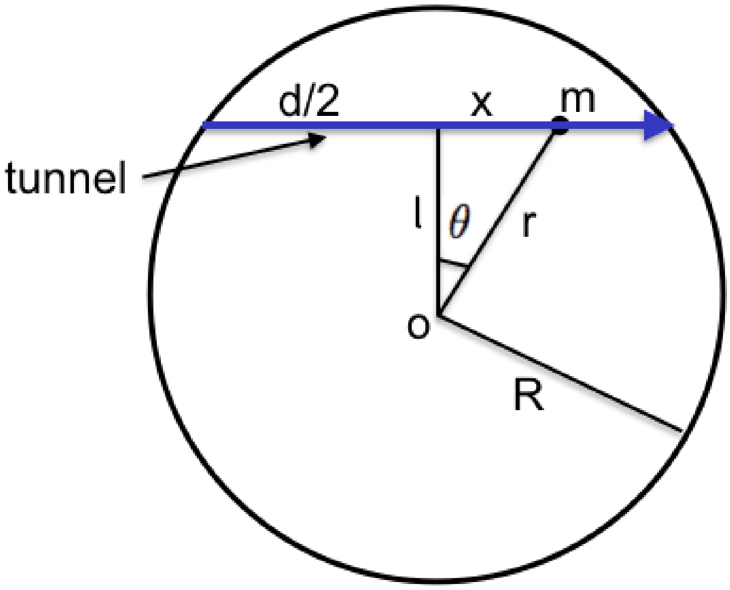


Figure 6: A tunnel through the Earth.

Decomposing the vectors in the above equation in the direction of motion (i.e. the x -axis along the tunnel), we get

$$\frac{d^2x}{dt^2} = -G\frac{M_\oplus r}{R_\oplus^3} \sin\theta \tag{1.130}$$

Using the fact that $\sin\theta = x/r$, we obtain

$$\frac{d^2x}{dt^2} + \left(G\frac{M_\oplus}{R_\oplus^3}\right) x = 0$$

which is the equation of a simple harmonic oscillator, with frequency

$$\omega = \sqrt{G\frac{M_\oplus}{R_\oplus^3}} \tag{1.131}$$

Thus, the time it takes to reach the other end of the tunnel, $\tau_{\text{one way}}$, is half the period of the oscillation, i.e.

$$\tau_{\text{one way}} = \frac{1}{2}\left(\frac{2\pi}{\omega}\right) = 2\pi\sqrt{\frac{R_\oplus^3}{GM_\oplus}} \tag{1.132}$$

Now using the values $G = 6.67 \times 10^{-11}$ m^3/kg.s^2 for the Newton constant, $M_\oplus = 5.98 \times 10^{24}$ kg for the mass of the Earth, and $R_\oplus = 6.38 \times 10^6$ m for the Earth's radius, we find

$$\tau_{\text{one way}} \simeq 42 \text{ min} \tag{1.133}$$

It is interesting to note that the time it takes to go from one end of the Earth to another is independent of the length of the tunnel.

- **Using Energy Conservation:**
This method of finding the time it takes to travel through the tunnel is much easier in the sense it does not involve vector components. We first note that the gravitational potential energy of the particle is

$$V(\boldsymbol{r}) = -\int_0^r \boldsymbol{F}_g.\, d\boldsymbol{r} = m\frac{GM_\oplus}{2R_\oplus^3}\, r^2 \tag{1.134}$$

where we have chosen the potential energy to be zero at center of the earth. In terms of the coordinate x, we have

$$V(x) = -m\frac{GM_\oplus}{2R_\oplus^3}\left(x^2 + l^2\right) \tag{1.135}$$

which is quadratic in the coordinate. Moreover, since we assumed that there are no friction or air resistance inside the tunnel, the total the energy of the mass-Earth system is conserved. Therefore, this system is exactly described by a simple harmonic oscillator with frequency given by the coefficient in front of x^2, i.e.

$$\omega = \sqrt{G\frac{M_\oplus}{R_\oplus^3}} \tag{1.136}$$

Thus we obtain the same $\tau_{\text{one way}}$ we found using Newton's law.

That we mentioned that the energy of this system is conserved, let us try to estimate the maximum speed that the particle can reach inside the tunnel. The conservation of energy equation reads

$$\frac{1}{2}mv^2(\boldsymbol{r}) + U(\boldsymbol{r}) = \text{constant} = E_{\text{initial}} \tag{1.137}$$

Since the mass m started from rest, the constant in the right hand side of (1.137) is just the potential energy of the mass at the surface of the Earth. Hence, we have

$$\frac{1}{2}mv^2(\boldsymbol{r}) + m\frac{GM_\oplus}{2R_\oplus^3}\, r^2 = m\frac{GM_\oplus}{2R_\oplus} \tag{1.138}$$

from which it follows

$$v(r) = \sqrt{\frac{GM_\oplus}{R_\oplus^3}\left(R_\oplus^2 - r^2\right)} \tag{1.139}$$

From Eq (1.139), we see that the particle reaches its maximum speed at the midpoint of the tunnel which is at distance $r_{\min}$ from the Earth center, given by

$$r_{\min} = \sqrt{R_{\oplus}^2 - \left(\frac{d}{2}\right)^2} \tag{1.140}$$

Thus,

$$\upsilon_{\max} = \sqrt{\frac{GM_{\oplus}}{4R_{\oplus}^3} d^2} \tag{1.141}$$

with the values of $G, M_{\oplus}$, and $R_{\oplus}$, we find

$$\boxed{\upsilon_{\max} \simeq 3.1 \text{ km}/s} \tag{1.142}$$

1.8 Non-conservative force

We will denote by $\boldsymbol{f}$ the non-conservative force acting on a particle of mass m, and by $\boldsymbol{F}_{\text{cons}}$ all the other forces that we assume to be conservative so that

$$V(\boldsymbol{r}) = -\int_{\boldsymbol{r}_*}^{\boldsymbol{r}} \boldsymbol{F}_{\text{cons}}.\, d\boldsymbol{r} \tag{1.143}$$

According to the work-kinetic energy theorem we have

$$\begin{aligned} \frac{1}{2}m\upsilon_2^2 - \frac{1}{2}m\upsilon_1^2 &= \int_{\boldsymbol{r}_1}^{\boldsymbol{r}_2} \boldsymbol{F}_{\text{cons}}.\, d\boldsymbol{r} + \int_{\boldsymbol{r}_1}^{\boldsymbol{r}_2} \boldsymbol{f}.\, d\boldsymbol{r} \\ &= V(\boldsymbol{r}_1) - V(\boldsymbol{r}_2) + \int_{\boldsymbol{r}_1}^{\boldsymbol{r}_2} \boldsymbol{f}.\, d\boldsymbol{r} \end{aligned} \tag{1.144}$$

which can be re-expressed as

$$\boxed{\Delta E = \int_{\boldsymbol{r}_1}^{\boldsymbol{r}_2} \boldsymbol{f}.\, d\boldsymbol{r}} \tag{1.145}$$

Non-conservative forces are generally complex interactions and can be modeled only phenomenologically. For instance, resistive forces can be parameterized by ($\kappa > 0$):

$$\boldsymbol{f} = -\kappa\, \upsilon^{\alpha}\, \hat{\upsilon} \tag{1.146}$$

where α is a real parameter, $\hat{\upsilon} = \boldsymbol{\upsilon}/\upsilon$ is a unit vector in the direction of the velocity of the particle. In particular, the force of friction between two blocks is approximately constant , which corresponds to $\alpha \simeq 0$. For a particle subject to a force of the form (1.146), the energy gets dissipated at a rate

$$\frac{dE}{dt} = -\kappa\, \upsilon^{\alpha+1} \tag{1.147}$$

It is a fundamental postulate in Physics that the **total energy of the universe** remains constant. Any loss of mechanical energy of a particle (or system of particles) implies that it has been transferred to the rest of the universe.

1.9 Angular momentum of a single particle

A particle with momentum $\boldsymbol{p}$ has angular momentum about a point $\boldsymbol{r}_0$ given by

$$\boldsymbol{L} = (\boldsymbol{r} - \boldsymbol{r}_0) \times \boldsymbol{p} \tag{1.148}$$

Very often we chose the point $\boldsymbol{r}_0$ at the origin, that is $\boldsymbol{r}_0 = \boldsymbol{0}$. The time derivative of the angular momentum vector is

$$\frac{d\boldsymbol{L}}{dt} = \frac{d\boldsymbol{r}}{dt} \times \boldsymbol{p} + \boldsymbol{r} \times \frac{d\boldsymbol{p}}{dt} = \frac{1}{m}\boldsymbol{p} \times \boldsymbol{p} + \boldsymbol{r} \times \boldsymbol{F}$$

Hence,

$$\frac{d\boldsymbol{L}}{dt} = \boldsymbol{r} \times \boldsymbol{F} \equiv \boldsymbol{\tau} \tag{1.149}$$

where $\boldsymbol{\tau}$ is called **the torque** of the force $\boldsymbol{F}$ with respect to the origin of coordinates. Thus, if the torque vanishes at all time, then the angular momentum is conserved, i..e

$$\text{If } \ \boldsymbol{\tau} = 0 \implies \Delta \boldsymbol{L} = 0 \tag{1.150}$$

This happens not only if the force is zero, but also if at all time the force points to the reference point (which in this case the origin of coordinates). For instance, a central force such as the gravitational force that the sun (S) exerts on a planet (P) is

$$\boldsymbol{F}_{S\to P} = -G\frac{M_S m_P}{|\boldsymbol{r}_S - \boldsymbol{r}_P|^3}.\ (\boldsymbol{r}_S - \boldsymbol{r}_P) \propto (\boldsymbol{r}_S - \boldsymbol{r}_P) \implies \boldsymbol{L}_P = \text{const} \tag{1.151}$$

1.10 System of particles

Consider N particles, each has mass m_i and at a position $\boldsymbol{r}_i$ with respect to an inertial frame. We define the position of the center of mass of this system of particles by

$$\boldsymbol{R}_{\rm CM} = \frac{1}{M}\sum_{i=1}^{N} m_i\, \boldsymbol{r}_i \tag{1.152}$$

where $M = \sum_i m_i$ is the total mass of the system. Newton's second law for each particle reads

$$\frac{d\boldsymbol{p}_i}{dt} = \boldsymbol{F}_i^{\rm (ext)} + \sum_{j\neq i} \boldsymbol{F}_{ij} \tag{1.153}$$

Here $\boldsymbol{p}_i$ is the linear momentum of the particle i, $\boldsymbol{F}_{ij}$ is the force acting on i due to j, $\boldsymbol{F}_i^{\rm (ext)}$ represents the resultant of the external forces that i is subject to. By summing over i in the equation above, we obtain

$$\frac{d\boldsymbol{P}}{dt} = \boldsymbol{F}^{\rm ext} + \sum_{i<j}(\boldsymbol{F}_{ij} + \boldsymbol{F}_{ji}), \qquad \boldsymbol{F}^{\rm ext} = \sum_i \boldsymbol{F}_i^{\rm (ext)} \tag{1.154}$$

with $\boldsymbol{P} = \sum_i \boldsymbol{p}_i = M\frac{d\boldsymbol{R}_{\rm CM}}{dt}$ is the total momentum of the N particle. If $\boldsymbol{F}_{ij} = -\boldsymbol{F}_{ji}$, which is the case in most of the physical systems, then we have

$$\frac{d\boldsymbol{P}}{dt} := M\,\frac{d^2\boldsymbol{R}_{\rm CM}}{dt^2} = \boldsymbol{F}^{\rm ext} \tag{1.155}$$

which states that the rate of change of the total linear momentum of the a system of particles (e.g. a body) is determined only by the external forces which act upon it. In particular, **if the total external force vanishes, then the total angular momentum of the system is conserved**, i.e.

$$\text{If} \qquad \boldsymbol{F}^{\rm ext} = 0 \implies \Delta\boldsymbol{P} = 0 \tag{1.156}$$

We define the angular momentum, about the origin of the coordinate system, for a system of N particles is

$$\boldsymbol{L} = \sum_i \boldsymbol{l}_i = \sum_i \boldsymbol{r}_i \times \boldsymbol{p}_i \tag{1.157}$$

Taking the derivative with respect to time of both sides of the above equation gives

$$\begin{aligned}\frac{d\boldsymbol{L}}{dt} &= \sum_i \boldsymbol{r}_i \times \boldsymbol{F}^{\rm ext} + \sum_{i<j}(\boldsymbol{r}_i - \boldsymbol{r}_j)\times \boldsymbol{F}_{ij}\\ &= \boldsymbol{\tau}^{\rm ext} + \sum_{i<j}(\boldsymbol{r}_i - \boldsymbol{r}_j)\times \boldsymbol{F}_{ij}\end{aligned} \tag{1.158}$$

If $\sum_{i<j}(\boldsymbol{r}_i - \boldsymbol{r}_j)\times \boldsymbol{F}_{ij} = 0$[14], then we have

[14]Which is the case, as example for the electrostatic or gravitational forces, where $\boldsymbol{F}_{ij} \propto (\boldsymbol{r}_i - \boldsymbol{r}_j)$.

$$\frac{d\boldsymbol{L}}{dt} = \boldsymbol{\tau}^{\text{(ext)}} \tag{1.159}$$

So, **when the total external torque vanishes, the total angular momentum of the system is conserved**, i.e.

$$\text{If} \quad \boldsymbol{\tau}^{\text{(ext)}} = 0 \implies \Delta \boldsymbol{L} = 0 \tag{1.160}$$

The total kinetic energy of the N particles is just the sum of the kinetic energy of each particle, i.e.

$$T = \frac{1}{2} \sum_i m_i \boldsymbol{r}_i{}^2 \tag{1.161}$$

Now let us write the position vector of a particle i, $\boldsymbol{r}_i$, as

$$\boldsymbol{r}_i = \boldsymbol{R}_{\text{CM}} + \tilde{\boldsymbol{r}}_i \tag{1.162}$$

that is $\tilde{\boldsymbol{r}}_i$ represents the position vector of the particle i with respect to the center of mass of the system. Then, the expression above of the kinetic energy of the system can be re-written as

$$T = \frac{1}{2} \left(\sum_i m_i \right) \left(\frac{d\boldsymbol{R}_{\text{CM}}}{dt} \right)^2 + \frac{1}{2} \sum_i m_i \left(\frac{d\tilde{\boldsymbol{r}}_i}{dt} \right)^2 + \left(\sum_i m_i \frac{d\tilde{\boldsymbol{r}}_i}{dt} \right) \frac{d\boldsymbol{R}_{CM}}{dt} \tag{1.163}$$

The last term in the equation above vanishes since

$$\begin{aligned} \left(\sum_i m_i \frac{d\tilde{\boldsymbol{r}}_i}{dt} \right) &= \sum_i m_i \frac{d\boldsymbol{r}_i}{dt} - \frac{d\boldsymbol{R}_{CM}}{dt} \sum_i m_i \\ &= M \frac{d\boldsymbol{R}_{CM}}{dt} - M \frac{d\boldsymbol{R}_{CM}}{dt} = 0 \end{aligned} \tag{1.164}$$

Thus, we find

$$T = \tfrac{1}{2} M \left(\tfrac{d\boldsymbol{R}_{\text{CM}}}{dt} \right)^2 + \tfrac{1}{2} \sum_i m_i \left(\tfrac{d\tilde{\boldsymbol{r}}_i}{dt} \right)^2 \tag{1.165}$$

This means that the total kinetic energy of a system of N particles can be split into into kinetic energy of the CM, together with the internal (with respect to the CM) kinetic energy.

Example 1.8: Walking on a Boat

A boat of mass M_B has its own stern just touching a dock. A person of mass m_p standing t the bow a distance L away begins walking towards the stern to get off the boat. Assuming that there is no friction between the water and the boat, we want to know how far is the person from the dock when he reached the stern.

First, we should realize that since there are no external forces applied to the person-boat system, the total linear momentum is conserved, i.e.

$$\boldsymbol{P} = \boldsymbol{p}_p + \boldsymbol{p}_B = \boldsymbol{P}(t=0) = 0 \implies M\frac{d\boldsymbol{R}_{\text{CM}}}{dt} = 0 \tag{1.166}$$

This is equivalent to saying that the CM must remain at the same point all the time. When the person was at the bow, the position of the CM, denoted by X_{CM} (since it's basically a motion in one dimension), with respect to the dock is

$$X_{\text{CM}} = \frac{m_p L + M_B L/2}{m_p + M_B} \tag{1.167}$$

When the person moves to the stern, the position of the CM is given by

$$X_{\text{CM}} = \frac{m_p d + M_B(d + L/2)}{m_p + M_B} \tag{1.168}$$

where d represents the distance the person is from the dock when he reaches the stern. Thus, equating the two expressions in (1.167) and (1.168), yields

$$d = \frac{m_p}{M_B + m_p} \tag{1.169}$$

If the mass of the person is very small compared to the mass of the boat we can approximate the distance d by

$$d \simeq \frac{m_p}{M_B} L << L \tag{1.170}$$

which means that the boat moves a tiny distance.

Example 1.9: Stacking blocks

In this example we would like to determine how far can a stack of N identical blocks, each of length L, be made to hang over the right edge of a table. We label the blocks $B_1, B_2, ...B_N$, starting from top to bottom, and denote by $d^{(k)}$ the maximum hang over of a stack of k blocks. We will consider the edge of the table to be the origin of coordinates, which is also the balance point, and x_i to be the coordinate of the center of the i^{th} blcok.

A configuration of a stack of N blocks is in equilibrium as long as the coordinate of its center of mass (CM), $X_{CM}^{(N)}$, is does not lie to the right of the edge of the table[24], with a maximum hang over when the the (CM) lies above the the balance point[25], i.e.

$$X_{\text{CM}}^{(N)}|_{\text{max-hang-over}} = 0 \tag{1.171}$$

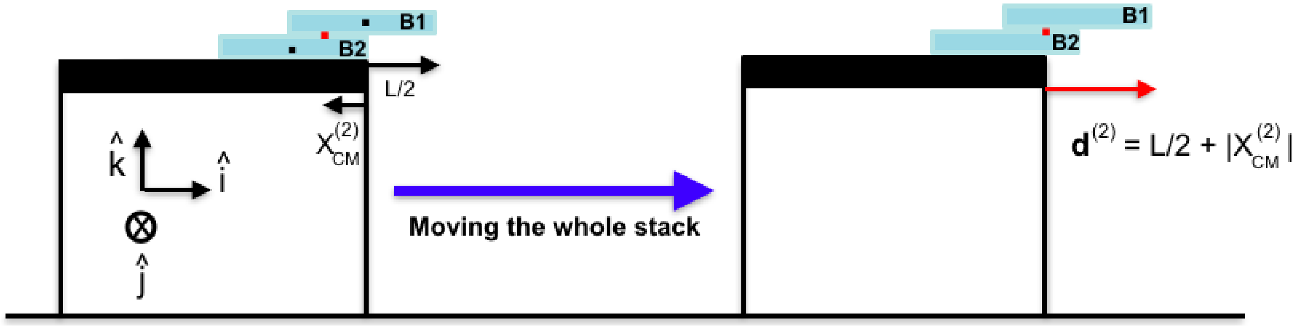


Figure 7: The maximum hang over of a stack of two blocks.

In the case of only one block, a maximum hang over can obtained by placing the center of the block above the edge of the table, and hence $d^{(1)} = L/2$. For $N = 2$, we first put the book B_2 with its right edge at the edge of the table. Then, on the top of it we place the book B_1 with its center on above the edge of the table, as shown in Fig.5. This configuration is in equilibrium, but it does not corresponds to a maximum hangover.

[24]This can be seen as follow: In equilibrium, the following conditions must be satisfied:

$$\text{Net force vanishes : } R\hat{k} - \sum_{i=1}^{N} m_i g \ \hat{k} = 0$$

$$\text{Net torque vanishes : } \sum_{i=1}^{N} m_i g x_i \ \hat{j} + \tau_{\mathbf{R}} = 0$$

where here $\mathbf{R}$ denotes the vector normal force, and $\hat{\imath}$ is a unit vector directed to the right hand side of the table, $\hat{k}$ is upward, opposite the direction of the gravitational acceleration, and $\hat{\jmath} = \hat{k} \times \hat{\imath}$. The second equation is equivalent to the following condition (M being the total mass of the the N blocks.)

$$Mg \ X_{\text{CM}}^{(N)} \ \hat{\jmath} = -|\tau_{\mathbf{R}}| \ \hat{\jmath} \quad \Rightarrow \quad X_{\text{CM}}^{(N)} \leq 0$$

[25]Since in this case the net torque around the balance point will be zero.

This is because the coordinate of the center of mass of the two books is

$$X_{\rm CM}^{(2)} = \frac{m\,(-L/2) + m(0)}{2m} = -\frac{L}{4} < 0 \tag{1.172}$$

Thus, a maximum hang over over can be obtained by moving the whole stack a distance $L/4$ to the right of the balance point. In this case, the maximum hang over is

$$d^{(2)} = \left(\frac{1}{2} + \frac{1}{4}\right) L = \frac{3}{4} L \tag{1.173}$$

Let us continue, and consider stacking three blocks. We first put the book B_3 on the table such that its right edge at the edge of the table. Then, we place on top of B_1 the stack $\{B_1, B_2\}$ on top of B_1 with the CM of $\{B_1, B_2\}$ above the edge of the table. Now, the coordinate of the CM of the stack $\{B_1, B_2, B_3\}$ is

$$X_{\rm CM}^{(3)} = \frac{m\,(-L/2) + 2m(0)}{3m} = -\frac{L}{6} < 0 \tag{1.174}$$

So, we can move the whole stack B_1, B_2, B_3 a maximum distance $L/6$ to the right edge of the table without that the blocks topple over, and we get a maximum hang over

$$d^{(3)} = \left(\frac{1}{2} + \frac{1}{4} + \frac{1}{6}\right) L = \frac{11}{12} L \tag{1.175}$$

We can repeat this for $N = 4$, where we put r the stack B_1, B_2, B_3 such that the CM of the later is above the edge of the table, and put block B_4 under the them such that its right edge is at the edge of the table, and then determine the CM of the whole stack of four blocks. We find that

$$X_{\rm CM}^{(3)} = \frac{m\,(-L/2) + 3m(0)}{4m} = -\frac{L}{8} < 0 \tag{1.176}$$

and so for a stack of 4 blocks the maximum hang over is

$$d^{(4)} = \left(\frac{1}{2} + \frac{1}{4} + \frac{1}{6} + \frac{1}{8}\right) L = \frac{25}{24} L \tag{1.177}$$

So, now the trend is clear! For a stack of N blocks, the maximum hang over is

$$d^{(N)} = \left(\tfrac{1}{2} + \tfrac{1}{4} + \tfrac{1}{6} + ...\tfrac{1}{2N}\right) L = \left(\textstyle\sum_{k=1}^{N} \tfrac{1}{2k}\right) L \tag{1.178}$$

Now, a comment is in order. If one take $N \to \infty$, i.e. infinite number of blocks, we get what is called **harmonic series** which is familiar to any one who has been exposed to Calculus II. The surprising thing is that it **diverges**[26]!, i.e. if you have enough blocks, you can extend the stack beyond the edge of the table as far as you want.

[26]However, it diverges logarithmically, i.e. it approaches infinity much slower. For instance,

2 Gravitating Bodies

2.1 Gravitational Potential Energy of a Spherical Mass Distribution

As we have shown in the Example 1.5, the potential energy of a point-like mass m located a distance r from a point-like test mass m_{test} is

$$V(r) = -\frac{G m_{\text{test}} m}{r} \tag{2.1}$$

If instead of a point like mass m we have an extended object, then we to calculate the potential energy of this system we decompose the object into an infinite number of point-like masses δm_i, and use the fact that the superposition principle to write

$$V(r) = -G m_{\text{test}} \sum_{i=1}^{\infty} \frac{\delta m_i}{r_i} \tag{2.2}$$

where r_i is the distance form δm_i to the location of the test mass, and $\hat{\boldsymbol{r}}_i$. For a continuous distribution of points with mass density ρ, the sum can be replaced by an integral over the whole volume $\mathcal{V}$ of the object, i.e.

$$\boxed{V(P) = -m_{\text{test}}\, G \int_{\mathcal{V}} \frac{\rho(r')}{r'}} \tag{2.3}$$

We note that the gravitational potential energy for this system is m_{test} times a quantity that depends on the density of the mass distribution but not the mass of the test particle. Thus, we define[27]

$$\boxed{U(P) = -G \int_{\mathcal{V}} \frac{\rho(r')}{r'}} \tag{2.5}$$

$\sum_{k=1}^{200} 1/k < 6$, whereas the arithmetic progression $\sum_{k=1}^{200} k = 20,100$. In fact, Leonard Euler, one of the most prolific mathematician, showed that for large N, the harmonic series can be approximated by

$$\sum_{k=1}^{N>>1} \frac{1}{2k} \simeq \ln N + \gamma$$

where $\gamma = 0.57721....$, called the Euler constant.

[27] In a choosing a system of coordinates, the gravitational potential at a point P with vector position $\boldsymbol{r}_P$, reads

$$U(P) = -G \int_{\mathcal{V}} \frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r}' - \boldsymbol{r}_P|}\, d^3\boldsymbol{r}' \tag{2.4}$$

such that which is called **the gravitational potential**, created by the mass distribution at a point P in space regardless if there is a test mass or not.

Next, we consider a spherical shell with a uniform surface mass density σ and radius R, with its center at the origin of coordinates, as shown in Fig.8. Then, if ds is an infinitesimal element of surface of the shell, and r' is the distance from the element ds to the point where the field is to be computed, the potential energy at P

$$U(P) = -G \int \frac{\sigma ds}{r'} \tag{2.6}$$

Since σ is assumed to be constant, all the surface elements at angle θ with respect to the line OP will produce potential at P[28]. Thus, we can decompose the shell into thin

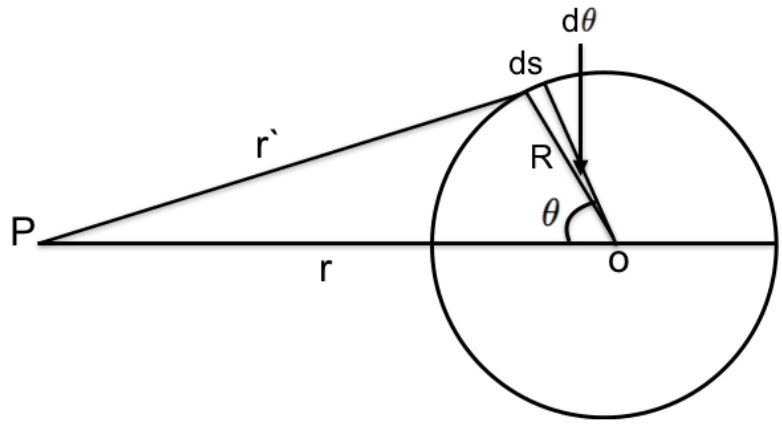


Figure 8: Gravitational potential of a spherical shell at point P.

rings each of thickness $Rd\theta$ and a radius $R\sin\theta$, so that the potential produced by a ring is

$$dU(P) = -G\frac{\sigma\, ds^{(\text{ring})}}{r'} = G\frac{\sigma\,(2\pi R)(R\sin\theta)\sigma d\theta}{r'} \tag{2.7}$$

The potential of the shell can be obtained by integrating over the angular variable from θ, i.e.

$$U(P) = -G(2\pi R^2\sigma)\int_{\theta_1}^{\theta_2} \frac{\sin\theta d\theta}{r'} \tag{2.8}$$

From Fig.8, we have

$$r'^2 = R^2 + r^2 - -2rR\cos\theta \tag{2.9}$$

[28]This also holds if the mass density σ depends just on the angle θ.

Now if differentiate both sides of the equation we get

$$\frac{\sin\theta d\theta}{r'} = \frac{dr'}{rR} \tag{2.10}$$

Thus, the gravitational potential produced at a point P by a surface reads

$$U(P) = G\frac{(2\pi R\sigma)}{r}\int_{r_1'}^{r_2'} dr' = Gm_{\text{test}}\frac{(2\pi R\sigma)}{r}\left(r_2' - r_1'\right) \tag{2.11}$$

where ρ_1 and ρ_2 are the closest and the farthest distance to the point P, respectively. We have two cases:

- **P outside the shell** ($r > R$)
  Then, $r_2' - r_1' = 2R$, and we obtain

  $$\boxed{U^{(\text{shell})}(r > R) = -\frac{GM}{r}} \tag{2.12}$$

  where $M = 4\pi R^2\sigma$ is the mass of the shell. This result shows that the gravitational potential energy at a point outside a shell with uniform mass distribution is the same as if all the mass of this shell is concentrated at its center

- **P inside the shell** ($r < R$)
  Then, $r_2' - r_1' = 2r$, and we obtain

  $$\boxed{U^{(\text{shell})}(r < R) = -\frac{GM}{R} = \text{Constant}} \tag{2.13}$$

The gravitational field produced by the shell at a point P is given by

$$\boldsymbol{g}(\boldsymbol{r}) = -\frac{\boldsymbol{\nabla} V(r)}{m_{\text{test}}} = -\boldsymbol{\nabla} U(r) \tag{2.14}$$

which yields

$$\boxed{\boldsymbol{g}^{(\text{shell})}(r) = \begin{cases} 0, & \text{for } r < R \\ -\frac{GM}{r^2}\,\hat{\boldsymbol{r}}, & \text{for } r > R \end{cases}} \tag{2.15}$$

Here $\hat{\boldsymbol{r}}$ is the unit vector directed from the center of the shell toward the test mass. The above results for the gravitational field at $r < R$ and $r > R$ are known as **Newton's first and second theorem**, respectively.

Now, we turn to the case of a solid sphere with uniform volume mass density of mass M and radius R. It can be considered to be made up of an infinite number of shells, and outside each shell of radius r' and mass $dM_{\text{shell}} = (4\pi r'2dr')\rho$, the gravitational potential is

$$dU^{(\text{solid sphere})}(r) = \begin{cases} -G\frac{dM_{\text{shell}}}{r'}, & \text{for } r < r' \\ -\frac{GdM_{\text{shell}}}{r} & \text{for } r > r' \end{cases} \tag{2.16}$$

So the gravitational potential at an observation point outside the solid sphere is

$$U^{(\text{solid sphere})}(r > R) = -\frac{G}{r}\int_{\text{whole sphere}} dM_{\text{shell}} \tag{2.17}$$

which yields

$$\boxed{U^{(\text{solide sphere})}(r > R) = -\frac{GM}{r}} \tag{2.18}$$

So in this case the gravitational potential is

$$\boxed{\boldsymbol{g}^{(\text{solid sphere})}(r > R) = -\frac{GM}{r^2}\hat{\boldsymbol{r}}} \tag{2.19}$$

Hence, similar to the case of spherical shell with a uniform mass distribution, the gravitation potential field created at a point outside a sphere is the same as if all the sphere?s matter were concentrated into a point at its center.

For a point inside the sphere, there are two contributions to the potential at P

$$U^{(\text{solide sphere})}(r < R) = -\frac{G}{r}\int_0^r \rho 4\pi r'^2 dr' - G\int_r^R \frac{\rho 4\pi r'^2 dr'}{r'} \tag{2.20}$$

$$= -\frac{G}{r}\left(\frac{4\pi}{3}\rho r^3\right) - G2\pi\rho\left(R^2 - r^2\right) \tag{2.21}$$

$$= -\frac{2\pi}{3}G\rho\left(3R^2 - r^2\right) \tag{2.22}$$

or, equivalently,

$$U^{(\text{solide sphere})}(r < R) = -\frac{G(\frac{4\pi}{3}\rho r^3)}{r} \tag{2.23}$$

or, equivalently

$$\boxed{U^{(\text{solide sphere})}(r < R) = -\frac{GM}{2R}\left(3 - \frac{r^2}{R^2}\right)} \tag{2.24}$$

and with the gravitational potential given by

$$\boxed{\boldsymbol{g}^{(\text{solid sphere})}(r < R) = -\frac{GM}{R^2}\left(\frac{r}{R}\right)\hat{\boldsymbol{r}}} \tag{2.25}$$

2.2 Central force problem for two body system

We will study an isolated system of two particles having masses m_1 and m_2, which interact with each other via a time independent central force, i.e.

$$\mathbf{F}_{12}(\boldsymbol{r}) = -\mathbf{F}_{21}(\boldsymbol{r}) = F(r)\frac{(\boldsymbol{r}_1 - \boldsymbol{r}_2)}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|} \tag{2.26}$$

So the equations of motion for the two masses are

$$m_1\ddot{\boldsymbol{r}}_1 = F(r)\frac{(\boldsymbol{r}_1 - \boldsymbol{r}_2)}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|}, \qquad m_2\ddot{\boldsymbol{r}}_1 = -F(r)\frac{(\boldsymbol{r}_1 - \boldsymbol{r}_2)}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|} \tag{2.27}$$

Instead of using $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$, it is more convenient to describe the system in terms of its vector position of its CM and the relative position of the two particles:

$$\boldsymbol{R}_{\mathrm{CM}} = \tfrac{m_1\boldsymbol{r}_1 + m_2\boldsymbol{r}_2}{M}, \qquad \boldsymbol{r} = (\boldsymbol{r}_1 - \boldsymbol{r}_2) \tag{2.28}$$

with $M = (m_1 + m_2)$ is the total mass of the system. Since we are assuming that the system is isolated it implies that the total momentum is conserved, i.e.

$$M\frac{d^2\boldsymbol{R}_{\mathrm{CM}}}{dt^2} = 0 \quad \Longrightarrow \quad \boldsymbol{R}_{\mathrm{CM}} = \boldsymbol{V}^{(0)}\, t \tag{2.29}$$

where

$$\boldsymbol{V}^{(0)} = \frac{m_1\boldsymbol{v}_1^{(0)} + m_2\boldsymbol{v}_2^{(0)}}{M} \tag{2.30}$$

Here $\boldsymbol{v}_1^{(0)}$ and $\boldsymbol{v}_2^{(0)}$ represent the initial velocities of the two masses. Since the velocity of the CM system is constant we can always make a Galilean transformation such that the total momentum vanishes. By choosing the CM of the system to be the origin of the new inertial frame, we have

$$R_{\mathrm{CM}}|_{\text{new frame}} = \mathbf{0} \tag{2.31}$$

In this case the momenta of m_1 and m_2 are given by

$$\boldsymbol{p}_1 = -\boldsymbol{p}_2 = \mu\frac{d\boldsymbol{r}}{dt} \tag{2.32}$$

with $\mu = m_1m_2/M$ is the reduced mass of the system. Hence, the equations of motion in (2.27) reduce to just the equation

$$\boxed{\mu\tfrac{d^2\boldsymbol{r}}{dt^2} = F(r)\ \hat{\boldsymbol{r}} := \mathbf{F}(\boldsymbol{r})} \tag{2.33}$$

Moreover, since $\mathbf{F}(\boldsymbol{r})$ is radial, the potential associated with it[29] is function of the magnitude of $\boldsymbol{r}$ only. Thus, the force can be written as $\mathbf{F}(\boldsymbol{r}) = -\boldsymbol{\nabla}V(r)$, and the above equation of motion can becomes

[29] It is straightforward to show that a central force is a conservative force

$$\mu \ddot{\boldsymbol{r}} = -\boldsymbol{\nabla} V(r) \tag{2.34}$$

In summary, in the inertial frame where the CM of the two masses, the motion of the system is reduced to that of a single particle of mass μ moving in the potential $V(r)$.

In addition to the conservation of the total linear momentum of the system, there are two other conserved quantities: angular momentum and total energy[30].This can be shown as follow:

- **Conservation of angular momentum**

$$\frac{d\boldsymbol{L}}{dt} = \frac{d\boldsymbol{r}}{dt} \times \boldsymbol{p} + \boldsymbol{r} \times \frac{d\boldsymbol{p}}{dt} = F(r)\ \boldsymbol{r} \times \boldsymbol{r} = 0 \tag{2.35}$$

A consequence of this conservation law is that the relative position vector $\boldsymbol{r}$ stays always in the plane perpendicular to the angular momentum vector. Hence, without loss of generality the motion can be described by a just two coordinates, say the polar coordinates (r, θ), in the plane perpendicular to $\boldsymbol{L}$. In these coordinates, the conservation of the angular momentum can expressed as

$$\dot{\theta} = \frac{l}{\mu r^2} \tag{2.36}$$

where l is the magnitude of the angular momentum and it is constant. The above equation has the geometrical interpretation (known as **Kepler's 2nd law**) that:

"**The relative position vector joining the two particles sweeps out an equal areas in equal amounts of time**."

We note that since the righthand side of (2.36) is positive, the angular velocity does not change sign, which means that, depending in the initial conditions, the particle once going around the center in some direction it will not change it.

- **Conservation of total energy**

This is a consequence of the fact that the interaction potential of the system is time-independent (see subsection 1.6). Therefore, we have

$$E := \frac{\mu}{2} v^2 + V(r) = \text{constant} \tag{2.37}$$

Equivalently, it can be expressed in terms of the polar coordinates as

$$\frac{\mu}{2}\left(\dot{r}^2 + r^2\dot{\theta}^2\right) + V(r) = \text{constant} \tag{2.38}$$

Using the conservation law of the angular momentum Eq. (2.36), the above equation reads

[30] A conserved quantity is associated with a symmetry that the Lagrangian of the system exhibits (see section 2.7 of chapter 2 in these notes) .

$$\frac{\mu}{2}\dot{r}^2 + \frac{l^2}{2mr^2} + V(r) = E \tag{2.39}$$

So we obtain a first order differential equation in the distance of the vector between the two particles. Since the second term on the right hand side of Eq. (2.39) is function of r, it is convenient to combine it with $V(r)$, and define an effective potential $V_{\text{eff}}(r)$:

$$V_{\text{eff}}(r) = V(r) + \frac{l^2}{2\mu r^2} \tag{2.40}$$

In this case, the equations describing the system take the simple forms:

$$\begin{cases} \frac{\mu}{2}\dot{r}^2 + V_{\text{eff}}(r) = E, \\ \\ \mu\ddot{r} = -\frac{dV_{\text{eff}}(r)}{dr} \equiv F_{\text{eff}} \end{cases} \tag{2.41}$$

which corresponds to a particle of mass μ in an external effective potential.

Let us discuss some general properties of such system. By solving for $r(t)$ from the equation of energy conservation, then substitute it into Eq. (2.36) we can determine $\theta(t)$ by integration over the time variable, i.e.

$$\theta(t) - \theta_0 = \frac{l}{\mu}\int_0^t \frac{dt'}{r^2(t')} \tag{2.42}$$

Now, instead of integrating over the time variable, we can make use the first Equation in (2.41) to express dt in term of a radial variable r', as

$$dt = \pm\frac{1}{\sqrt{\frac{2}{\mu}\left[E - V_{\text{eff}}(r')\right]}}dr' \tag{2.43}$$

so that the equation of θ will be a function of r,

$$\theta(r) - \theta_0 = \pm\frac{l}{\sqrt{2\mu}}\int_{r_0}^{r} \frac{dr'/r'^2}{\sqrt{[E - V_{\text{eff}}(r')]}}dr' \tag{2.44}$$

from which one can infer $r(\theta)$, i.e. the orbit of the particle. Note that the signs $+$ and $-$ in front of the integral correspond to the part of the trajectory where the radial velocity is positive and negative, respectively.

For a given energy energy E, there could exist values of r at which the radial velocity vanishes[31]. Such points, if they exist, are called **turning points**, and correspond to the solution of the equation

[31] Of course, the particle does not come to rest since $\theta(t)$ is a monotonic function in time.

$$U_{\text{eff}}(r_*) = E \tag{2.45}$$

If r_* is a minimum (maximum), then the point $(r_*, \theta(r_*) \equiv \theta_*)$ is called **perigee** (**apogee**), and in this case we have

$$\theta(r) = \begin{cases} \theta_* \pm \frac{l}{\sqrt{2\mu}} \int_{r_*}^{r} \frac{dr'/r'^2}{\sqrt{[E - V_{\text{eff}}(r')]}} dr' & \longleftarrow \quad \textbf{perigee} \\ \theta_* \pm \frac{l}{\sqrt{2\mu}} \int_{r}^{r_*} \frac{dr'/r'^2}{\sqrt{[E - V_{\text{eff}}(r')]}} dr' & \longleftarrow \quad \textbf{apogee} \end{cases} \tag{2.46}$$

Note that the above expressions have the form $r(\theta) = \mathcal{F}(\theta - \theta_*)$, and as a result $r(\theta)$ will have the property

$$r(\theta_* - \theta) = r(\theta_* + \theta), \quad \forall \theta \tag{2.47}$$

This means that the orbit is invariant under a reflectional symmetry with respect to any axis defined by the origin and a perigee (apogee).

If the system has two turning points with radial distances r_1 and r_2 (say $r_1 < r_2$), then the system is bound and has confined orbit. So, during the time in which the radial variable r passes through one cycle, the angle θ changes by the amount

$$\Delta\theta = \frac{l}{\sqrt{2\mu}} \int_{r_1}^{r_2} \frac{dr'/r'^2}{\sqrt{(E - U_{\text{eff}})}} \tag{2.48}$$

However, for the the trajectory of the particle to define a closed orbit, $\Delta\theta$ must be a rational multiple of 2π, i.e.

$$\textbf{close orbit} \ \Rightarrow \ \Delta\theta = \frac{n}{m} 2\pi, \quad \forall \ n, m \in \mathbb{Z} \tag{2.49}$$

such that when the radial variable makes m cycles, the change in the angle θ will be $2\pi n$, which means the particle will have made n revolutions. Thus, the above requirement put a severe restriction on the type of potentials that the particle is subjected to for its path to be a close orbit. Indeed, almost all potentials do not satisfy the above condition. In the example 1.9 we will see that the orbits of a particle subject to the inverse square force of gravity can be close orbits.

A special case of interest is when the closed orbit is circular. Indeed, for any attractive central force (i.e. $F(r) < 0$), the orbital equation of motion in Eq. (2.41) always admits circular orbits with radius r_0 given by

$$r_0 = \left[-\frac{l^2}{\mu F(r_0)} \right]^{1/3} = \left[\frac{l^2}{\mu \frac{dV(r)}{dr}|_{r_0}} \right]^{1/3} \tag{2.50}$$

Thus for each value of the orbit's radius r_0, there is a unique value of the orbital angular momentum. Which is just the statement that in the case of the circular orbit the angular velocity is such that the centrifugal and the central force balance each other exactly.

2.3 Kepler problem: Gravitational potential

In this subsection we will study the possible orbits of a particle subject to an attractive with a central potential of the for

$$V(r) = -\frac{\lambda}{r} \tag{2.51}$$

with λ determines the strength of the central force, and it is a real positive number[32]. Then, the effective potential reads

$$V_{\text{eff}}(r) = -\frac{\lambda}{r} + \frac{l^2}{2mr^2} \tag{2.52}$$

We see that as $r \to 0$ the effective potential tends to $+\infty$, and approaches 0 from negative values as as r goes to infinity. This implies that, depending on the energy of the particle, the system can be bound (i.e. two turning points) or unbound (i.e. just one turning point). It is straightforward to show that there is a single value r_0 at which the potential $V_{\text{eff}}(r)$ has a global minimum given by

$$[V_{\text{eff}}]_{\text{min}} = -\frac{\mu\lambda^2}{2l^2} \tag{2.53}$$

Hence, when the system has energy E_{min} that is equal to $[V_{\text{eff}}]_{\text{min}}$, then the particle moves in a circular orbit with radius[33]

$$r_0 = \frac{l^2}{\lambda\mu} = -\frac{\lambda}{2E_{\text{min}}} \tag{2.54}$$

which has a period

$$T = \frac{2\pi}{\dot{\theta}(r_0)} = \frac{2\pi l^3}{\mu\lambda^2} \tag{2.55}$$

Now let us consider the more general solution for the particle's orbit. For that we make a change of variable $u = 1/r$ in the second equation in (2.41), and we get the differential equation[34]

---

[32] For the case of gravitational force due to heavy spherical body of mass M, we have $\lambda = GmM$, whereas in the case of Coulomb force generated by an electric charge Q, we have $\lambda = qQ/4\pi\epsilon_0$.

[33] Note that E_c is negative.

[34] Another way is by performing the integral relating $\theta(r)$ to the radial distance, i.e

$$\begin{aligned}\theta(r) &= \pm\frac{l}{\sqrt{2\mu}}\int\frac{dr'/r'^2}{\sqrt{\left[E+\frac{\lambda}{r'}-\frac{l^2}{2\mu r'^2}\right]}}\\ &= \pm\int\frac{d\rho'/\rho^2}{\sqrt{\left[\epsilon-\frac{1}{\rho'^2}+\frac{2}{\rho'}\right]}}\end{aligned}$$

$$\frac{d^2u}{d\theta^2} + u = \frac{m\lambda}{l^2} \tag{2.57}$$

We note that $u = \lambda m l^2$ is a particular solution to the above equation, and without the term right hand side the solution has the form of cosine (or a sine). Hence, we have the general solution

$$u(\theta) = A\cos(\theta - \theta_0) + \frac{m\lambda}{l^2} \tag{2.58}$$

Without loss of generality we can chose the origin of the θ coordinates to be such that the largest value that $u(\theta)$ can take is at $\theta = 0$, which equivalent to setting $\theta_0 = 0$ in the above expression, i.e.

$$u(\theta) = A\cos\theta + \frac{m\lambda}{l^2} \tag{2.59}$$

Now by setting

$$\frac{1}{p} = \frac{m\lambda}{l^2}, \qquad e = Ap \tag{2.60}$$

we obtain the orbital equation[35]

$$r = \frac{p}{e\cos\theta + 1} \tag{2.61}$$

This equation is known as a conic section, which is the curve obtained from the intersection of the surface of a cone with a plane(see Appendix 10.2).

There are three types of conic sections depending on the range/value of the dimensionless parameter e; namely:

---

where we made the change variables $\rho = r'/r_0, \epsilon = E/|E_{\min}|$. The above integral gives

$$\theta(\rho) = \pm \arccos\left[\frac{1/\rho - 1}{\sqrt{1+\epsilon}}\right] + \text{cosntant} \tag{2.56}$$

Without loss of generality we can set the constant on the right hand side of Eq. (2.56) to be equal to zero, since we can always rotate the system of coordinates (or the axis in the plane of motion) by an angle equal to the value of the constant above. .

[35]For a repulsive force (i.e. $\lambda < 0$), such as the case of the electric force between like electric charges, the orbital equation reads

$$r = \frac{p}{e\cos\theta - \mathbf{1}}$$

with $1/p = |\lambda|/ml^2$. So, although the solution for the repulsive force have almost the same form as the attractive force there is a crucial difference in relative sign between the terms in the bracket.

- $\mathbf{0 \leq e < 1}$: the orbit is an **ellipse**. The case $e = 0$ corresponds to a **circle**[36].
- $\mathbf{e = 1}$: the orbit is an **parabola**.
- $\mathbf{e > 1}$: the orbit is an **hyperbola**.

With the help of the equation for the conservation of angular momentum, we can rewrite the energy equation in (2.41) in terms of the radial variable $r(\theta)$ and its derivative with respect to the angle θ as

$$E = \frac{l^2}{2m}\left[\frac{d(1/r)}{d\theta} + (1/r)^2\right] - \lambda\,(1/r) \tag{2.62}$$

After substituting the expression of $r(\theta)$ given in Eq. (2.61) into the above equation we find that, as expected, the energy is constant, and is given by

$$\boxed{E = \frac{l^2}{2mp^2}\left(e^2 - 1\right)} \tag{2.63}$$

Related to the above discussion about the orbits, the following comments are in order:

(a) For $\mathbf{e > 1}$ [Orbit = **Hyperbola**] the particle's **energy is positive**, and so its energy is greater than the escape velocity[37] at every point on its trajectory.

(b) For $\mathbf{0 \leq e < 1}$ [Orbit = **Ellipse/Circle**] the particle's **energy is negative**, and so the particle is bounded by the central force. If $0 > E > V_{\rm eff}^{(\rm min)} = -\lambda^2\mu/2l^2$, the orbit is an ellipse, whereas if $E = V_{\rm eff}^{(\rm min)}$ the particle's trajectory is a circle of radius $r_0 = l^2/\mu\lambda$.

(c) For $\mathbf{e = 1}$ [Orbit = **Parabola**] the particle's **energy is equal to zero**, and so its speed at every point on its trajectory is equal to the escape velocity.

For bound orbits, i.e the particle has negative energy, we can find the period of revolution, T, by integrating the first Kepler's law of equal areas given in Eq. (2.36), and write .

$$\int_{\rm orbit} d\mathcal{A} = \int_0^T \frac{l}{2m}dt \quad \Rightarrow \quad T = \frac{2m\mathcal{A}}{l} \tag{2.64}$$

where $d\mathcal{A} == r^2 d\theta/2$ is the element of area swept by the vector $\mathbf{r}$ in a time interval dt. For an ellipse, $\mathcal{A} = \pi ab = \pi a^2\sqrt{1-e^2}$, with a and b are semi-major axis[38] and

[36] Note for the case of repulsive potential, the parameter can not be in the range $\mathbf{0 \leq e < 1}$ because then the radial distance r will be negative, which is not allowed.

[37] The escape velocity is the minimum velocity that the particle needs to have to break free from the attractive force.

[38] The semi-major axis can be thought of as the average radius of the orbit of the ellipse, i.e. $a = (r_{\rm min} + r_{\rm max}/2$.

semi-minor axis, respectively[39]. Hence,

$$T = \frac{2m\pi a^2\sqrt{1-e^2}}{l} \tag{2.65}$$

Using the definition of p in Eq. (2.60) and the relation between p and a given in the footnote 31, we obtain

$$\boxed{T = 2\pi\sqrt{\frac{m}{\lambda}}\, a^{3/2}} \tag{2.66}$$

We have just proved **Kepler's third law** for inverse-square force, which applies to the motion of the planets around the sun, stating that

"**The square of the period of any planet is proportional to the cube of its semi-major axis of its orbit**."

2.4 Virial Theorem

Consider a system of N point-like particles with mass m_i, position vector $\boldsymbol{r}_i$, velocity $\boldsymbol{v}_i$ and momentum vector $\boldsymbol{p}_i$, and subject to force $\boldsymbol{F}_i$, obeying Newton's equations of motion

$$\dot{\boldsymbol{p}}_i = \boldsymbol{F} \tag{2.67}$$

We define[40]

$$G = \sum_{i=1}^{N} \boldsymbol{p}_i.\boldsymbol{r}_i = \sum_{1}^{N} m_i \dot{\boldsymbol{r}}_i.\boldsymbol{r}_i = \frac{1}{2}\frac{d\mathrm{I}}{dt} \tag{2.68}$$

Here $\mathrm{I} = \sum_{i=1}^{N} m_i r_i^2$ is the moment of inertia of the N particles about the origin of of the coordinate system. Taking the time derivative of the virial gives[41]

$$\frac{1}{2}\frac{dG}{dt} = \frac{1}{2}\frac{d^2\mathrm{I}}{dt^2} + \sum_{i=1}^{N} \dot{\boldsymbol{p}}_i.\boldsymbol{r}_i + \sum_{i=1}^{N} \boldsymbol{p}_i.\dot{\boldsymbol{r}}_i = \Omega + 2T \tag{2.69}$$

[39] The semi-major and semi-minor axis can be be expressed in terms of p and e by rewriting the equation (2.61) in cartesian coordinates as follow:

$$x^2 + y^2 = (p - ex)^2 \quad \Rightarrow \frac{(x-x_0)^2}{a^2} + \frac{y^2}{b^2}$$

which is the equation of ellipse with

$$x_0 = \left(\frac{e}{1-e^2}\right)p, \quad a = \left(\frac{1}{1-e^2}\right)p, \quad b = \left(\frac{1}{\sqrt{1-e^2}}\right)p$$

[40] The quantity G is called the **virial**.

[41] This equation is also known as **Lagrange 'sidentity**.

where T is the total kinetic energy of the system, and $\Omega = \sum_{i=1}^{N} \boldsymbol{F}_i.\dot{\boldsymbol{r}}_i$ is known as the **virial of Clausius**. Now let us compute the time average of the above expression over a long period of time τ, i.e.

$$< \frac{dG}{dt} >=< \Omega > +2 < T > \tag{2.70}$$

with

$$< X >= \lim_{\tau\to\infty} \frac{1}{\tau} \int_0^{\tau} X(t)dt \tag{2.71}$$

The left hand side of (2.70) yields

$$< \frac{dG}{dt} >= \lim_{\tau\to\infty} \frac{1}{\tau} \big[G(\tau) - G(0)\big] \tag{2.72}$$

If G is a bounded function, then its values will be oscillate between a minimum $[G_{\min}$ an a maximum $G_{\max}]$, and so $< \frac{dG}{dt} >$ vanishes as τ tends to infinity. This statement defines a virial equilibrium of a state of a system. In this case, the system obeys the **virial theorem**:

$$\boxed{< T >= -\tfrac{1}{2} < \Omega >} \tag{2.73}$$

For forces that are derivable from a potential V, the viral theorem reads[42]

$$\boxed{< T >= -\tfrac{1}{2} < \textstyle\sum_{i=1}^{N} \boldsymbol{r}_i.\boldsymbol{\nabla}_i V >} \tag{2.74}$$

It is important to point out that the Virial theorem applies only for bound system[43]. Hence it is useful to define define the virial parameter

$$\mathcal{Q}_{\text{virial}} = -\frac{< T >}{< \Omega >} \tag{2.75}$$

which measures the relative importance of the kinetic and gravitational energy of a system of particles. For bound system that was left to itself for long time it will eventually reach virial equilibrium and will have $\mathcal{Q}_{\text{virial}} = 1/2$, whereas if it is unbound the kinetic energy overcomes the potential energy and in this case the virlal parameter will be larger than unity. Moreover, a gravitationally bound system of very large number

[42]The expression $\boldsymbol{\nabla}_i V.\boldsymbol{r}_i$ is just symbolic, for ease of notation. It should be understood as

$$\frac{\partial V}{\partial \boldsymbol{r}_i}.\boldsymbol{r}_i = \frac{\partial V}{\partial x_i}.x_i + \frac{\partial V}{\partial y_i}.y_i + \frac{\partial V}{\partial z_i}.z_i$$

[43]For example, it does not apply to a the particles of an explosion.

of particles such as cluster of galaxies or galaxies tends very rapidly into viria equilibrium[44] and hence for such systems we the time averages of the kinetic and potential energy are close to their current values.

Now let us the case of central forces, with a potential of the form

$$V = \frac{1}{2}\sum_{j\neq i}\lambda_{ij}(r_{ij})^n, \qquad \lambda_{ij} = \lambda_{ji} \tag{2.76}$$

where $r_{ij} = |\boldsymbol{r}_i - \boldsymbol{r}_j|$ is the distance between the i^{th} and i^{th} particles. Then, the force acting the k^{th} particle is given by $\boldsymbol{F}_k = -\frac{1}{2}\boldsymbol{\nabla}_k\sum_{j\neq i}\lambda_{ij}(r_{ij})^n$, which can be equivalently written as

$$\boldsymbol{F}_k = -\frac{1}{2}\boldsymbol{\nabla}_k\sum_{j\neq k}\lambda_{kj}(r_{kj})^n - \frac{1}{2}\boldsymbol{\nabla}_k\sum_{j\neq k}\lambda_{jk}(r_{jk})^n$$

By using

$$\boldsymbol{\nabla}_k r_{kj} = \boldsymbol{\nabla}_k r_{jk} = \boldsymbol{\nabla}_k\big[(\boldsymbol{r}_k - \boldsymbol{r}_j).(\boldsymbol{r}_k - \boldsymbol{r}_j)\big]^{1/2} = \frac{1}{r_{kj}}(\boldsymbol{r}_k - \boldsymbol{r}_j)$$

Thus, we obtain

$$\boldsymbol{F}_k = -\sum_{j\neq k}\lambda_{kj}n(r_{kj})^{n-2}(\boldsymbol{r}_k - \boldsymbol{r}_j)$$

Then, the virial Clausius is

$$\sum_k \boldsymbol{F}_k.\boldsymbol{r}_k = -\sum_{j,k;j\neq k}\lambda_{kj}n(r_{kj})^{n-2}(\boldsymbol{r}_k - \boldsymbol{r}_j).\boldsymbol{r}_k \tag{2.77}$$

which can be written in a symmetric form as

$$\begin{aligned}\sum_k \boldsymbol{F}_k.\boldsymbol{r}_k &= -\frac{1}{2}\sum_{j,k;j\neq k}\lambda_{kj}n(r_{kj})^{n-2}(\boldsymbol{r}_k - \boldsymbol{r}_j).\boldsymbol{r}_k - \frac{1}{2}\sum_{j,k;j\neq k}\lambda_{kj}n(r_{kj})^{n-2}(\boldsymbol{r}_j - \boldsymbol{r}_k)\boldsymbol{r}_j \\ &= -nV\end{aligned} \tag{2.78}$$

where in obtaining the second term in the right hand side of the above equation we first made a change of labeling of $(j,k) \to (k,j)$, then used the fact that $r_{jk} = r_{kj}$ and $\lambda_{jk} = \lambda_{kj}$. Hence, for a bound system subject to central potential that has the form of a power law in the distances between the particles, we have

$$\boxed{< T >= \tfrac{n}{2} < V >} \tag{2.79}$$

[44] In fact, any bound system that is out of virial equilibrium tends to evolve toward it, while conserving its total energy.

Example 2.1: Some Applications of the Virial Theorem

Below we will look into few example where can virial theorem can be used to study the properties of bound systems:

- **Collapsing Cloud of Gas**

  As an application of the virial theorem to Astrophysical systems, we consider a spherically symmetric cloud of gas of radius R and mass M, collapsing under its own gravity. First we will compute the gravitational potential energy of the gas cloud for an infinitesimal shell of thickness dr at a distance r from the center of the spherical cloud, which is given by

$$dV = -G\frac{M(r)4\pi r^2\rho(r)dr}{r} \tag{2.80}$$

  where $\rho(r)$ is the mass density of the gas in the shell and $M(r)$ is the mass of the gas inside the shell. Thus, the gravitation potential energy is

$$V = -4\pi G\int_0^R \frac{M(r)4\pi r^2\rho(r)dr}{r} \tag{2.81}$$

  Assuming constant mass density, we can integrate the above expression, and get

$$V = -\frac{3GM^2}{5R} \tag{2.82}$$

  Now with the use of the Virial theorem, the total energy of the cloud of gas is

$$E = \frac{1}{2}V = -\frac{3GM^2}{10R} \tag{2.83}$$

  Note that this energy depends on the size of the cloud. So if the collapse started with an initial radius R_i to a final radius R_f, where does the energy difference go? the answer is that this energy difference gets radiated away by the gas, i.e.

$$\Delta E_{\text{radiation}} = \frac{3GM^2}{10}[\frac{1}{R_f} - \frac{1}{R_i}] \tag{2.84}$$

  For instance, taking $R_i \to \infty$ (i.e. the collapse started at infinity), the amount of radiation emitted by the cloud when it reaches a radius R_f is[45]

$$\Delta E_{\text{radiation}} = \frac{3GM^2}{10R_f} = -\frac{1}{2}V(R_f) \tag{2.85}$$

[45] It turns out when applying this calculation for stars, the energy radiated due the gravitational collapse yields an age for the star much smaller than 100 million years, which is clearly in conflict with the observations. It was only after the discovery of quantum mechanics that it was understood that the main source of energy radiated is due to thermonuclear reactions inside of the star. For more details, see reference [26]

- **The Law of Ideal Gas**

In this example we will apply the virial theorem to the kinetic theory of gases and derive ideal gas law. We will consider N in a container of volume V at a temperature T. Then, according to equipartition theorem the average kinetic energy per molecule is $3/2k_{\mathrm{B}T}$, and so for a system of N particles we have

$$\mathrm{KE} = \frac{3}{2} N k_B T \tag{2.86}$$

where we denoted the kinetic energy KE to avoid confusion with the temperature. To compute the virial Clausius, we note that the collision between the molecules is negligible since we are assuming an ideal gas, and the only collision that the molecules have is the wall of the container. Thus, for an infinitesimal area $d\boldsymbol{A} = dA\hat{\boldsymbol{n}}$, with

$$d\boldsymbol{F}_i = -PdA\hat{\boldsymbol{n}} \tag{2.87}$$

Here P is the gas pressure. Since the number of particles in a gas is huge (of order $\mathcal{N}_{\mathrm{Avogadro}} \sim 6 \times 10^{23}$), the summation over all particles colliding with the walls can be replaced by an integration over the surface of the container, i.e.

$$\frac{1}{2}\sum_{i=1}^{N} \boldsymbol{F}_i.\boldsymbol{r}_i = -\frac{P}{2}\oint \hat{\boldsymbol{n}}.\boldsymbol{r}\, dA = -\frac{P}{2}\int_{\mathrm{V}} (\boldsymbol{\nabla}.\boldsymbol{r})\, dV = -\frac{3}{2}PV \tag{2.88}$$

where we used the fact that $\boldsymbol{\nabla}.\boldsymbol{r} = 3$. Thus, with the results in Equations (2.86) and (2.87), the expression of the virial theorem yields

$$PV = Nk_BT \tag{2.89}$$

which is the familiar formula of an ideal gas law.

- **Inferring the Existence of Dark Matter**

Another use of virial theorem is to extract the mass of stationary stable systems, such as galaxy cluster bound by the gravitational potential. For a set of N galaxies of similar mass m, we have

$$T = \frac{1}{2}\sum_i m_i v_i^2 = \frac{1}{2}M_{\mathrm{cluster}} < v^2 > \tag{2.90}$$

where $M = Nm$ is the total gravitational mass of the cluster and

$$< v^2 >= \frac{\sum_i v_i^2}{N} \tag{2.91}$$

represents the mean square speed of the galaxies in the cluster. The gravitational potential energy of the cluster is

$$V = -G\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\frac{m_i m_j}{r_{ij}} = -GM^2\frac{1}{N^2}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\frac{1}{r_{ij}} \tag{2.92}$$

We can define[46]

$$< \frac{1}{r} >= \frac{1}{N(N-1)/2}\sum_{i}\sum_{j>i}\frac{1}{r_{ij}} \tag{2.93}$$

to be the average of the inverse distance between two galaxies in the cluster. Then,

$$V = -GM^2\frac{N(N-1)/2}{N^2} < \frac{1}{r} >$$

If we define $r_g := 2/< \frac{1}{r} >$, called the gravitational radius, and take N to be very large, then the potential can be written in the simple form

$$V = -\frac{GM^2_{\text{cluster}}}{r_g},$$

We can now apply the virial theorem in Eq.(2.79) with $n = -1$, and get

$$M = \frac{< v^2 > r_g}{G} \tag{2.94}$$

Hence, one can infer the mass of the galaxy cluster from its velocity dispersion and its size[47]. Indeed, this method has been used and still used for both cluster of galaxies and galaxies[48]. For instance, in 1933 Fritz Zwicky, an Astrophysicist at Caltech, estimated the mass of the Coma cluster[49], by measuring the total amount of light emitted by the stars in the galaxies. To his surprise, he found that the observed value of the mass pf the cluster was much smaller than the

[46]Since there are $N(N-1)/2$ pairs of galaxies in a system of N galaxies in the cluster.

[47]To be more precise, what astrophysicists measure is instead of $< v^2 >$, the line of sight velocity dispersion $< v_\parallel^2 >$, and if we assume isotropy $< v_\parallel^2 >= \frac{1}{3} < v^2 >$. For the gravitational radius r_g one instead introduces an "effective size of the system", R_{eff}, as a rough estimate of the size of the system. In addition. an over all fudge parameter, α, of order unity, which accounts for the radial distribution of the galaxies. So, the expression that astrophysicist usually use for the virial mass is

$$M = \alpha\frac{< v_\parallel^2 > R_{\text{eff}}}{G}$$

[48]In 1921, Sir Eddington, used the virial theorem to estimate the mass of the Globular clusters

[49]Coma cluster is at mean distance of about 336 million light year (or equivalently about 100 Mpc) from from Earth, and consists of approximately 1000 galaxies

gravitational mass computed using virial theorem. The missing mass is ascribed to dark matter, which it turns out to be about 90% of the visible matter in the cluster[50].

2.5 Restricted three body problem

[50] Also, measurements of rotation curves of spiral galaxies also indicate that visible matter account for about 10% of the matter in the galaxies.

3 D'Alembert Principle and Euler-Lagrange Equations

In this chapter, I will discuss the Lagrangian formulation of mechanics which will enable us to derive the equations of motion in a much more easy and elegant way than in the Newtonian mechanics that we encountered in the previous chapter. As we will show, the equations of motion can be obtained from the two basic forms of energies contained in the system: the kinetic energy and the potential energy, and without solving explicitly for the constraint forces acting on the system. Furthermore, the equations of motion take the same form in all coordinate systems.

3.1 Constraints and generalized coordinates

Almost every physical system is subject to constraints which comes in two classes: holonomic and non-holonomic constraints, as will be discussed below.

3.1.1 Holonomic constraints

For a system of N particles in a 3-dimensional space, the holonomic constraints are defined by relations that involve the coordinates and time only:

$$\Phi_\alpha(x_1, x_2, ..., x_{3N}; t) = 0; \quad \alpha = 1, ..., m \tag{3.1}$$

A system whose all constraints equations are of the holonomic form is called **holonomic system**. If the time does not appear explicitly in the above equation, the constraints are known as **scleoronomic** constraints, otherwise they are called **rheonomic** constraints.

The key point of holonomic constraints is that the coordinates are no longer independent; they are related via the constraint equations. So, in this case we can describe the configuration of a system of N particles with $n = (3N - M)$ independent coordinates, denoted by $\{q_1, q_2, ..., q_n\}$, called the **generalized coordinates**. It is important to realize that **n is independent of the particular set of coordinates used to describe the system**.

The process of obtaining the cartesian coordinates in terms of the generalized coordinates is called coordinates transformation:

$$x_i = x_i(q_1, q_2, ..., q_n); \qquad i = 1, 2, ..., 3N \tag{3.2}$$

It is desirable that one and only one set of coordinates of q's corresponds to each possible configuration of the system. That is, there should be one to one correspondence between points in the allowable domain of the x's and points in the allowable domain of q's for each value of time. The necessary condition for this to be the case is that the Jacobian determinant of the coordinates transformation (3.2) must be non zero, i.e. the transformation is invertible. For instance, suppose that the 3 N coordinates x_i are subject to m holonomic constraint equations of the form:

$$f_\alpha(x_1, x_2, ...x_{3N}; t) = C_\alpha; \qquad \alpha = 3N - m + 1, ..., 3N \tag{3.3}$$

where the C_α's are constants. So the system can be described with $n = (3N - m)$ independent generalized coordinates $\{q_a;\ a = 1, ..., n\}$. Now, we define an additional set of m q's and identify them with m constant functions f_α of Eq (3.3), i.e.

$$q_\alpha = f_\alpha(x_1, x_2, ..., x_{3N}; t); \qquad \alpha = n + 1, ..., 3N \tag{3.4}$$

Thus, the transformations (3.2) can be now considered to of the form

$$x_i = x_i(q_1, q_2, ..., q_{3N}; t); \qquad i = 1, 2, ..., 3N \tag{3.5}$$

if the Jacobian determinant is non-zero, i.e.

$$\mathcal{J} := \left| \frac{\partial\,(x_1, x_2, ..., x_{3N})}{\partial\,(q_1, q_2, ..., q_{3N})} \right| \neq 0 \tag{3.6}$$

In this case one can solve for the q's as functions of the x's and time, as

$$q_a = g_a(x_1, x_2, ..., x_{3N}; t); \qquad a = 1, ..., n \tag{3.7}$$

where the remaining constants $q_\alpha, \alpha = n + 1, ..., 3N$ are given by equation (3.4).

Example 3.1: Particle on a circular path

Consider a particle which is constrained to move on a circular path of radius R, as show in the figure below. The equation of the constraint is

$$x_1^2 + x_2^2 = R^2 \tag{3.8}$$

So this system has only one degree of freedom. We choose the generalized coordinate describing the motion of the particle to be the polar angle, i.e.

$$q_1 \equiv \theta \tag{3.9}$$

which can vary freely without violating the constraint. Now we define a second generalized coordinate q_2 and which will be equal to the radius:

$$q_2 = R \tag{3.10}$$

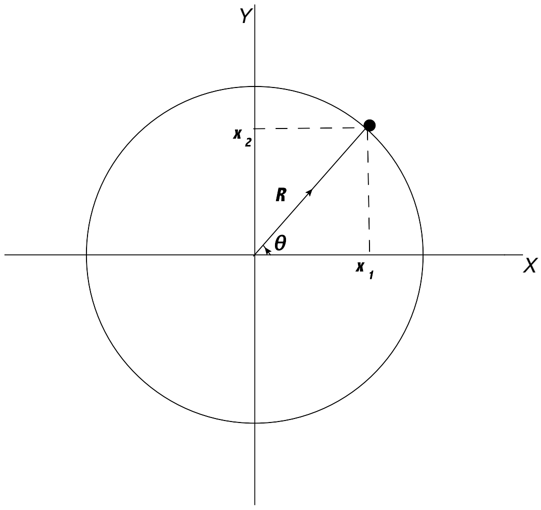


Figure 9: Paricle on a circular path.

Then, the transformation equations read

$$x_1 = q_2 \cos q_1; \quad x_2 = q_2 \sin q_1 \tag{3.11}$$

and has the Jacobian given by

$$\mathcal{J} = \left| \frac{\partial (x_1, x_2)}{\partial (q_1, q_2)} \right| = \begin{vmatrix} \frac{\partial x_1}{\partial q_1} & \frac{\partial x_1}{\partial q_2} \\ \\ \frac{\partial x_2}{\partial q_1} & \frac{\partial x_2}{\partial q_2} \end{vmatrix} \implies \mathcal{J} = -q_2 \tag{3.12}$$

Hence the q's may be expressed as functions of the x's except when the jacobian is zero, i.e. $q_2 = 0.$, and in this case the radius of the circle is zero and the angle q_1 is undefined. The transformation of q's in terms of x's is given by

$$q_1 = \tan^{-1} \frac{x_2}{x_1}; \qquad q_2 = (x_1^2 + x_2^2)^{1/2} \tag{3.13}$$

where $0 \leq q_1 < 2\pi$ and $q_2 > 0$ so that the q's be single-valued functions of the x's.

Example 3.2: Two particle at constant distance

Consider two particles connected by a rigid rod of length L in the plane xy, with its ends at coordinates $(x_1, y_1), (x_2, y_2)$, respectively. The constraint equation is

$$(x_2 - x_1)^2 + (y_2 - y_1)^2 - L^2 = 0 \tag{3.14}$$

Since L is constant this is a holonomic system with scleoronomic constraint. It has three degrees of freedom (4 cartesian coordinates with one constraint), and so can be described by 3 generalized coordinates, which we choose to be

$$\begin{cases} (x_{CM}, y_{CM}) : & \text{the coordinate of the center of mass of the system} \\ \theta \quad : & \text{the angle between the rod and the x-axis} \end{cases} \tag{3.15}$$

3.1.2 Non-holonomic constraints

Theis type of constraints can not be put in the form of equation (3.1), and there are two kinds of them:

- **Inequality constraints**[10]
  In this case the coordinates are restricted by inequalities. Examples:
  (i) A particle on a sphere which is allowed to roll off: $\sum_{i=1}^{3} x_i^2 - R^2 \geq 0$;
  (ii) A particle moving in a spherical container of radius R: $\sum_{i=1}^{3} x_i^2 \leq R^2$.

- **Non-integrable (or history dependent) constraints**
  This kind of constraints are written as differential expressions of the form[11]

$$\sum_{i=1}^{n} \boldsymbol{a}_\alpha . d\boldsymbol{q} + a_{\alpha t}\, dt = 0; \qquad \alpha = 1, 2, ..., m \tag{3.16}$$

  where $\boldsymbol{q} \equiv (q_1, q_2, ..., q_n)$, denotes a point in the the space of generalized coordinates, and $\boldsymbol{a}_{\alpha i} \equiv (a_{\alpha 1}, a_{\alpha 2}, ..., a_{\alpha n})$, and $a_{\alpha t}$ are certain functions of the n $q_i's$ and the time variable. Such constraints have to be non-integrable, otherwise,upon integration they could reduce to holonomic constraints.

  The integrability of (3.16) means that the left hand side of the equation can be written as an exact differential:

[10] Also called unilateral constraints.

[11] In these lectures we will restrict our self only to the non-holonomic constraints that are linear in velocities, called linear non-holonomic constants.

$$\sum_{i=1}^{n} \boldsymbol{a}_{\alpha} . d\boldsymbol{q} + a_{\alpha t}\ dt = d\mathcal{F}(\mathbf{q}; t) \tag{3.17}$$

A necessary and sufficient conditions for this to occur are[12]

$$\begin{aligned} \frac{\partial a_{\alpha i}}{\partial q_j} - \frac{\partial a_{\alpha j}}{\partial q_i} &= 0 \\ \frac{\partial a_{\alpha i}}{\partial t} - \frac{\partial a_{\alpha t}}{\partial q_i} &= 0 \end{aligned} \tag{3.19}$$

In general, the approach to showing whether or not a differential constraint is holonomic is based on finding a function (called integration factor) which will turn the differential constraints equation into an exact differential. In other words, a differential constraints of the form

$$\sum_{i=1}^{n} a_i(\boldsymbol{q}; t)\ dq_i + a_t(\boldsymbol{q}; t)\ dt = 0 \tag{3.20}$$

will is holonomic if there exists a function $I(q_1, q_2, ...q_n, t)$, such that

$$I(\boldsymbol{q}, t) \left[\sum_{i=1}^{n} a_i(\boldsymbol{q}t)\ dq_i + a_t(\boldsymbol{q}; t)\ dt \right] = d\mathcal{F}(\boldsymbol{q}; t) \tag{3.21}$$

where

$$\begin{aligned} \frac{\partial \mathcal{F}}{\partial q_i} &= I(\boldsymbol{q}, t)\ a_i(\boldsymbol{q}, t) \\ \frac{\partial \mathcal{F}}{\partial t} &= I(\boldsymbol{q}, t)\ a_t(\boldsymbol{q}, t) \end{aligned} \tag{3.22}$$

---

[12]**Proof:**
According to equation (3.17), we hav

$$\sum_{i=1}^{n} \frac{\partial f}{\partial q_i}\ dq_i + \frac{\partial f}{\partial t}\ dt = \sum_{i=1}^{n} a_{\alpha i}\ dq_i + a_{\alpha t}\ dt \tag{3.18}$$

Equating the coefficients in front of dq_i and dt on both sides of the equation yields

$$\begin{aligned} a_{\alpha i} &= \frac{\partial f}{\partial q_i} \implies \frac{\partial a_{\alpha i}}{\partial q_j} - \frac{\partial a_{\alpha j}}{\partial q_i} = 0 \\ a_{\alpha t} &= \frac{\partial f}{\partial t} \implies \frac{\partial a_{\alpha i}}{\partial t} - \frac{\partial a_{\alpha t}}{\partial q_i} = 0 \end{aligned}$$

So a consequence of their non-integrability, the **non-holonomic constraints can't be used to reduce the the number of generalized coordinates**. Hence, generalized coordinates are useful only for holonomic systems.

Example 3.3: Cylinder Rolling Without Slipping

As an example of an integrable differential constraint, consider a cylinder rolling without slipping on a table. In this case the constraint reads

$$\frac{dx(t)}{dt} = R\,\frac{d\theta(t)}{dt} \implies dx = R\;d\theta \tag{3.23}$$

which after integrating both sides of the above equation we get

$$x - R\;\theta = \text{constant} \implies \text{holonomic constraint} \tag{3.24}$$

Example 3.4: Vertical Disk Rolling Without Slipping

The standard example of a non-holonomic constraint is a vertical disk of radius R which rolls without slipping on a fixed horizontal plane xy (see the figure). We can choose as generalized coordinates, the center of the disk, (x, y), on the plane (which is the same coordinates of the contact point), together with angle of rotation ϕ of the disk about a perpendicular axis through its center, and the angle θ between the axis of the symmetry of disk and the x-axis. As the disk rolls through an angle $d\phi$, the point of contact moves a distance $Rd\phi$ in a direction depending on θ:

$$\begin{aligned} dx - R\sin\theta d\phi &= 0 \\ dy - R\cos\theta d\phi &= 0 \end{aligned} \tag{3.25}$$

or, equivalently,

$$\begin{aligned} R\dot{\phi}\sin\theta - \dot{x} &= 0 \\ R\dot{\phi}\cos\theta - \dot{y} &= 0 \end{aligned} \tag{3.26}$$

So this system can be described by four generalized coordinates, (x, y, θ, ϕ), and two independent differential constraints. Unlike the simple 1-dimensional case (example 1. 5) where the disk is confined to the yz plane, rolling along x-axis subject to holonomic constraint, in this example the differential constraints are non-integrable. It is straightforward to check that the coefficients $a_{1x} = a_{2y} = 1, a_{1\theta} = a_{2\theta} = 0, a_{1\phi} = -R\sin\theta$, and $a_{2\phi} = -R\cos\theta$ do not satisfy Eq (3.19), because the coefficients in front of $d\phi$ depend on θ, where as the coefficients in front of $d\theta$ are zero[12]. Therefore, all 4 generalized coordinates have to be given to describe the configuration of the system.

3.2 D'Alembert principle of virtual work

Consider a system of N particles subject to K holonomic constraints:

$$\Phi_\alpha(\boldsymbol{r}_1, \boldsymbol{r}_2, ..., \boldsymbol{r}_N; t) = 0; \qquad \alpha = 1, ..., K \tag{3.27}$$

Each constraint of this form imposes a restriction on the possible velocities:

$$\sum_{i=1}^{N} \left(\frac{\partial \Phi_\alpha}{\partial \boldsymbol{r}_i} \right) \boldsymbol{v}_i + \frac{\partial \Phi_\alpha}{\partial t} = 0; \qquad \alpha = 1, ..., K \tag{3.28}$$

So there are many (in fact infinitely many) allowed velocities since we have imposed K number of constraints on the $3N$ components of the velocities . An infinitesimal displacement over time dt due to allowed velocities will be called **the allowed infinitesimal displacement** and it is given by

$$d\boldsymbol{r}_i = \boldsymbol{v}_i dt \tag{3.29}$$

These allowed displacements together with dt satisfy the constraint equations:

$$\sum_{i=1}^{N} \left(\frac{\partial \Phi_\alpha}{\partial \boldsymbol{r}_i} \right) d\boldsymbol{r}_i + \frac{\partial \Phi_\alpha}{\partial t} = 0; \qquad \alpha = 1, ..., K \tag{3.30}$$

As there are many allowed velocities there are many allowed infinitesimal displacements. Now, let us consider a small change $\delta \boldsymbol{r}$ in the configuration of the system defined by[15]

$$\delta \boldsymbol{r} = d\boldsymbol{r}_i - d\boldsymbol{r}'_i \tag{3.31}$$

which is known as the **virtual displacement**. Note that unlike the definition of virtual displacement given in text books where $\delta \boldsymbol{r}$ takes place without any passage of

[12]Another way to see that theses constraints can not be integrable is that if they were then they would give two holonomic constraints and that would be used to eliminate two of the variables, say θ and $\phi$and hence the configuration of the system would be uniquely determined by (x, y) alone. However, this is clearly can not be the case since θ can be anything for a given (x, y) consistent with the constraints. Similarly, one can any ϕ for a given values of x, y, and θ) by rolling the cylinder around circles of different diameter.

[15]This definition of the virtual displacement was introduced by S. Ray, J. Shamanna in 2003 (for more details see [7]).

time, the above definition is rather the difference of any two **allowed displacements** during a time dt, i.e. $\delta \boldsymbol{r} = (\boldsymbol{v}_i - \boldsymbol{v}'_i)dt$. The difference of the two allowed velocities $\bar{\boldsymbol{v}}_i = (\boldsymbol{v}_i - \boldsymbol{v}'_i) = \frac{\delta r}{dt}$ is called the **virtual velocity**. Since both $d\boldsymbol{r}_i$ and $d\boldsymbol{r}'_i$ satisfy the constraint equation (3.30), then the virtual displacement satisfy

$$\boxed{\sum_{i=1}^{N} \left(\frac{\partial \Phi_\alpha}{\partial \boldsymbol{r}_i}\right) \delta \boldsymbol{r}_i = 0; \qquad \alpha = 1, ..., K} \tag{3.32}$$

Thus, the absence of $\frac{\partial \Phi_\alpha}{\partial t}$ in Eq (3.32) means that virtual displacements are the allowed displacements in the case of frozen constraints, in the sense that we make $\frac{\partial \Phi_\alpha}{\partial t} \to 0$, though $\frac{\partial \Phi_\alpha}{\partial \boldsymbol{r}_i}$ still involves time. In the case of scleronomic hololomic constraints, the virtual displacements and the allowed displacements are collinear as $\frac{\partial \Phi_\alpha}{\partial t} = 0$.

In a similar way, for a system that has $\tilde{k}$ non-holonomic constraints of the form

$$\sum_{i=1}^{3N} a_{\alpha i}\, dx_i + a_{\alpha t} dt = 0; \qquad \alpha = 1, ..., \tilde{k} \tag{3.33}$$

the virtual displacements $\delta \boldsymbol{r}_i$ satisfy

$$\boxed{\sum_{i=1}^{3N} a_{\alpha i}\, \delta x_i = 0} \tag{3.34}$$

The net force $\boldsymbol{F}_i$, i=1, 2, ..., N, that each particle is acted upon, can be written as

$$\boldsymbol{F}_i = \boldsymbol{F}_i^{\,(\mathrm{app})} + \boldsymbol{F}_i^{\,(\mathrm{c})} \tag{3.35}$$

where $\boldsymbol{F}_i^{\,\mathrm{app}}$ and $\boldsymbol{F}_i^{\,\mathrm{c}}$ are the applied and the constraint force on the ith particle. The virtual work of these forces in virtual displacements $\delta \boldsymbol{r}_i$ reads

$$\delta W = \sum_{i=1}^{N} \boldsymbol{F}_i .\ \delta \boldsymbol{r}_i \tag{3.36}$$

We will restrict our self to a common class of constraints, known as the **ideal constraints**, for which the virtual work done by the constraint forces is zero. In this case, the virtual work of the system reads

$$\delta W = \sum_{i=1}^{N} \boldsymbol{F}_i^{\,(\mathrm{app})} .\ \delta \boldsymbol{r}_i \tag{3.37}$$

Now suppose that each particle of the system is in equilibrium; i.e. $\boldsymbol{F}_i = 0, i = 1, 2, , ..., N$, then according to (3.36) we have

$$\sum_{i=1}^{N} \boldsymbol{F}_i^{\text{(app)}} . \, \delta \boldsymbol{r}_i = 0 \tag{3.38}$$

This is known as **the principle of virtual work** which applies only to **static situations**.

Comments:

- The applied forces $\boldsymbol{F}_i^{\text{(app)}}$. will in general be non zero and the virtual displacements will be related to one another via the constraint equations.
- For a system of particles, it is **the net virtual work** done by the constraint forces that vanish, i.e.

$$\sum_{i=1}^{N} \boldsymbol{F}_i^{\text{(c)}} . \, \delta \boldsymbol{r}_i = 0 \tag{3.39}$$

and not necessarily the virtual works of the constraint on each particle.

Example 3.5: Virtual work of the constraint forces in Atwood machine

The Atwood machine consists of two masses (considered to be point-like) connected by a string passing over a pulley. If the string is massless and inextensible, and the pulley is assumed to be massless and frictionless, then the forces of constraint will reduce to the force of tension in the string.

The virtual displacements compatible with the constraint will be in the vertical direction (i.e. along the y-axis) and so the forces of constraints are the tensions $\boldsymbol{T}_1$ and $\boldsymbol{T}_2$ acting on each mass. The virtual work of the constraint forces are

$$\begin{aligned} \delta W_1^{\text{(const)}} &= \boldsymbol{T}_1 . \, \delta \boldsymbol{r}_1 = T \delta y_1 \neq 0 \\ \delta W_2^{\text{(const)}} &= \boldsymbol{T}_2 . \, \delta \boldsymbol{r}_2 = T \delta y_2 \neq 0 \end{aligned} \tag{3.40}$$

Here we used the fact that $T_1 = T_2 = T$ since the string is massless. The displacements δy_1 and δy_2 must be compatible with the constraint

$$y_1 + y_2 = l - \pi R \implies \delta y_1 = -\delta y_2 \tag{3.41}$$

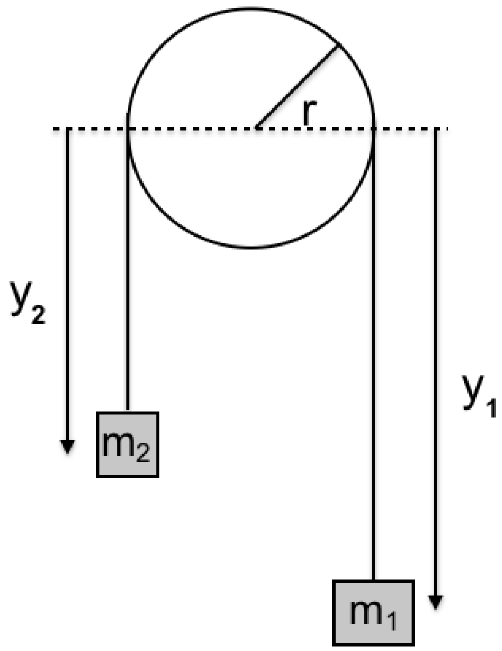


Figure 10: Atwood.

where l is the total length of the string and R is the radius of the pulley. Substituting the above relation between the virtual displacements into (3.40) we obtain the total virtual work done by the virtual forces to

$$\delta W_{\text{tot}}^{(\text{const})} = 0 \tag{3.42}$$

3.3 D'Alembert principle and equations of motion

Now, for a dynamical system of N-particles, where particle i is subject to a force $\boldsymbol{F}_i$, **Bernoulli** (1700- 1782) and **D'Alembert** (1713- 1783) noticed that since $\boldsymbol{F}_i - \frac{d\boldsymbol{p}_i}{dt} = 0$, one can imagine that if every particle i were given an effective force " $-\dot{\boldsymbol{p}}_i$", then the system would be in equilibrium. In this case, if we restrict our selves to systems for which the net virtual work of the constraint forces vanish, then

$$\sum_{i=1}^{N} (\boldsymbol{F}_i - \dot{\boldsymbol{p}}_i)\ \delta\boldsymbol{r}_i = \sum_{i=1}^{N} \left(\boldsymbol{F}_i^{(\text{app})} - \dot{\boldsymbol{p}}_i\right)\ \delta\boldsymbol{r}_i = 0 \tag{3.43}$$

Thus, we get

$$\delta W^{(\text{app})} := \sum_{i=1}^{N} \boldsymbol{F}_i^{(\text{app})}\ \delta\boldsymbol{r}_i = \sum_{i=1}^{N} \dot{\boldsymbol{p}}_i\ \delta\boldsymbol{r}_i \tag{3.44}$$

which is known as **D'Alembert principle**.

In terms of generalized coordinates $\{q_a, a = 1, ..., n\}$, the virtual work of the applied forces can be written as

$$\delta W^{(\text{app})} = \sum_{a=1}^{n} Q_a .\ \delta q_a \tag{3.45}$$

where the quantities Q_a are called the **generalized forces** defined as[16]

$$Q_a = \sum_{i=1}^{N} \boldsymbol{F}_i^{(\text{app})} .\ \frac{\partial \boldsymbol{r}_i}{\partial q_a} \tag{3.46}$$

The right hand side of equation (3.44) can written as

$$\sum_{i=1}^{N} \dot{\boldsymbol{p}}_i\ \delta \boldsymbol{r}_i = \sum_{i=1}^{N} \sum_{a=1}^{n} m_i \ddot{\boldsymbol{r}}_i\ \frac{\partial \boldsymbol{r}_i}{\partial q_a}\ \delta q_a \tag{3.47}$$

We consider the equation:

$$\frac{d}{dt}\left[m_i \dot{\boldsymbol{r}}_i \frac{\partial \boldsymbol{r}_i}{\partial q_a}\right] = \left[m_i \ddot{\boldsymbol{r}}_i \frac{\partial \boldsymbol{r}_i}{\partial q_a}\right] + \left[m_i \dot{\boldsymbol{r}}_i \frac{d}{dt}\frac{\partial \boldsymbol{r}_i}{\partial q_a}\right] \tag{3.48}$$

When summed over the index i, the second term on the right hand side gives

$$\sum_{i=1}^{N}\left(m_i \ddot{\boldsymbol{r}}_i \frac{\partial \boldsymbol{r}_i}{\partial q_a}\right) = \sum_{i=1}^{N}\left[\frac{d}{dt}\left(m_i \dot{\boldsymbol{r}}_i \frac{\partial \boldsymbol{r}_i}{\partial q_a}\right) - m_i \dot{\boldsymbol{r}}_i \frac{d}{dt}\frac{\partial \boldsymbol{r}_i}{\partial q_a}\right] \tag{3.49}$$

The cartesian components of the velocity vector, $\dot{\boldsymbol{r}}_i$, can be written in terms of the generalized velocities as follows:

$$\dot{\boldsymbol{r}}_i = \sum_{b=1}^{n} \frac{\partial \boldsymbol{r}_i}{\partial q_b} + \frac{\partial \boldsymbol{r}_i}{\partial t} \tag{3.50}$$

Since $\boldsymbol{r}_i$ are only functions of $q_a's$ and time, we have

$$\frac{\partial \dot{\boldsymbol{r}}_i}{\partial \dot{q}_a} = \frac{\partial \boldsymbol{r}_i}{\partial q_a} \tag{3.51}$$

[16] In deriving equation (3.46) we used the relation between $\delta \boldsymbol{r}_i = \sum_{a=1}^{n} \frac{\partial \boldsymbol{r}_i}{\partial q_a}\ \delta q_a$

Applying the chain rule to the time derivative of $\left(\frac{\partial \boldsymbol{r}_i}{\partial q_a}\right)$ and using Eq (3.50), we can write

$$\frac{d}{dt}\left(\frac{\partial \boldsymbol{r}_i}{\partial q_a}\right) = \sum_{b=1}^{n} \frac{\partial^2 \boldsymbol{r}_i}{\partial q_a \partial q_b}\dot{q}_b + \frac{\partial^2 \boldsymbol{r}_i}{\partial q_a \partial t} = \frac{\partial \dot{\boldsymbol{r}}_i}{\partial q_a} \tag{3.52}$$

Next, we substitute the above expression and the one of $\left(\frac{\partial \boldsymbol{r}_i}{\partial q_a}\right)$ and $\frac{d}{dt}\left(\frac{\partial \boldsymbol{r}_i}{\partial q_a}\right)$ given in in (3.51) to the right hand side of (3.49), yields

$$\sum_{i=1}^{N}\left[m_i \ddot{\boldsymbol{r}}_i \frac{\partial \boldsymbol{r}_i}{\partial q_a}\right] = \sum_{i=1}^{N}\left[\frac{d}{dt}\left(m_i \boldsymbol{v}_i \frac{\partial \boldsymbol{v}_i}{\partial \dot{q}_a}\right) - \left(m_i \boldsymbol{v}_i \frac{\partial \boldsymbol{v}_i}{\partial q_a}\right)\right] \tag{3.53}$$

By recognizing that

$$\begin{aligned} \sum_{i=1}^{N} m_i \boldsymbol{v}_i \frac{\partial \boldsymbol{v}_i}{\partial \dot{q}_a} &= \frac{\partial}{\partial \dot{q}_a}\left(\sum_{i=1}^{N} \frac{1}{2} m_i v_i^2\right) = \frac{\partial}{\partial \dot{q}_a} T, \\ \sum_{i=1}^{N} m_i \boldsymbol{v}_i \frac{\partial \boldsymbol{v}_i}{\partial q_a} &= \frac{\partial}{\partial q_a}\left(\sum_{i=1}^{N} \frac{1}{2} m_i v_i^2\right) = \frac{\partial}{\partial q_a} T \end{aligned} \tag{3.54}$$

we see that

$$\sum_{i=1}^{N} \dot{\boldsymbol{p}}_i \delta \boldsymbol{r}_i = \sum_{a=1}^{n}\left[\frac{d}{dt}\left(\frac{\partial T}{\partial q_a}\right) - \frac{\partial T}{\partial q_a}\right]\delta q_a \tag{3.55}$$

which according to D'Alembert principle implies that

$$\sum_{i=1}^{N}\left[\frac{d}{dt}\left(\frac{\partial T}{\partial q_a}\right) - \frac{\partial T}{\partial q_a} - Q_a\right]\delta q_a = 0 \tag{3.56}$$

If the only constraints that the system is subject to are of holonomic kind, such that the $\delta q's$ are independent, then the coefficients of each δq_a in Eq (3.56) must be zero, which leads to

$$\frac{d}{dt}\left(\frac{\partial T}{\partial q_a}\right) - \frac{\partial T}{\partial q_a} = Q_a, \quad a = 1, 2, ..., n \tag{3.57}$$

Thus, if we know the kinetic energy function $T(q, \dot{q}, t)$ of the system, we can write down the generalized forces without computing any constraints.

From now on, we will be interested in the applied forces that are conservative, i.e.

$$\boldsymbol{F}_i^{(\mathrm{app})} = -\boldsymbol{\nabla}_i \, U(\boldsymbol{r}_1, ..., \boldsymbol{r}_N) \tag{3.58}$$

then the generalized forces can written as

$$Q_a = -\sum_{i=1}^{N} \boldsymbol{\nabla}_i \, V . \frac{\partial \boldsymbol{r}_i}{\partial q_a} = -\frac{\partial V(q)}{\partial q_a}, \qquad a = 1, 2, ..., n \tag{3.59}$$

where $V(q) \equiv U(q_1, q_2, .., q_n)$ is the generalized potential which is independent on the generalized velocities. By defining the **Lagrangian function**:

$$\boxed{L(q, \dot{q}, t) := T(q, \dot{q}, t) - U(q)} \tag{3.60}$$

we can then write the expression (3.57) as

$$\boxed{\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_a} - \frac{\partial L}{\partial q_a} = 0, \qquad a = 1, 2, ..., n} \tag{3.61}$$

The above set of n-differential equations (typically of second order) are called the **Euler-Lagrange** equations of motion.

For a system whose configuration is described by n-generalized coordinates and a Lagrangian $L(q, \dot{q}, t)$, the generalized momentum p_a associated with q_a, also called **conjugate momentum** of the coordinate q_a, is defined by

$$p_a = \frac{\partial L}{\partial \dot{q}_a} = p_a(q, \dot{q}, t) \tag{3.62}$$

In particular, for a velocity independent potential, that is $\frac{\partial V}{\partial \dot{q}} = 0$, the generalized momenta reads

$$\boxed{p_a = \frac{\partial T}{\partial \dot{q}_a}} \tag{3.63}$$

If the Lagrangian of the system does not depend explicitly on a given q_b, then this coordinate is said to be **cyclic**. In this case, the equation of motion for q_b is

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_a}\right) = 0 \implies \frac{dp_b}{dt} \tag{3.64}$$

Thus, **conjugate momentum of a cyclic coordinate is conserved**.

Example 3.6: Particle in a plane

Consider a free particle of mass m moving on plane. It can be described by the generalized coordinates r and θ (the polar coordinates) and its kinetic energy reads

$$T = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) \tag{3.65}$$

Thus, the generalized forces are given by

$$\begin{aligned} Q_r &= \boldsymbol{F}^{\text{ (app)}}.\frac{\partial \boldsymbol{r}}{\partial r} = \boldsymbol{F}^{\ app}.\, \hat{r} \equiv F_r^{(\text{app})} \\ Q_r &= \boldsymbol{F}^{\text{ (app)}}.\frac{\partial \boldsymbol{r}}{\partial \theta} = \boldsymbol{F}^{\text{ (app)}}.\, r\hat{\theta} \equiv rF_\theta^{(\text{app})} \end{aligned} \tag{3.66}$$

where $\hat{r} = \boldsymbol{r}/r$, and $\hat{\theta}$ is a unit vector perpendicular to $\hat{r}$ in increasing θ. Using Eq (3.57), we obtain the equations of motion

$$m\ddot{r} - mr\dot{\theta}^2 = F_r^{(\text{app})} \tag{3.67}$$

$$\frac{d}{dt}(mr^2\dot{\theta}) = rF_\theta^{(\text{app})} \tag{3.68}$$

The first term in Eq (3.67) represents the radial acceleration and the second term is the centripetal acceleration. The second equation above is the statement that the rate of change of the angular momentum is equal to the torque.

Example 4.7: Block moving on an incline

Consider a block of mass m held motionless on frictionless incline plane of mass M and angle of inclination θ, which rests on a frictionless horizontal surface (see figure below). We wish to determine the acceleration of the two blocks after the release of the block m. For that, we denote by $X(t)$ the absolute displacement of the inclined plane with respect to a static frame of reference on the ground, and by $x(t)$ the displacement of the block m relative to an observer on the inclined plane. The position of m is

$$\begin{aligned} \overrightarrow{OP} &= \overrightarrow{OA} + \overrightarrow{AB} + \overrightarrow{BP} \\ &= X\ \boldsymbol{e}_X + h\ \boldsymbol{e}_Y + x\ \boldsymbol{u} \end{aligned} \tag{3.69}$$

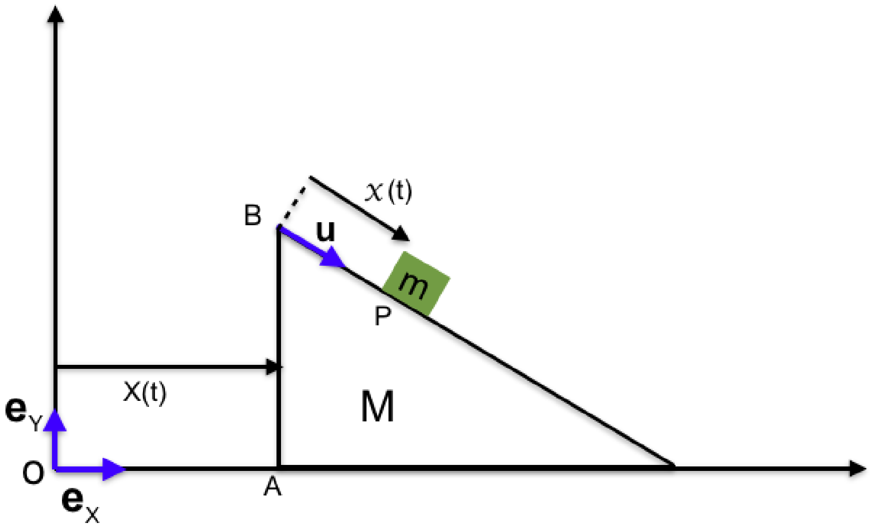


Figure 11: Block moving on an incline

where h is a constant which represents the initial hight on the inclined plane where the block was released. The vectors $\boldsymbol{e}_X, \boldsymbol{e}_Y$ and $\boldsymbol{u}$ are unit vectors along the X-axis, Y-axis, and the axis along the incline. Thus, the velocity of the block m reads

$$\boldsymbol{v} = \dot{X}\ \boldsymbol{e}_X + \dot{x}\ \boldsymbol{u} \tag{3.70}$$

The kinetic energy of the system block-incline is

$$\begin{aligned} T &= \frac{1}{2} M \dot{X}^2 + \frac{1}{2} m v^2 \\ &= \frac{1}{2} M \dot{X}^2 + \frac{1}{2} m \left(\dot{X}^2 + \dot{x}^2 - 2\dot{x}\dot{X}\cos\theta \right) \end{aligned} \tag{3.71}$$

If we choose the gravitational potential to be zero at $Y = h$, then the potential energy of the system is given by

$$V = -mgx\sin\theta + V^{(\text{incline})} \tag{3.72}$$

Here $V^{(\text{incline})}$ is the gravitational potential energy of the incline which is constant, and hence does not enter into the equation of motion. So, we can write the Lagrangian of the system block-incline as

$$L = \frac{1}{2} M \dot{X}^2 + \frac{1}{2} m \left(\dot{X}^2 + \dot{x}^2 - 2\dot{x}\dot{X}\cos\theta \right) + mg\sin\theta \tag{3.73}$$

from which we obtain the equations of motion

$$
\begin{aligned}
(M+m)\ddot{X} - m\cos\theta\ddot{x} &= 0 \\
-m\cos\theta\ddot{X} + m\ddot{x} &= mg\sin\theta
\end{aligned}
\tag{3.74}
$$

Hence, the block and the inclined plane move with the constant accelerations

$$
a^{(\text{incline})} := \ddot{X} = \frac{m}{(M+m\sin^2\theta)\sin\theta\cos\theta}g \tag{3.75}
$$

$$
a^{(\text{block})} := \ddot{x} = \frac{(M+m)}{(M+m\sin^2\theta)\sin\theta}g \tag{3.76}
$$

Note that for $M >> m$, then $a^{(\text{block})} \simeq g\sin\theta$ and $a^{(\text{incline})} \simeq \frac{m}{2M}\sin 2\theta$. So, in the limit where the ratio m/M is negligibly small, we can consider, to a good approximation, the incline plane to be static if it was initially at rest.

3.4 Non-holonomic constraints and Lagrange multipliers

In deriving the Euler-Lagrange equations (3.61) we assumed that the generalized coordinates are independent, which is the case if the system is subject only to holonomic constraints. For non-holonomic constraints, there will be more generalized coordinates than the number of degrees of freedom. Thus, the δq's are no longer independent for a virtual displacement consistent with the constraints.

Let us suppose that there are a set of "m" non-holonomic equations of the form

$$
\sum_{a=1}^{n} A_{\alpha a}\, dq_a + A_{\alpha t}\, dt = 0; \qquad \alpha = 1,..,m \tag{3.77}
$$

where n is the number of chosen generalized coordinates to describe the system. The displacements δq's must satisfy the following condition

$$
\sum_{a=1}^{n} A_{\alpha a}\, dq_a = 0, \qquad \alpha = 1,..,m \tag{3.78}
$$

In addition, we assume that each generalized applied force Q_a is obtained from a potential function $U(\mathbf{q})$. For **the generalized constraint forces**, $\mathcal{C}_a$, we take their net virtual work equal to zero, i.e.

$$
\sum_{a=1}^{n} C_a\, \delta q_a = 0 \tag{3.79}
$$

We Multiply both sides of the equation (3.78) by factors λ_α, known as **Lagrange multipliers**, and obtain

$$\lambda_\alpha \sum_{a=1}^{n} C_a \; \delta q_a = 0 \tag{3.80}$$

Now, we subtract the sum of these m equations from equation (3.79) gives

$$\sum_{a=1}^{n} \left[C_a - \sum_{\alpha=1}^{m} \lambda_\alpha A_{\alpha a} \right] \; \delta q_a = 0 \tag{3.81}$$

Up to this point the $\lambda' s$ have been considered to be arbitrary while the $\delta q' s$ must satisfy the constraint equations. However, if one chooses the Lagrange multipliers such that

$$C_a = \sum_{\alpha=1}^{m} \lambda_\alpha A_{\alpha a} \tag{3.82}$$

then in this case the relation (3.81) will be satisfied for arbitrary $\delta q' s$. Thus, the $\delta q' s$ can be chosen independently, which implies that

$$-\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_a}\right) + \frac{\partial T}{\partial q_a} = -\frac{\partial V}{\partial q_a} + C_a \tag{3.83}$$

The above equation can be re-written as

$$\boxed{\tfrac{d}{dt}\left(\tfrac{\partial L}{\partial \dot{q}_a}\right) - \tfrac{\partial L}{\partial q_a} = \textstyle\sum_{\alpha=1}^{m} \lambda_\alpha A_{\alpha a}; \qquad a = 1, ..., n} \tag{3.84}$$

These equations are the Lagrange's equations for a system with a non-holonomic constraints that have the form given in Eq (3.77). So, in total we have (n + m) equations with which to solve for the (n+ m) independent variables, namely n generalized coordinates and m Lagrange multipliers.

Note that the above Lagrange equations can also be applied to holonomic systems in which there are more generalized coordinates than the number of degrees of freedom. For instance, suppose there are m holonomic constraint equations of the form

$$\Phi_\alpha(q_1, ..., q_n; t) = 0 \tag{3.85}$$

By differentiating the left hand side of the above equation, and comparing with equation (3.77), we obtain

$$A_{\alpha a} = \frac{\partial \Phi_\alpha}{\partial q_a}, \qquad A_{\alpha t} = \frac{\partial \Phi_\alpha}{\partial t} \tag{3.86}$$

If we apply the Lagrange multiplier method to this holonomic system and the resulting differential equations are completely solved, then the $q's$ and $\lambda's$ are expressed as explicit functions of time. Hence, the generalized constraint forces $\mathcal{C}_a$ can also be obtained as explicit functions of time.

Example 3.8: Block moving on an incline: Lagrange multiplier

Let us consider the same system of the example 2.7, but now we wish to use the method of Lagrange multiplier to solve the interaction force between the two blocks. This force is normal to the frictionless contact surface and may be considered as generalized constraint force associated with the y coordinate.

Although we are using three generalized coordinates, $(x,\ X,\ y)$, the system has only two degrees of freedom. This is because there is one equation of holonomic constraint $y = 0$ which can be expressed as

$$\dot{y} = 0 \implies A_{11} = A_{12} = A_{1t} = 0; \qquad A_{13} = 1 \tag{3.87}$$

implying

$$C_1 = C_2 = 0, \qquad C_3 = \lambda \tag{3.88}$$

By writing separately the horizontal and vertical velocity components of the block that is sliding on the incline, the kinetic and the potential energies of the system read

$$\begin{aligned} T &= \frac{1}{2}M\dot{X}^2 + \frac{1}{2}m\left[\Big(\dot{X} - (\dot{x}\cos\theta + \dot{y}\sin\theta)\Big)^2 + (\dot{y}\cos\theta + \dot{x}\sin\theta)^2\right] \\ V &= mg(y\cos\theta - x\sin\theta) + \text{constant} \end{aligned} \tag{3.89}$$

Thus, the Lagrangian is given by

$$\begin{aligned} L = \frac{1}{2}(M+m)\dot{X}^2 + \frac{1}{2}m\left[\dot{x}^2 1\dot{y}^2 - 2(\dot{x}\cos\theta + \dot{y}\sin\theta)\dot{X}\right] \\ - mg(y\cos\theta - x\sin\theta) + \text{cosntant} \end{aligned} \tag{3.90}$$

Hence, we derive the following equations of motion

$$
\begin{aligned}
(M+m)\ddot{X} - m\cos\theta\ \ddot{x} - m\sin\theta\ \ddot{y} &= 0 \\
-m\cos\theta\ \ddot{X} + m\ \ddot{x} - mg\sin\theta &= 0 \\
-m\sin\theta\ \ddot{X} + m\ \ddot{y} - mg\cos\theta &= \lambda_1
\end{aligned}
\tag{3.91}
$$

with the constraint equation $\dot{y} = 0$, we obtain

$$
\ddot{X} = \frac{m\sin\theta\cos\theta}{(M+m\sin^2\theta)}g, \qquad \ddot{x} = \frac{m\sin\theta}{(M+m\sin^2\theta)}g \tag{3.92}
$$

and the constraint force given by

$$
C_3 = \lambda_1 = \frac{Mm\cos\theta}{(M+m\sin^2\theta)}g \tag{3.93}
$$

Example 3.9: Rigid rod rotating: Lagrange multiplier

In this example we consider two particles of identical masses, m, connected by a rigid rod of length l which rotates in a horizontal plane with a constant angular velocity ω. A knife-edge supports at the two particles prevent the particle from having a velocity component along the the rod, but the particles can slide without friction in a direction perpendicular to the rod. We would like to solve for the coordinates of the center of mass system (CMS), (x_{cm}, y_{cm}) and the constraint force as function of time assuming that the CM is initially at the origin and has a speed v_0 in the positive y-direction. This system has two independent equations of holonomic constraints:

$$
\begin{aligned}
(x_2 - x_1)^2 + (y_2 - y_1)^2 &= l^2 \\
(y_2 - y_1) - (x_2 - x_1)\tan\omega t &= 0
\end{aligned}
\tag{3.94}
$$

and one non-holonomic constraint equation

$$
(\dot{x}_1 + \dot{x}_2)\cos\omega t + (\dot{y}_1 + \dot{y}_2)\sin\omega t \tag{3.95}
$$

which restricts the velocity of the CM of the rod to a direction which is perpendicular to the rod. So, we can choose (x_{cm}, y_{cm}) as the generalized coordinates that describes the rod. The coordinates of the two masses can be expressed in terms of the CMS coordinates as

$$x_1 = x_{cm} - \frac{l}{2}\cos\omega t, \quad y_1 = y_{cm} - \frac{l}{2}\sin\omega t \tag{3.96}$$
$$x_2 = x_{cm} + \frac{l}{2}\cos\omega t, \quad y_2 = y_{cm} + \frac{l}{2}\sin\omega t$$

The total kinetic energy of the system is

$$T = \frac{1}{2}m(\dot{x}_{cm}^2 + \dot{x}_{cm}^2) + \frac{1}{4}ml^2\omega^2 \tag{3.97}$$

whereas the potential is zero. So, we can write

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_a}\right) - \frac{\partial T}{\partial q_a} = \lambda_1 A_{1a}, \qquad a = 1, 2 \tag{3.98}$$

where here $q_1 = x_{cm}, q_2 = y_{cm}, A_{11} = \cos\omega t, A_{12} = \sin\omega t$, and $A_{22} = 0$. This yields

$$2m\ddot{x}_{cm} = \lambda_1 \cos\omega t, \quad 2m\ddot{y}_{cm} = \lambda_1 \sin\omega t \tag{3.99}$$

By using the constraint equation

$$\dot{y}_{cm} = \dot{x}_{cm}\tan\omega t \implies \ddot{y}_{cm} = \ddot{x}_{cm}\tan\omega t \tag{3.100}$$

we get

$$\frac{d}{dt}(\dot{x}_{cm}^2 + \dot{y}_{cm}^2) = 0 \implies (\dot{x}_{cm}^2 + \dot{y}_{cm}^2) = \upsilon_0^2 \tag{3.101}$$

which shows that the CM has a uniform motion with a constant speed υ_0. Since the direction of the motion is always perpendicular to the rod we have

$$\dot{x}_{cm} = -\upsilon_0 \sin\omega t, \quad \dot{y}_{cm} = \upsilon_0 \cos\omega t, \tag{3.102}$$

Integrating the equations above and using the initial conditions yields

$$x_{cm} = \frac{\upsilon_0}{\omega}(\cos\omega t - 1), \qquad y_{cm} = \frac{\upsilon_0}{\omega}\sin\omega t, \tag{3.103}$$

Thus, this system moves in a circular path of radius υ_0/ω at constant speed υ_0.

The Lagrange multiplier is given by

$$\lambda_1 = -2mv_0\omega \tag{3.104}$$

which leads to the constraint forces:

$$C_1 = -2mv_0\omega\cos\omega t, \quad C_2 = -2mv_0\omega\sin\omega t \tag{3.105}$$

3.5 Rayleigh's dissipation function

Consider systems for which the generalized forces are not totally derivable from a potential function, but have the form

$$\mathcal{Q}_a = -\frac{\partial V}{\partial q_a} + \mathcal{Q}'_a \tag{3.106}$$

Then, according to (3.57), the equations of motion read

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_a}\right) - \frac{\partial L}{\partial q_a} = \mathcal{Q}'_a, \quad a = 1, 2, .., n \tag{3.107}$$

with $L = T - V$ is the Lagrangian which we introduced earlier in this chapter. In addition we will assume that the generalized forces $\mathcal{Q}'_a$ are function of the generalized velocities $\dot{q}_a$ and have the form

$$\mathcal{Q}'_a = -\sum_{b=1}^{n} \mathcal{C}_{ab}\,\dot{q}_b \tag{3.108}$$

where the coefficients $\mathcal{C}_{ab}$ form a real, symmetric matrix, known as **the damping coefficients**. The $\mathcal{Q}'_a$ are **generalized friction forces** which are dissipative in nature and result in a loss if energy.

We define **Rayleigh**'s dissipation function, $\mathcal{F}(q, \dot{q}, t)$, as follows:

$$\mathcal{F}(q, \dot{q}, t) = \tfrac{1}{2}\sum_{a,b=1}^{n} \mathcal{C}_{ab}\,\dot{q}_a\dot{q}_b \tag{3.109}$$

In this case the equations of motion can be written as

$$\tfrac{d}{dt}\left(\tfrac{\partial L}{\partial \dot{q}_a}\right) - \tfrac{\partial L}{\partial q_a} + \tfrac{\partial \mathcal{F}}{\partial \dot{q}_a} = 0, \quad a = 1, 2, ..., n \tag{3.110}$$

Note that

$$\sum_{a=1}^{n} \mathcal{Q}'_a \dot{q}_a = -2\mathcal{F} \tag{3.111}$$

which means that **Rayleigh**'s dissipation function represents the rate at which the friction forces do work on the system.

3.6 Velocity dependent potentials

If the generalizes forces $\mathcal{Q}'_a$, are obtained from a velocity dependent potential function $U(q, \dot{q}, t)$ in accordance with

$$\mathcal{Q}_a = \frac{d}{dt}\left(\frac{\partial U}{\partial \dot{q}_a}\right) - \frac{\partial U}{\partial q_a} \tag{3.112}$$

then, we can write

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_a}\right) - \frac{\partial L}{\partial q_a} = 0, \quad a = 1, 2, ..., n \tag{3.113}$$

where

$$L = T - U \tag{3.114}$$

As an example of such kind of velocity dependent potential, consider the electromagnetic (EM) force that acts on a charged particle[51]

$$\boldsymbol{F} = e_q(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}) \tag{3.115}$$

Here e_q is the electric charge of the particle, $\boldsymbol{E}$ is the electric field and $\boldsymbol{B}$ is the magnetic field that the particle is subject to. The fields $\boldsymbol{E}$ and $\boldsymbol{B}$ can be obtained from a scalar potential $\phi(\vec{x}, t)$ and a vector potential $\boldsymbol{A}(\boldsymbol{x}, t)$ as:

$$\begin{aligned} \boldsymbol{E} &= -\boldsymbol{\nabla}\phi - \frac{\partial \boldsymbol{A}}{\partial t}, \\ \boldsymbol{B} &= \boldsymbol{\nabla} \times \boldsymbol{A} \end{aligned} \tag{3.116}$$

Substituting the above expressions into Eq (3.115), we get

[51] I am using the S.I system of units.

$$\boldsymbol{F} = e_q\left(-\boldsymbol{\nabla}\phi - \frac{\partial \boldsymbol{A}}{\partial t} + \boldsymbol{v} \times \boldsymbol{\nabla} \times \boldsymbol{A}\right) \tag{3.117}$$

Now let us write down explicitly one of the components, say the x-component, of the last term in the bracket in terms of the components of $\boldsymbol{v}$, and $\boldsymbol{A}$:

$$\begin{aligned}(\boldsymbol{v} \times \boldsymbol{\nabla} \times \boldsymbol{A})_x &= v_y\partial_x A_y + v_z\partial_x A_z - v_y\partial_y A_x - v_z\partial_z A_x \qquad (3.118)\\ &= (v_y\partial_x A_y + v_z\partial_x A_z + v_x\partial_x A_x) - (v_x\partial_x A_x + v_y\partial_y A_x + v_z\partial_z A_x)\end{aligned}$$

or, equivalently,

$$(\boldsymbol{v} \times \boldsymbol{\nabla} \boldsymbol{A})_x = \frac{\partial}{\partial x}(\boldsymbol{v}.\boldsymbol{A}) - \frac{dA_x}{dt} + \frac{\partial A_x}{\partial t} \tag{3.119}$$

Thus, the x-component of the electromagnetic force reads

$$F_x = e_q\left(-\frac{\partial \phi}{\partial x} + \frac{\partial}{\partial x}(\boldsymbol{v}.\boldsymbol{A}) - \frac{dA_x}{dt}\right) \tag{3.120}$$

Noting that

$$\frac{dA_x}{dt} = \frac{d}{dt}\left[\frac{\partial}{\partial \dot{x}}(\boldsymbol{v}.\boldsymbol{A})\right] \tag{3.121}$$

That implies that F_x can be re-written in the form

$$F_x = \frac{d}{dt}\left(\frac{\partial U}{\partial \dot{x}} - \frac{\partial U}{\partial x}\right) \tag{3.122}$$

where

$$U = e\left(\phi - \boldsymbol{v}.\boldsymbol{A}\right) \tag{3.123}$$

Similarly, we find that the y and the z components of $\boldsymbol{F}$ can be written in exactly the same form as F_x, i.e.

$$\begin{aligned}F_y &= \frac{d}{dt}\left(\frac{\partial U}{\partial \dot{y}} - \frac{\partial U}{\partial y}\right)\\ F_z &= \frac{d}{dt}\left(\frac{\partial U}{\partial \dot{z}} - \frac{\partial U}{\partial z}\right)\end{aligned}$$

Hence, the the Lagrangian of a charged particle of mass m in an electromagnetic field is given by

$$\boxed{L^{(\mathrm{EM})} = \tfrac{1}{2} m \dot{\boldsymbol{r}}^2 - e_q \left(\phi - \boldsymbol{v}.\boldsymbol{A}\right)} \tag{3.124}$$

The generalized momenta associated with a charged particle moving in an electromagnetic field is then[52]

$$p_i := \frac{\partial L}{\partial \dot{x}_i} = m\dot{x}_i + eA_i \tag{3.125}$$

That is

$$\boxed{\boldsymbol{p} = m\boldsymbol{v} + e_q \boldsymbol{A}} \tag{3.126}$$

This shows that part of the particle's momentum is associated with the electromagnetic field. Note that since $L^{(\mathrm{EM})}$ does not depend explicitly on the position, we have

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}_i}\right) = 0 \implies \frac{dp_i}{dt} = 0 \tag{3.127}$$

which means that **the generalized momentum**, rather than $m\dot{\boldsymbol{r}}$, of a charged particle in an electromagnetic field is a conserved quantity.

Example 3.10: Charged Particle in Constant Magnetic Field

Consider a charged particle of mass m and charge e_q moving in a uniform magnetic field directed along the z-axis, i.e. $\boldsymbol{B} = B_0 \, \hat{\boldsymbol{e}}_z$.

From the definition of the magnetic field vector as the curl of a vector potential, we have

$$B_0 = \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \tag{3.128}$$

By choosing the vector potential $\boldsymbol{A}$ and the scalar potential ϕ to have the form

$$\boldsymbol{A} = xB_0 \, \hat{\boldsymbol{e}}_y, \qquad \phi = 0 \tag{3.129}$$

[52] In the relativistic limit, the Lagrangian of a charged in an electromagnetic field reads

$$L = -mc^2 \sqrt{1 - \left(\frac{\dot{\boldsymbol{r}}}{c}\right)^2} + e_q \boldsymbol{A}(\boldsymbol{r}).\dot{\boldsymbol{r}} - e\phi(\boldsymbol{r})$$

yields $\boldsymbol{B} = B_0\, \hat{\boldsymbol{e}}_z$ and vanishing electric field. So, the Lagrangian reads

$$L = \frac{m}{2}\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right) + e_q B_0 x\dot{y} \tag{3.130}$$

Then, the equations of motion are

$$\left.\begin{aligned} m\ddot{x} - e_q B\dot{y} = 0, \\ m\ddot{y} + e_q B\dot{x} = 0, \\ m\ddot{z} = 0 \end{aligned}\right\} \quad \Rightarrow \begin{aligned} \frac{d}{dt}(m\dot{x}) - e_q B\dot{y} = 0, \\ \frac{d}{dt}\left(m\dot{y} + e_q Bx\right) = 0, \\ m\ddot{z} = 0 \end{aligned} \tag{3.131}$$

The second and the third equations yield

$$\dot{y} = a - \frac{e_q B}{m} x \tag{3.132}$$

$$z = z_0 + bt \tag{3.133}$$

where a, b, and z_0 are constant that can be determined from the initial conditions. With the use of the above expression of $\dot{y}$ the first equation in (3.131) becomes a second order differential equation just in the variable x:

$$\ddot{x} + \omega^2 (x - x_0) = 0, \quad \omega = \frac{|e_q| B}{m} \tag{3.134}$$

It has the general solution

$$x = x_0 + A\cos\left(\omega t + \delta\right) \tag{3.135}$$

Substituting (3.135) into equation (3.132) and then integrating over the time variable, yields

$$y = y_0 - A\sin\left(\omega t + \delta\right) \tag{3.136}$$

Note that the component of the particle's velocity in the xy plane are

$$\upsilon_x = -A\omega \sin\left(\omega t + \delta\right), \qquad \upsilon_y = -A\omega\cos\left(\omega t + \delta\right) \tag{3.137}$$

Hence, the constant A can written in terms of the oscillation frequency ω and the transverse velocity $\upsilon_\perp$ as

$$A^2 = \frac{v_\perp^2}{\omega^2} \tag{3.138}$$

By combining the above equation with Eq (3.135), we obtain

$$(x - x_0)^2 + (y - y_0)^2 = R_L^2, \qquad R_L = \frac{|v_\perp|}{\omega} \tag{3.139}$$

and so the equation of circle with radius R_L, known as the **Larmor radius**, in the plane xy. Now, together with the equation of z given by the third equation in (3.131), the trajectory of the charged particle defines a **helix**[53].

[53]Of course we obtain the same result by just applying newton's second law. Below I give the details of how to show that he trajectory is a helix. The equation of motion is

$$m\frac{d\boldsymbol{v}}{dt} = e_q \boldsymbol{v} \times \boldsymbol{B}$$

which in components form reads

$$m\dot{v_x} = e_q B_0 v_y, \quad m\dot{v_y} = -e_q B_0 v_x, \quad m\dot{v_z} = 0$$

The last equation implies that v_z is constant, which is not surprising since there is no force in this direction due to the fact that the magnetic field, which along the z-direction, is perpendicular to the particle's velocity. Taking the time derivative of the first equation and using the second equation for v_y, we get

$$\ddot{v_x} + \omega^2 v_x = 0, \quad \omega = \frac{|e_q| B}{m}$$

which has the solution

$$v_x = C\cos(\omega t + \delta) \quad \Rightarrow v_y = \mp C\cos(\omega t + \delta)$$

where minus sign in the righthand side of the v_y is for positive e_q and the positive sign for negative charge. The constant C is determined by the initial condition and physically is the component of the particle's velocity in the direction perpendicular to the direction of the magnetic field, i.e $C = v_\perp$.

Integrating v_x and v_y over the time variable, we obtain

$$x - x_0 = \frac{v_\perp}{\omega}\sin(\omega t + \delta), \quad y - y_0 = \pm\frac{v_\perp}{\omega}\cos(\omega t + \delta)$$

Combining the two equations above we obtain

$$(x - x_0)^2 + (y - y_0)^2 = R_L^2, \quad R_L = \frac{|v_\perp|}{\omega}$$

which with the equation $(z - z_0) \propto t$ implies that the trajectory of particle is a helix.

3.7 Symmetries and conservation laws

For a system of n degrees of freedom there are n differential equations of a second order in time. The solution of each of them requires two integrations, and hence the system needs 2n constants of integration, which can be determined by the initial conditions of the system (e.g.,the initial values of the generalized positions and velocities). However, sometimes, one can learn about a system by resorting to a complete solution of the equations of motion. This will occur whenever the Lagrangian of the system exhibits some symmetry. In this case, first integrals can be found easily, and take the form of equations like

$$f(q, \dot{q}, t) = \text{constant} \tag{3.140}$$

Let us consider the coordinate transformation

$$q_a(t) \;\rightarrow\; \tilde{q}_a(q, \xi) \tag{3.141}$$

where ξ is a continuous parameter. Without loss of generality, we take $\tilde{q}_a(q, 0) = q_a$. If the Lagrangian of the system is invariant under the transformation above, then

$$0 = \frac{dL}{d\xi}(\tilde{q}, \dot{\tilde{q}}, t)|_{\xi=0} = \frac{\partial L}{\partial q_a}\frac{\partial \tilde{q}_a}{\partial \xi}|_{\xi=0} + \frac{\partial L}{\partial \dot{q}_a}\frac{\partial \dot{\tilde{q}}_a}{\partial \xi}|_{\xi=0} \tag{3.142}$$

Using the equation of motion for q_a, we can replace $\frac{\partial L}{\partial q_a}$ by $\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_a}\right)$ in the above equation, and obtain

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_a}\right)\frac{\partial \tilde{q}_a}{\partial \xi}|_{\xi=0} + \frac{\partial L}{\partial \dot{q}_a}\frac{\partial \dot{\tilde{q}}_a}{\partial \xi}|_{\xi=0} = 0 \tag{3.143}$$

We also note that

$$\frac{d}{dt}\left(\frac{\partial \tilde{q}_a}{\partial \xi}\right) = \frac{\partial \dot{\tilde{q}}_a}{\partial \xi} \tag{3.144}$$

which implies that equation (3.143) can be written as

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_a}.\frac{\partial \tilde{q}_a}{\partial \xi}\right)|_{\xi=0} = 0 \tag{3.145}$$

This means that the quantity

$$\boxed{\mathcal{Q}_a = \frac{\partial L}{\partial \dot{q}_a}.\frac{\partial \tilde{q}_a}{\partial \xi}|_{\xi=0} = \text{Constant}} \tag{3.146}$$

We say that $\mathcal{Q}_a$ is a **conserved charge**, associated with the symmetry transformation (3.141). This statement is known as **Noether's theorem**[54]. Below we will discuss the charges associated with different space time symmetries.

3.7.1 Invariance under translation

Consider a system with a Lagrangian L which does not depend explicitly on some generalized coordinate q_*; that is

$$\frac{\partial L}{\partial q_*} = 0 \tag{3.147}$$

Hence, q_* is a cyclic coordinate. This is equivalent to saying that L is invariant under the translation of q_* by an arbitrary constant, c, i.e.

$$q_* \ \rightarrow \ q_* + c \tag{3.148}$$

Thus,

$$\frac{\partial \tilde{q}}{\partial \xi}|_{\xi=0} = 1 \ \Longrightarrow \ \mathcal{Q}_* = \frac{\partial L}{\partial \dot{q}_*} = p_* \tag{3.149}$$

So, the conserved charge associated with the generalized coordinate is the corresponding conjugate momentum.
Now let us generalize this result to a system of N particles described by the Lagrangian

$$L = \frac{1}{2}\sum_{i=1}^{N} m_i(\dot{\boldsymbol{r}}_i)^2 - \sum_{i\neq j} V(\boldsymbol{r}_i - \boldsymbol{r}_j) \tag{3.150}$$

This Lagrangian is invariant under the transformation

$$\boldsymbol{r}_i \ \rightarrow \ \boldsymbol{r}_i + a\ \boldsymbol{n} \tag{3.151}$$

for any vector $\boldsymbol{n}$ and arbitrary real number a. The above transformation is just the translations of a vector in three dimension, which are elements of the Galilean group. According to Noether theorem we have

$$\sum_{i=1}^{N} \frac{\partial L}{\partial \dot{\boldsymbol{r}}_i}.\ \boldsymbol{n} = \sum_{i=1}^{N} \boldsymbol{p}_i.\ \boldsymbol{n} = \text{constant} \tag{3.152}$$

which is just the total linear momentum in the direction along $\boldsymbol{n}$. But since $\boldsymbol{n}$ is arbitrary vector, the total momentum of the system of N particles is conserved, i.e.

[54]Named after **Emmy Noether** (1882- 1935), one of the great mathematicians of the 20 th century.

$$\boxed{\boldsymbol{P}_{\rm tot} = \sum_{i=1}^{N} \boldsymbol{p}_i = \text{constant}} \tag{3.153}$$

Thus, if a Lagrangian is invariant is invariant under the spatial translations[55], then the **total linear momentum of the system is conserved**.

3.7.2 Invariance under rotation

Suppose that the potential considered of a system of N particles only depends on the mutual distances $|\boldsymbol{r}_i - \boldsymbol{r}_j|$ between the particles, i.e.

$$V = V(|\boldsymbol{r}_1 - \boldsymbol{r}_2|, ..., |\boldsymbol{r}_1 - \boldsymbol{r}_N|, |\boldsymbol{r}_2 - \boldsymbol{r}_N|, ...|\boldsymbol{r}_{N-1} - \boldsymbol{r}_N|) \tag{3.154}$$

then the Lagrangian is invariant under over all rotation of the system. For an infinitesimal rotation of angle ϵ about an arbitrary axis parameterized by an unit vector $\boldsymbol{n}$, is given by

$$\boldsymbol{r}_i \;\rightarrow\; \boldsymbol{r}_i + \epsilon\; \boldsymbol{n} \times \boldsymbol{r}_i \tag{3.155}$$

The corresponding conserved charge is

$$\sum_{i=1}^{N} \frac{\partial L}{\partial \dot{\boldsymbol{r}}_i}.\; (\boldsymbol{n} \times \boldsymbol{r}_i) = \sum_{i=1}^{N} \boldsymbol{p}_i.\; (\boldsymbol{n} \times \boldsymbol{r}_i) = \text{constant} \tag{3.156}$$

Using the vector identity $\boldsymbol{A}.(\boldsymbol{B}\times\boldsymbol{C}) = \boldsymbol{B}.(\boldsymbol{C}\times\boldsymbol{A})$, we can write the above conservation law as

$$\sum_{i=1}^{N} \boldsymbol{n}.\; (\boldsymbol{r}_i \times \boldsymbol{p}_i) = \boldsymbol{n}.\; \boldsymbol{L}_{\rm tot} = \text{constant} \tag{3.157}$$

Here $\vec{L}_{\rm tot} = \sum_i (\boldsymbol{r}_i \times \boldsymbol{p}_i)$ is the total orbital angular momentum of the system. Since the vector $\boldsymbol{n}$ is arbitrary, we conclude that

$$\boxed{\boldsymbol{L}_{\rm tot} = \text{constant}} \tag{3.158}$$

Therefore, the conservation of the total angular momentum of a system is a direct consequence of the invariance of the Lagrangian under arbitrary over all rotations.

[55] This means that the space is homogeneous.

3.7.3 Invariance under time-translation

We consider the transformation that shifts the time by a constant τ but leaves the generalized coordinates and their velocities unchanged, i.e.

$$\begin{aligned} t &\longrightarrow t' = t + \tau \\ q_a(t) &\longrightarrow q_a(t)' = q_a(t) \\ \dot{q}_a(t) &\longrightarrow \dot{q}_a(t') = \dot{q}_a(t) \end{aligned} \tag{3.159}$$

The invariance of a Lagrangian under the above transformation, means that

$$\frac{dL}{d\tau}|_{\tau=0} := \frac{\partial t}{\partial \tau}|_{\tau=0}\frac{\partial L}{\partial t} + \sum_a \frac{\partial q_a}{\partial \tau}|_{\tau=0}\frac{\partial L}{\partial q_a} + \sum_a \frac{\partial \dot{q}_a}{\partial \tau}|_{\tau=0}\frac{\partial L}{\partial \dot{q}_a} = 0 \tag{3.160}$$

which implies that

$$\frac{\partial L}{\partial t} = 0 \tag{3.161}$$

Thus, a Lagrangian of a system is invariant under time translation if does not depend explicitly on time. Now let us write the total derivative of L with respect to time as[18]

$$\begin{aligned} \frac{dL}{dt} &= \frac{\partial L}{\partial t} + \sum_{a=1}^{n}\left[\frac{\partial L}{\partial \dot{q}_a}\ \ddot{q}_a + \frac{\partial L}{\partial q_a}\ \dot{q}_a\right] \\ &= \frac{\partial L}{\partial t} + \sum_{a=1}^{n}\left[\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_a}\right)\ \ddot{q}_a + \frac{\partial L}{\partial q_a}\ \ddot{q}_a\right] \end{aligned} \tag{3.162}$$

where in obtaining the second equality we used the Euler-Lagrange equations of motion. We can re-write (3.162) in the following form

$$\frac{d\mathcal{H}}{dt} = -\frac{\partial L}{\partial t} \tag{3.163}$$

where the function h is defined by

$$h = \sum_{a=1}^{n} \frac{\partial L}{\partial \dot{q}_a}\ \dot{q}_a - L \tag{3.164}$$

Hence, if the Lagrangian does not explicitly depend on time, i.e. invariant under time translation, then

[18]Note that in general dL/dt does not vanish because for any given solution of the equations of motion the value of the Lagrangian depends on time through the time dependence of the $q's$ and $\dot{q}'s$.

$$h := \sum_{a=1}^{n} \frac{\partial L}{\partial \dot{q}_a}\, \dot{q}_a - L = \text{constant} \tag{3.165}$$

Under certain assumptions the function h is **the total energy** of the system. This is the case if the coordinate transformation equations are not explicitly dependent on time and if the potential is independent of the generalized velocities. As we will see later in these notes, the function h is identical in value with the Hamiltonian $\mathcal{H}$ of a system. The reason that we are here using a lower case for the energy function is that it is a function of $q's$ and $\dot{q}'s$, whereas $\mathcal{H}$ is a function of the $q's$ and their corresponding conjugate momenta $p's$.

3.7.4 Scale invariance of equations of motion

Let L be the Lagrangian L of some system. Then, we can describe obtain the same solutions to the equations of motion for the system if instead of using L we use a new Lagrangian $L' = \alpha L$, where α is a constant. So, one can ask under which kind of transformations of the coordinates and/or time will result in a rescaling of the Lagrangian. For that we will consider the transformation of the form

$$q_a \;\longrightarrow\; A\, q_a; \qquad t \;\longrightarrow\; B\, t \tag{3.166}$$

where A and B are constants. Then, the kinetic energy transforms as

$$T(q, \dot{q}) \;\longrightarrow\; \left(\frac{A}{B}\right)^2 T(q, \dot{q}) \tag{3.167}$$

If assume that the potential, V, of the system is an homogeneous function of the generalized coordinates of degree D, then

$$V(q_1, q_2, .., q_n) \;\longrightarrow\; V(Aq_1, Aq_2, .., Aq_n) = A^D V(q_1, q_2, .., q_n) \tag{3.168}$$

Thus, under the transformation (3.166), the Lagrangian would be changed by an over all constant, provided that

$$\frac{A^2}{B^2} = A^D \;\Longrightarrow\; B = A^{1-\frac{D}{2}} \tag{3.169}$$

So, in this case the transformation

$$q_a \;\longrightarrow\; A\, q_a; \qquad t \;\longrightarrow\; A^{1-\frac{D}{2}}\, t \tag{3.170}$$

yields that $L \;\rightarrow\; A^D\, L$, and hence leaves the equation of motion invariant, provided that V is an homogeneous function of degree D.

4 Hamilton's Principle

4.1 Basics of variational calculus

We are all familiar with the notion of real, or complex, function which takes a variable (or several variables) and return it (them) to a number. A function that takes a function (or several functions), say y, and return it to a number is called **functional** and it is denoted by $\mathcal{F}[y]$. It can be seen as a generalization of ordinary function of several variables to a function of infinitely many variables. An example of a functional is the integral of a continuous function (of one or several variable) over some domain. Of particular importance is finding the extremum of functionals, i.e. finding some particular function which renders $\mathcal{F}(\{y\})$ minimum or maximum. For ordinary functions that means setting its derivative to zero and solving for the variable x. To apply similar principle to $\mathcal{F}(\{y\})$, we need to introduce the concept of variation, which will be discussed in following subsection.

4.1.1 Functional derivative

We begin by considering a general function $f(x)$ that is twice differentiable. The variation of a function $f(x)$ by an amount $h(x)$ is denoted by

$$\delta f(x) = h(x) \tag{4.1}$$

This results in a variation $\delta\mathcal{F}[f]$ of the functional $\mathcal{F}[f]$, defined as

$$\delta\mathcal{F}[f] := \mathcal{F}[f + \delta f] - \mathcal{F}[f] \tag{4.2}$$

If δf is infinitesimal, then we can write

$$\delta f(x) = \epsilon\eta(x) \tag{4.3}$$

where ϵ is an infinitesimal number and $\eta(x)$ is an arbitrary function. In this case we can Taylor expand the functional $\mathcal{F}[f + \epsilon\eta]$ in terms of powers of ϵ, i.e.

$$\mathcal{F}[f + \epsilon\eta] = \mathcal{F}[f] + \frac{d\mathcal{F}[f + \epsilon\eta]}{d\epsilon}\Big|_{\epsilon=0}\epsilon + \frac{1}{2}\frac{d^2\mathcal{F}[f + \epsilon\eta]}{d\epsilon^2}\Big|_{\epsilon=0}\epsilon^2.... \tag{4.4}$$

Note that if this expansion is infinite, then $F(\epsilon)$ has to be differentiable with respect to ϵ any number of times. We define **the functional derivative**[56] of $\mathcal{F}[f]$ by

$$\boxed{\delta\mathcal{F}[f;\eta] := \frac{d\mathcal{F}[f+\epsilon\eta]}{d\epsilon}\Big|_{\epsilon=0}} \tag{4.5}$$

[56]It is also called the **variational derivative**.

The right hand side can be brought into a form linear of a linear functional of η as[57]

$$\delta\mathcal{F}[f;\eta] \equiv \int dx \frac{\delta\mathcal{F}[f]}{\delta f(x)}\eta(x) \tag{4.6}$$

or, equivalently,

$$\boxed{\left.\frac{d\mathcal{F}[f+\epsilon\eta]}{d\epsilon}\right|_{\epsilon=0} := \int dx \frac{\delta\mathcal{F}[f]}{\delta f(x)}\eta(x)} \tag{4.7}$$

where $\frac{\delta\mathcal{F}[f]}{\delta f(x)}$ is called **Frechet derivative** of $\mathcal{F}[f]$ with respect to $f(x)$.

For two functionals, $\mathcal{F}_1[f]$ and $\mathcal{F}_2[f]$, we have

$$\left.\frac{d\left(\mathcal{F}_1[f+\epsilon\eta]\mathcal{F}_2[f+\epsilon\eta]\right)}{d\epsilon}\right|_{\epsilon=0} = \left.\left\{\frac{d\mathcal{F}_1[f+\epsilon\eta]}{d\epsilon}\mathcal{F}_2[f+\epsilon\eta]\right\}\right|_{\epsilon=0} + \left.\left\{\mathcal{F}_1[f+\epsilon\eta]\frac{d\mathcal{F}_2[f+\epsilon\eta]}{d\epsilon}\right\}\right|_{\epsilon=0} \tag{4.8}$$

which according to (4.7) it follows that

$$\boxed{\frac{\delta(\mathcal{F}_1[f]\mathcal{F}_2[f])}{\delta f(x)} = \frac{\delta\mathcal{F}_1[f]}{\delta f(x)}\mathcal{F}_2[f] + \mathcal{F}_2[f]\frac{\delta\mathcal{F}_2[f]}{\delta f(x)}} \tag{4.9}$$

Similarly, one can show that the chain rule for functions also applies to functionals:

$$\boxed{\frac{\delta\mathcal{F}[\mathcal{G}[f]]}{\delta f(x)} = \int dy \frac{\delta\mathcal{F}[\mathcal{G}]}{\delta\mathcal{G}(y)}\frac{\delta\mathcal{G}[f]}{\delta f(x)}} \tag{4.10}$$

In general, we define the functional derivatives at some order n as follows:

$$\left.\frac{d^n\mathcal{F}[f+\epsilon\eta]}{d\epsilon^n}\right|_{\epsilon=0} := \int dx_1...dx_n \frac{\delta^n\mathcal{F}[f]}{\delta f(x_1)...\delta f(x_n)}\eta(x_1)...\eta(x_n) \tag{4.11}$$

so that $\mathcal{F}$ can be expressed as a Taylor expansion in terms of $\delta f(x) = \epsilon\eta(x)$ as

[57]For a function of several variables, $f(x_1, x_2, ..)$, its first total differential is

$$df = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}dx_i$$

By analogy, for a functional, $\mathcal{F}[f]$, which can be thought of as a function of an infinitely many variables, we have

$$\delta\mathcal{F}[f;\eta] = \int dx \frac{\mathcal{F}[f]}{\delta f(x)}\eta(x)$$

$$\delta \mathcal{F}[f] = \sum_{n=0}^{\infty} \frac{1}{n!} \int dx_1 ... dx_n \frac{\delta^n \mathcal{F}}{\delta f(x_1)...\delta f(x_n)} \delta f(x_1) ... \delta f(x_n) \tag{4.12}$$

If we choose the variation of the function $f(x)$ to be localized at some point y, i.e.

$$\delta f(x) = \epsilon \delta(x - y) \tag{4.13}$$

and inserting it into equation (4.12) yields

$$\frac{\delta \mathcal{F}[f]}{\delta f(y)} = lim_{\epsilon \to 0} \frac{\mathcal{F}[f(x) + \epsilon \delta(x-y)] - F[f(x)]}{\epsilon} \tag{4.14}$$

which is similar to the definition of the ordinary differentiation for functions.

Now let us consider the following functional

$$\mathcal{F}[f] = \int_{x_1}^{x_2} \delta(x - x_0) f(x)\, dx = f(x_0) \tag{4.15}$$

which associates to a function f a value of this function at a particular point $x_{\in}[x_1, x_2]$. Its variational derivative is then given by

$$\left. \frac{d\mathcal{F}}{d\epsilon} \right|_{\epsilon=0} = \frac{d}{d\epsilon} \left(f(x_0) + \epsilon \eta(x_0) \right) \Big|_{\epsilon=0} = \eta(x_0) = \int_{x_1}^{x_2} \delta(x - x_0) \eta(x)\, dx \tag{4.16}$$

and so,

$$\frac{\delta^n \mathcal{F}[f]}{\delta f(x_1)...\delta f(x_n)} = \begin{cases} \delta(x - x_0) & \text{if } n = 1 \\ 0 & \text{if } n > 1 \end{cases} \tag{4.17}$$

Using the explicit expression of $\mathcal{F}[f] = f(x_0)$, implies that

$$\frac{\delta f(x_0)}{\delta f(x)} = \delta(x - x_0) \tag{4.18}$$

We can generalize this result to the case of a functional of the form

$$\mathcal{F}[f] = \int_{x_1}^{x_2} \delta(x - x_0) f^{\alpha}(x)\, dx = f^{\alpha}(x_0) \tag{4.19}$$

Then,

$$\left. \frac{d^n \mathcal{F}(f + \epsilon \eta)}{d\epsilon} \right|_{\epsilon=0} = \alpha(\alpha - 1)...(\alpha - n + 1) \eta^n(x_0) f^{(\alpha-n)}(x_0) \tag{4.20}$$

or, equivalently

$$\left. \frac{d^n \mathcal{F}(f + \epsilon \eta)}{d\epsilon} \right|_{\epsilon=0} = \int \prod_{k=1}^{n} (\alpha - k + 1) \delta(x_k - x_0) f^{(\alpha-n)}(x_0) \eta(x_k)\, dx_k \tag{4.21}$$

which implies

$$\frac{\delta^n f^\alpha(x_0)}{\delta f(x_1)...f(x_n)} = f^{(\alpha-n)}(x_0)\prod_{k=1}^{n}(\alpha-k+1)\delta(x_k-x_0) \tag{4.22}$$

4.1.2 Euler-Lagrange equation

In mechanics, we will be interested in a class of functionals of the form

$$\mathcal{F}[y] = \int_{x_1}^{x_2} \Phi(x, y(x), y'(x))\ dx \tag{4.23}$$

where we assume that f is a smooth function of all three variables, with y' represent the derivative of y with respect to x. Then

$$\mathcal{F}[y+\epsilon\eta(x)] = \int_{x_1}^{x_2} \Phi(x, y(x)+\epsilon\eta(x), y'(x)+\epsilon\eta'(x))\ dx \tag{4.24}$$

Using the definition of the variation in (4.5), and applying Leibnitz's rule, gives

$$\delta\mathcal{F}[y;\eta] = \int_{x_1}^{x_2}\left[\frac{\partial\Phi}{\partial y}(x,y,y')\eta(x) + \frac{\partial\Phi}{\partial y'}(x,y,y')\eta'(x)\right]\ dx \tag{4.25}$$

We can get rid of the derivative η' by integrating the above integral by parts:

$$\begin{aligned}\int_{x_1}^{x_2}\frac{\partial\Phi}{\partial y'}\eta'(x)dx &= \int_{x_1}^{x_2}\frac{d}{dx}\left(\frac{\partial\Phi}{\partial y'}\eta\right)dx - \int_{x_1}^{x_2}\eta\frac{d}{dx}\left(\frac{\partial\Phi}{\partial y'}\right)dx \\ &= \left(\frac{\partial\Phi}{\partial y'}\eta\right)\Big|_{x_1}^{x_2} - \int_{x_1}^{x_2}\eta(x)\frac{d}{dx}\left(\frac{\partial\Phi}{\partial y'}\right)dx\end{aligned} \tag{4.26}$$

So, we have

$$\delta\mathcal{F}[y;\eta] = \int_{x_1}^{x_2}\left[\frac{\partial\Phi}{\partial y} - \frac{d}{dx}\left(\frac{\partial\Phi}{\partial y'}\right)\right]\eta(x)\ dx + \left(\frac{\partial\Phi}{\partial y'}\eta(x)\right)\Big|_{x_1}^{x_2} \tag{4.27}$$

In order to determine the function $y(x)$ uniquely we need to specify its boundary conditions. Here we will require that the solution satisfies the Dirichlet boundary conditions[58], i.e. $y(x_1) = c_1$ and $y(x_2) = c_2$, with $c_{1,2}$ are constants. This corresponds to requiring that $\eta(x)$ vanishes at the end points. So, according to Eq (4.6), we have

$$\frac{\delta\mathcal{F}}{\delta y(x)} = \frac{\partial\Phi}{\partial y} - \frac{d}{dx}\left(\frac{\partial\Phi}{\partial y'}\right) \tag{4.28}$$

[58] We could instead impose Neumann boundary conditions where the end points $y(x_1)$ and $y(x_2)$ are free but the derivatives at theses points are fixed.

Since the variation is arbitrary, the extremal condition $\delta\mathcal{F}[f;\eta] = 0$ amounts to having $\frac{\delta\mathcal{F}}{\delta y(x)} = 0$, which implies that the function $y(x)$ must be the solution to the equation[59]:

$$\boxed{\frac{\partial\Phi}{\partial y} - \frac{d}{dx}\left(\frac{\partial\Phi}{\partial y'}\right) = 0} \tag{4.29}$$

This equation has a similar form of the equation of motion for a system described by a Lagrangian $L(q, \dot{q}, t)$ which we derived in the previous chapter (see (3.61)). So, it is not surprising that (4.29) is known as the **Euler-Lagrange** equation.

The above result can be generalized to functionals of several functions, i.e.

$$\mathcal{F}[y_1, ...y_n] := \int_{x_1}^{x_2} \Phi\left(y_1(x), y_1'(x); ...; y_n(x), y_n'(x); x\right), \tag{4.30}$$

Writing $\delta y_i = \epsilon\eta_i(x)$, and following the same steps as before, we get

$$\delta\mathcal{F}[y_1, ...y_n;\eta] = \int_{x_1}^{x_2} \sum_i \left[\frac{\partial\Phi}{\partial y_i} - \frac{d}{dx}\left(\frac{\partial\Phi}{\partial y_i'}\right)\right] \eta_i(x)\ dx \tag{4.31}$$

Thus, the Frechet derivative with respect to y_i is given by

$$\frac{\delta\mathcal{F}[y_1, ...y_n]}{\delta y_i} = \frac{\partial\Phi}{\partial y_i} - \frac{d}{dx}\left(\frac{\partial\Phi}{\partial y_i'}\right) \tag{4.32}$$

Because the functions $\eta_i(x)$ are all independent and arbitrary, the necessary condition for $\mathcal{F}[y_1, ...y_n]$ is to be extremal is that each expression in the bracket must vanish, i.e.

$$\boxed{\frac{\partial\Phi}{\partial y_i} - \frac{d}{dx}\left(\frac{\partial\Phi}{\partial y_i'}\right) = 0, \qquad i = 1, 2, ...} \tag{4.33}$$

This method can be extended to the case of functions that depend not only on x, y and y' but also on higher derivatives of y, denoted by $y^{(n)}$, with $n > 1$:

$$\mathcal{F}[y] = \int_{x_1}^{x_2} \Phi\left(x, y(x), y'(x), ..., y^{(n)}(x)\right)\ dx \tag{4.34}$$

[59]This results from the following theorem:
"If $G(x)$ is a continuous function in $[x_1, x_2]$ and $\int_{x_1}^{x_2} G(x)h(x)\ dx = 0$ for every continuous differentiable h (x) in $[x_1, x_2]$, with $h(x_1) = h(x_2) = 0$, then

$$G(x) = 0, \quad \forall x \in [x_1, x_2]$$

Then, the functional derivative reads

$$\delta\mathcal{F}[y;\eta] = \int_{x_1}^{x_2} \left[\frac{\partial\Phi}{\partial y}\ \eta(x) + \frac{\partial\Phi}{\partial y'}\ \eta'(x) + ... \frac{\partial\Phi}{\partial y^{(n)}}\ \eta^{(n)}(x) \right]\ dx \tag{4.35}$$

As before, considering the end points fixed, i.e. $\eta(x_1) = \eta(x_2) = 0$, and assuming that all partial derivatives up to $\eta^{(n-1)}(x)$ vanish at x_1 and x_2, we obtain

$$\delta\mathcal{F}[y;\eta] = \int_{x_1}^{x_2} \left[\frac{\partial\Phi}{\partial y} - \frac{d}{dx}\left(\frac{\partial\Phi}{\partial y'}\right) + ... + (-1)^n \frac{d^n}{dx^n}\left(\frac{\partial\Phi}{\partial y^{(n)}}\right)\right]\ \eta(x)dx \tag{4.36}$$

which implies that

$$\frac{\delta\mathcal{F}[y]}{\delta y} = \frac{\partial\Phi}{\partial y} - \frac{d}{dx}\left(\frac{\partial\Phi}{\partial y'}\right) + ... + (-1)^n \frac{d^n}{dx^n}\left(\frac{\partial\Phi}{\partial y^{(n)}}\right) \tag{4.37}$$

In order for the above argument to be consistent, Φ need to be to be $(n+1)$-differentiable function and must have $2n$ derivatives. Therefore, the function $y(x)$ that renders the functional (4.34) extremal must satisfy

$$\boxed{\frac{\partial\Phi}{\partial y} - \frac{d}{dx}\left(\frac{\partial\Phi}{\partial y'}\right) + ... + (-1)^n \frac{d^n}{dx^n}\left(\frac{\partial\Phi}{\partial y^{(n)}}\right) = 0} \tag{4.38}$$

Unlike the Euler-Lagrange equations (4.29) and (4.33), there is **no existence or uniqueness theorem** in this case.

4.2 Fermat's Principle of Least Time

In a medium where the speed of light is $v(x, y)$, the time taken for it to propagate from point $A = (x_A, y_A)$ to the point $B = (x_B, y_B)$ is

$$T[y] = \int_A^B \frac{dl}{v} = \int_A^B \frac{\sqrt{dx^2 + dy^2}}{v} \tag{4.39}$$

or, equivalently,

$$cT[y] = \int_A^B n\ dl \tag{4.40}$$

with $n = c/v$ is the refraction index of the medium, which in general is a function of position. The term "cT" is called **the optical length** of the path taken by the ray of light. Writing the element of length as

$$dl = \sqrt{dx^2 + dy^2} = \sqrt{1 + y'^2}\ dx, \qquad y' = \frac{dy}{dx} \tag{4.41}$$

implies that the optical length has the form

$$cT[y] = \int_{x_A}^{x_B} \Phi(x, y, y')\, dx \tag{4.42}$$

where

$$\Phi(x, y, y') = n(x, y)\sqrt{1 + y'^2} \tag{4.43}$$

The path that light follows between two points is the one for which the optical length is minimal (i.e. that has the shortest time), also known as a **Fermat's principle**. Then, the path $y(x)$ is solution of the Euler-Lagrange equation, i.e.

$$\left[1 + y'^2\right] \frac{\partial n}{\partial y} - \left[1 + y'^2\right] y' \frac{\partial n}{\partial x} - ny'' = 0 \tag{4.44}$$

4.2.1 Propagation of Light in Homogeneous Medium

For a homogeneous medium, the index of refraction is independent of position, i.e. constant, and in this case Eq (4.44) reads

$$\frac{dy}{dx} = \text{constant} \equiv a \tag{4.45}$$

After integrating over the variable x, we obtain

$$y = ax + b \tag{4.46}$$

which is the equation of straight line. Thus, we conclude that in a homogeneous medium the path taken by a ray of light between two points is the shortest distance between them.

Next, consider two regions with index of refractions n_1 and n_2. Then in any of the these two media, the light propagates in a straight line, and the question is whether is it the same line from the start till the end or will it refract when crossing from one medium to another. If the ray starts at in region 1 a point (x_1, y_1) and ends at (x_2, y_2) in the region 2, and without loss of generality we choose the crossing point to have coordinates $(x, 0)$, then the time of travel is

$$\begin{aligned} cT[x] &= n_1 d_1 + n_2 d_2 \\ &= n_1\sqrt{(x - x_1)^2 + y_1^2} + n_2\sqrt{(x - x_2)^2 + y_2^2} \end{aligned} \tag{4.47}$$

Now the path which minimizes the time taken, corresponds to the x coordinate of the crossing point for which derivative of $T[x]$ with respect to x vanishes, i.e.

$$n_1 \frac{(x - x_1)}{\sqrt{(x - x_1)^2 + y_1^2}} + n_2 \frac{(x - x_2)}{\sqrt{(x - x_2)^2 + y_2^2}} = 0 \tag{4.48}$$

But this equation is just

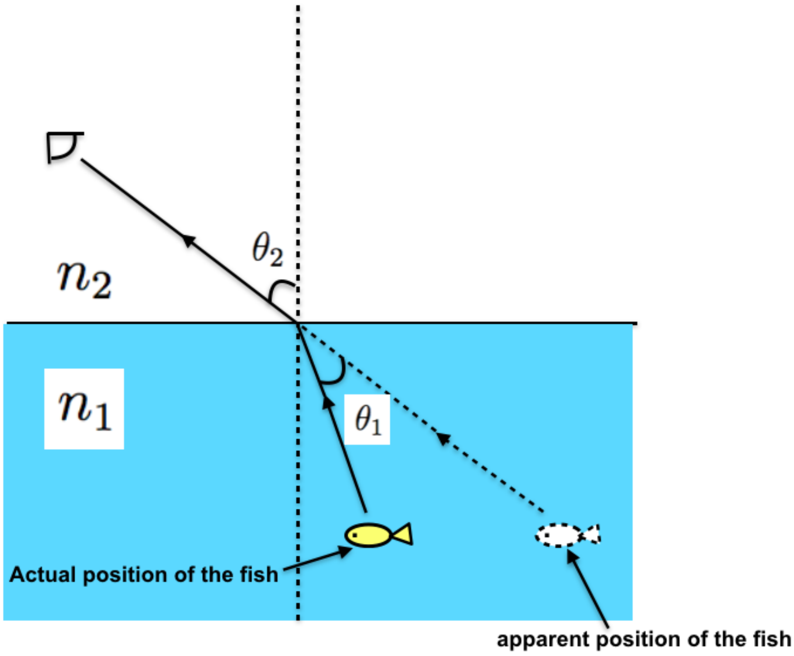


Figure 12: The refraction light when passing from one medium to another. What you see is not what get: you might be pointing to a fish in the water, but what you are looking at is an image not the fish.

$$n_1 \sin\theta_1 = n_2 \sin\theta_2 \tag{4.49}$$

with θ_1 and θ_2 are the angles that the light ray in each region makes with the normal to the surface separating the two media. This a a famous relation in geometric optics known as **Snell's law**.

4.2.2 Propagation of light in non-homogeneous medium

Let us now consider the case of non-homogeneous medium. An example of such medium is the atmosphere, where the air temperature and density change with the height, y , resulting in index of refraction depending on altitude, i.e.: $n(x, y) = n(y)$. Then, the optical length has the form

$$\Phi(y, y') = n(y)\sqrt{1 + y'^2} \tag{4.50}$$

Now we use the fact that $\Phi(y, y')$ is independent of x, and write

$$\begin{aligned} \frac{d\Phi}{dx} &:= \frac{\partial\Phi}{\partial x} + \frac{\partial\Phi}{\partial y}y' + \frac{\partial\Phi}{\partial y'} \\ &= \frac{\partial\Phi}{\partial y}y' + \frac{\partial\Phi}{\partial y'} \end{aligned} \tag{4.51}$$

Using Euler-Lagrange equation, we can re-write the last equation as

$$\begin{aligned}\frac{d\Phi}{dx} &= \frac{d}{dx}\left(\frac{\partial\Phi}{\partial y'}\right)y' + \frac{\partial\Phi}{\partial y'} \\ &= \frac{d}{dx}\left(y'\frac{\partial\Phi}{\partial y'}\right)\end{aligned} \tag{4.52}$$

Thus,

$$\frac{d}{dx}\left(\Phi - y'\frac{\partial\Phi}{\partial y'}\right) = 0 \implies \Phi - y'\frac{\partial\Phi}{\partial y'} = k \tag{4.53}$$

where k is a constant. Substituting the explicit form of Φ, we obtain

$$\frac{n(y)}{\sqrt{1+y'^2}} = k \tag{4.54}$$

The constant k can be interpreted as the value of the index of refraction at a point where the light ray becomes horizontal (i.e. $y' = 0$)[60]. Now squaring the above equation and re-rearranging it yields

$$\frac{dy}{dx} = \sqrt{\frac{n^2}{k} - 1} \tag{4.55}$$

The integration over y gives

$$\boxed{x = \int \frac{dy}{\sqrt{\frac{n^2}{k^2}-1}} + C_1} \tag{4.56}$$

where C_1 is a constant of integration to be determined from the initial conditions.

As an example, consider a medium for which the index of refraction depends linearly on the y-coordinate, i.e.

$$n(y) = n_0 + \lambda y \tag{4.57}$$

where n_0 is the index of refraction at the point where the ray starts propagating and λ is a constant. Then, Eq (4.56) reads

$$x = \frac{k}{\lambda}\int \frac{dy}{\left(y+\frac{n_0}{\lambda}\right)^2 - \frac{k^2}{\lambda^2}} + \text{constant} \tag{4.58}$$

[60]By writing $dx/dy = \tan\theta(y)$, the above equation reads

$$n(y)\sin\theta(y) = k$$

which is just the generalization of Snell's law.

Making the change of variables

$$y + \frac{n_0}{\lambda} = \frac{k}{\lambda} \cosh\theta \tag{4.59}$$

we can write the above integral as

$$x = \frac{k}{\lambda} \int \frac{\sinh\theta}{\cosh^2\theta - 1}\, d\theta + \text{const} = \frac{k}{\lambda}\theta + \text{const} \tag{4.60}$$

After expressing θ in terms of x and substituting it into Eq (4.59), we get

$$y = -\frac{n_0}{\lambda}\left[1 - \frac{k}{n_0}\cosh\left(\frac{\lambda}{k}\,(x - x_0)\right)\right] \tag{4.61}$$

Here x_0 is constant related to the integration constant above, and it is the center of symmetry of the path. For small λ, we can expand the cosh-function and keep the leading order, and we obtain

$$y \simeq -\frac{n_0 - k}{\lambda} + \frac{\lambda}{k}x^2 \tag{4.62}$$

So, the light trajectory is a parabola. There are two interesting phenomena that occur depending on the sign of λ:

(a) Refraction index decreasing with altitude ($\lambda < 0$)
In this case, the apparent height of the objects is different from the actual height. For example, if you are at the ground level looking at a tall building ($>$ 10 meters) or a mountain, will appear to you to taller than its actual height due to the bending of light caused by the increase of the refraction index as the light is coming toward your eyes.
(b) Refraction index increasing with altitude ($\lambda > 0$):
This is for example the case for the index of refraction at heights smaller than few meters from the ground. This is because the closer the air is to the ground will be, the hotter it is, hence less dense than the air above it, and therefore $n(y)$ is smaller closer to the ground. This can give rise to two phenomena:

- **Mirage**: An observer looking at an object might register two images, one them is upside which corresponds to a mirror image. For example when the light comes from the sky comes down toward the road with very small angle, it curves up that it could reach your eyes so that it looks to you that water is on the (hot) ground or desert, where as in fact it is a **mirage**.

- **Vanishing zone**: For every object there is a region, called the vanishing zone, where the observer can be such that the light can not be detected.

4.3 Brachistochrone problem

The Brachistochrone problem, which derives from from two Greek words, brachistos, meaning "shortest", and chronos, meaning "time", is one of the classical examples of applying the calculus of variation. It was posed by Johann Bernoulli to the readers of Acta Eruditorum in June 1696, who introduced it as follows:

"I, Johann Bernoulli, address the most brilliant mathematicians in the world. Nothing is more attractive to intelligent people than an honest, challenging problem, whose possible solution will l bestow fame and remain as a lasting monument. Following the example set by Pascal, Fermat, etc., I hope to gain the gratitude of the whole scientific community by placing before the finest mathematicians of our time a problem which will l test their methods and the strength of their intellect. If someone communicates to me the solution of the proposed problem, I shall publicly declare him worthy of praise". The problem he posed was the following: Given two points A and B in a vertical plane, what is the curve traced out by a point acted on only by gravity, which starts at A and reaches B in the shortest time."

Within a year five solutions were obtained, **Newton**, **Jacob Bernoulli**, **Leibniz** and **de L'Hopital** solving the problem in addition to **Johann Bernoulli**.

So, we would like to determine the curve through fixed points A and B that allows a ball rolling on the curve under the effect of gravity (without friction) to reach the final point B in the shortest possible time. The time interval from A to B reads

$$T = \int_A^B \frac{dl}{v} = \int_{x_A}^{x_B} \frac{\sqrt{1+y'^2}}{v}\, dx \tag{4.63}$$

where x and y are coordinate of a point on the curve, with the y-axis taken to be directed downward, and v is the speed of the ball and $y' = dy/dx$. We can obtain the velocity of the ball using the conservation of the total energy:

$$\frac{1}{2}mv^2 - mgy = 0 \;\Rightarrow\; v = \sqrt{2gy} \tag{4.64}$$

Here we have chosen the gravitational energy at $y = 0$ to be zero and assumed that the particle started at rest. Thus, T has the form

$$T = \int_{x_A}^{x_B} f(y, y')\, dx \tag{4.65}$$

with

$$f(y, y') = \sqrt{\frac{1+y'^2}{2gy}} \tag{4.66}$$

So, for the functional to be minimal, the $f(y, y')$ must be a solution to Euler-Lagrange equation, which, after some simplifications gives

$$2yy'' + 1 + y'^2 = 0 \tag{4.67}$$

This non-linear second order differential equation can be reduced to a first order equation as follows, if we multiply it by y' and re-arrange it to obtain

$$\frac{d}{dx}\left(y + yy'^2\right) = 0 \tag{4.68}$$

Thus,

$$y' = \sqrt{\frac{a}{y} - 1} \tag{4.69}$$

where a is a constant of integration chosen to be positive. To solve this differential equation we use the following parametrization for the coordinate y :

$$y(\phi) = a \sin^2 \theta \tag{4.70}$$

where $\theta = 0$ corresponds to the starting point A for which $x_A = y_A = 0$. Substituting in the equation above gives

$$\frac{dx}{d\phi} = a(1 - \cos 2\theta) \tag{4.71}$$

which after integrating over ϕ yields

$$x(\phi) = a(\theta - \frac{1}{2} \sin 2\theta) \tag{4.72}$$

To summarize, our results for the coordinates of a point on the curve are

$$x(\theta) = \frac{a}{2}(\theta - \sin\theta), \quad y(\theta) = \frac{a}{2}(1 - \cos\theta) \tag{4.73}$$

where $0 \leq \theta \leq 2\pi$ and the constant a is chosen such that the curve passes though the point B of coordinates (x_B, y_B)[61]. These are the parametric equations of an inverted cycloid, i.e the trajectory traced by a point on a rotating circle along a fixed axis (see figure 5). It is worth to note that the above result could be derived in a straight forward way using Beltrami identity[62].

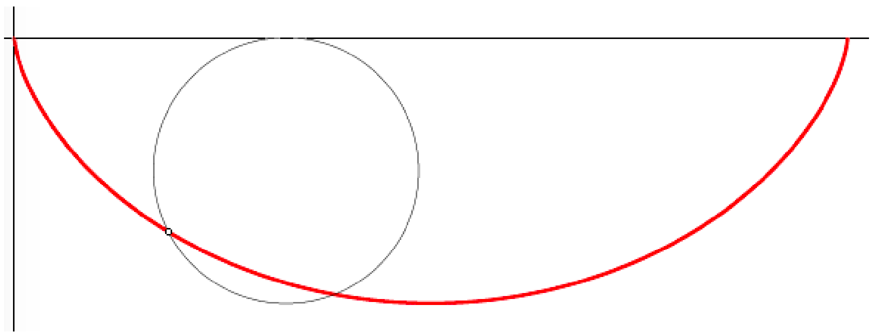


Figure 13: Cycloid.

Before closing this discussion, I would to mention one remarkable feature of this curve, and that is if you release the ball from some point on the curve from rest, the time it takes

[61]In principle, since we considered y increasing with x, the angle θ is allowed to vary only in the interval $[0 - \pi]$. However, repeating the calculation for $y' < 0$ one can check that the above solution is also valid.

[62] I should note that to solve this problem, Bernoulli did not use Euler-Lagrange equation which was discovered about 60 years later.Instead he reduced, in a very clever way, the brachistochrone problem to Snell's law in optics.

until it reaches the lowest point will be independent of the location where it was released from. To show this, suppose that the point of release is $P_0 = (x_0, y_0)$, corresponding to the angle θ_0. Then,

$$x(\theta) = x_0 + \frac{a}{2}(\phi - \sin\theta), \quad y(\theta) = y_0 + \frac{a}{2}(1 - \cos\theta) \tag{4.74}$$

So, the time it takes the ball to reach the lowest point is

$$\begin{aligned} T &= \int_{x_0}^{\pi R} \sqrt{\frac{1 + y'^2}{2g(y - y_0)}} dx = \sqrt{\frac{R}{g}} \int_{\theta_0}^{\pi} \sqrt{\frac{1 - \cos\theta}{(\cos\theta_0 - \cos\theta)}} d\theta \\ &= \sqrt{\frac{R}{g}} \int_{\theta_0}^{\pi} \frac{\sin\theta/2}{\sqrt{(\cos^2\theta_0/2 - \cos^2\theta/2)}} d\theta \end{aligned} \tag{4.75}$$

Making the change of variable $u = \cos\theta/2/\cos\theta_0/2$, we get[63]

$$T = 2\sqrt{\frac{R}{g}} \int_0^1 \frac{du}{\sqrt{(1 - u^2}} = \frac{\pi}{\sqrt{g}}\sqrt{R} \simeq 0.99\ \sqrt{R} \tag{4.76}$$

which is independent of θ_0, and is, practically, equals $\sqrt{R}$ seconds. This means that the oscillations of the ball along the cycloid are "exactly" **isochronous**, i.e. the period is independent of the oscillation amplitude (equals approximately $4\sqrt{R}$ seconds). This is unlike the oscillations of a pendulum, which as we will see in example 5.1, are only isochronous, for very small oscillation amplitudes.

4.4 Hamilton's principle

In this subsection we will introduce an alternative simple and elegant formulation of classical mechanics for obtaining Newton's equations of motion, based on variational principle that we discussed in the previous subsection. The alternative approach is called "**Hamilton's principle**" also known as "**least action principle**", formulated by some of the greatest mind of the 18 th and 19 th centuries, like **de Maupertuis**, **Euler**, **Lagrange**, **Hamilton** and **Jacobi**[64]. Although Hamilton's principle is not a new physical theory, it is one of the most profound results, not only in classical but also in the formulation of quantum Physics.

4.4.1 Least action principle for a holonomic system

Let us consider the variation of the kinetic energy of a system of N particles:

$$\begin{aligned} \delta T &= \delta\left[\frac{1}{2}\sum_{i=1}^{N} m_i \dot{\boldsymbol{r}}_i^2\right] = \sum_{i=1}^{N} m_i \dot{\boldsymbol{r}}_i.\ \delta\dot{\boldsymbol{r}}_i \\ &= \frac{d}{dt}\left[\sum_{i=1}^{N} m_i \dot{\boldsymbol{r}}_i.\ \delta\boldsymbol{r}_i\right] - \sum_{i=1}^{N} m_i \ddot{\boldsymbol{r}}_i.\ \delta\boldsymbol{r}_i \end{aligned} \tag{4.77}$$

[63]The integral over the variable u is equal to $\pi/2$. To show this we make a change of variable $u = \sin\beta$ and we get

$$\int_0^1 \frac{du}{\sqrt{(1 - u^2}} = \int_0^{\pi/2} \frac{\cos\beta}{\sqrt{1 - \sin^2\beta}} d\beta = \frac{\pi}{2}$$

[64]According to some historians, **Leibniz** first formulated this principle in a letter dated 1707

where the property $\frac{d}{dt}[\delta f(t)] = \delta[\frac{d}{dt}f]$ has been used. With the use of D'Alembert principle, i.e. $\sum_{i=1}^{N}(\boldsymbol{F}_i - m_i\ddot{\boldsymbol{r}}_i).\ \delta\boldsymbol{r}_i = 0$, we can re-write the above equation as

$$\delta T + \delta W = \frac{d}{dt}\left[\sum_{i=1}^{N} m_i\dot{\boldsymbol{r}}_i.\ \delta\boldsymbol{r}_i\right] \tag{4.78}$$

where $\delta W = \sum_{i=1}^{N} \boldsymbol{F}_i.\ \delta\boldsymbol{r}_i$ is the net work done on the system. Now, integrating this equation with respect to time between the instances t_1 and t_2, yields

$$\int_{t_1}^{t_2} (\delta T + \delta W)\, dt = \left[\sum_{i=1}^{N} m_i\dot{\boldsymbol{r}}_i.\ \delta\boldsymbol{r}_i\right] \tag{4.79}$$

If we assume that the configuration of the system is specified at the times t_1 and t_2, that is $\delta\boldsymbol{r}_i = 0$ at these times, then

$$\boxed{\int_{t_1}^{t_2} (\delta T + \delta W)\, dt = 0} \tag{4.80}$$

which is the most general form of the Hamilton's principle.
Now, let us restricts ourself to a system which is subject to a set of holonomic constraints, so that it is described by the generalized coordinates $q_1, q_2, .., q_n$, with n is the number of degrees of freedom. The variation of the net applied work can be written as

$$\delta W = \sum_{a=1}^{n} \mathcal{Q}_a \delta q_a \tag{4.81}$$

where $\mathcal{Q}_a' s$ are the applied generalized forces. If all the applied forces derive from a potential function $V(\vec{q}, t)$; i.e. $\mathcal{Q}_a = -\frac{\partial V}{\partial q_a}$, then

$$\delta W = \delta V \tag{4.82}$$

In this case, Eq (4.80) reads

$$\int_{t_1}^{t_2} \delta\,(T - V)\, dt := \int_{t_1}^{t_2} \delta L(\mathbf{q}, \dot{\mathbf{q}}, t)\ dt = 0 \tag{4.83}$$

where $\mathbf{q} \equiv (q_1, ...q_n)$ and $\dot{\mathbf{q}} = (\dot{q}_1, ...\dot{q}_n)$. For holonomic systems, the operations of integration and variation can be interchanged[65]. Then, we have

$$\delta \int_{t_1}^{t_2} L(\mathbf{q}, \dot{\mathbf{q}}, t)\ dt = 0 \tag{4.84}$$

[65] For a non-holonomic system, the varied path is not in general a possible path. Thus, in general, $\int_{t_1}^{t_2} \delta L dt \neq \delta \int_{t_1}^{t_2} L dt$ for a system with non-holonomic constraints.

where both the actual and the varied paths meet the conditions imposed by any holonomic constraint. The above equation is known as the **principle of least action**. It states that:

The actual path in configuration space followed by a holonomic system during the time interval $[t_1, t_2]$ is such that the action integral:

$$\boxed{S[q, \dot{q}, t] = \int_{t_1}^{t_2} L(\mathbf{q}, \dot{\mathbf{q}}, t)\ dt} \tag{4.85}$$

is stationary with respect to path variations which vanishes at the end points. That is:

$$\boxed{\delta S = 0} \tag{4.86}$$

Therefore, as we saw in the previous subsection, the functions $q_a(t)$ that renders $\mathcal{S}$ stationary must be a solution of the Euler Lagrange equation, i.e.

$$\boxed{\frac{\partial \mathcal{L}}{\partial q_a} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_a}\right) = 0, \qquad a = 1, ...n} \tag{4.87}$$

which is exactly the equations of motions that we derived in the previous chapter using D'Alembert principle. Thus, the dynamic of a system can be determined from the principle of least action extremizes the action functional of one single function, the Lagrangian. In fact, this principle can be used not only in classical mechanics of point like particles, but also for continuous systems such as the electromagnetic filed.

It should be noted that in the derivation above we assumed that all the forces applied to the system are conservative, i.e. they are derivable from a potential. If it is not the case, then actual path must satisfy:

$$\int_{t_1}^{t_2} \left[\delta L + \delta W^{(\text{non-conservative})}\right] dt = 0 \tag{4.88}$$

or, equivalently,

$$\int_{t_1}^{t_2} \left[\frac{\partial \mathcal{L}}{\partial q_a} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_a}\right) + Q_a^{(\text{non-conservative})}\right] \delta q_a = 0 \tag{4.89}$$

which for arbitrary variations δq_a, implies that

$$\boxed{\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_a}\right) - \frac{\partial \mathcal{L}}{\partial q_a} = Q_a^{(\text{non-conservative})}} \tag{4.90}$$

Example 4.1: **Simple pendulum**

A simple pendulum is composed of a point-like object of mass m and a massless string of constant length l (see the figure below). The coordinate of the object in the plane XY are

$$x = l \sin\theta \quad \text{and} \quad y = l \cos\theta \tag{4.91}$$

So, this system has just one degree of freedom, which is described by the angle θ. Taking the time derivative of the above coordinates, we find the components of the velocity:

$$\dot{x} = l\,\dot{\theta} \sin\theta \quad \text{and} \quad \dot{y} = -l\,\dot{\theta} \cos\theta \tag{4.92}$$

Hence, the kinetic energy of this pendulum is

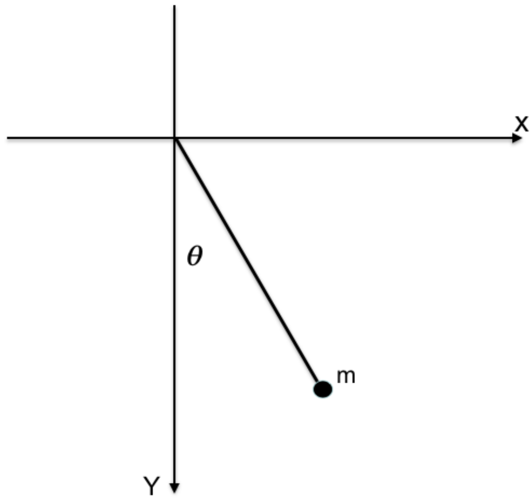


Figure 14: Simple pendulum.

$$K = \frac{m}{2}\left(\dot{x}^2 + \dot{y}^2\right) = \frac{m}{2}\, l^2\, \dot{\theta}^2 \tag{4.93}$$

If we choose the the plane $y = 0$ to be the zero potential energy, then the gravitational potential energy of this system is

$$U = -mgl\cos\theta \tag{4.94}$$

So, the Lagrangian of the simple pendulum is given by

$$U = \frac{m}{2}\, l^2\, \dot{\theta}^2 + mgl\cos\theta \tag{4.95}$$

Using Euler-Lagrange equation, the equation of motion for θ reads

$$\ddot{\theta} + \frac{g}{l}\, \sin\theta = 0 \tag{4.96}$$

For small oscillations, we can approximate $\sin\theta$ by θ, and we get

$$\ddot{\theta} + \omega_0^2 \, \theta = 0, \qquad \omega_0 = \sqrt{\frac{g}{l}} \tag{4.97}$$

which has the solution

$$\theta(t) = A \sin(\omega_0 t + \delta) \tag{4.98}$$

where A and δ are constant to be determined from the initial conditions.

Example 4.2: Spherical Pendulum

A spherical pendulum consists of a point-like particle of mass m attached to a string or a massless rod of constant length l, such that the particle is constrained to move on a sphere of radius l. The coordinate of the particle (the z-axis is pointed downward, see figure below) can expressed in terms of the polar angle θ and the azimuthal angle ϕ, as

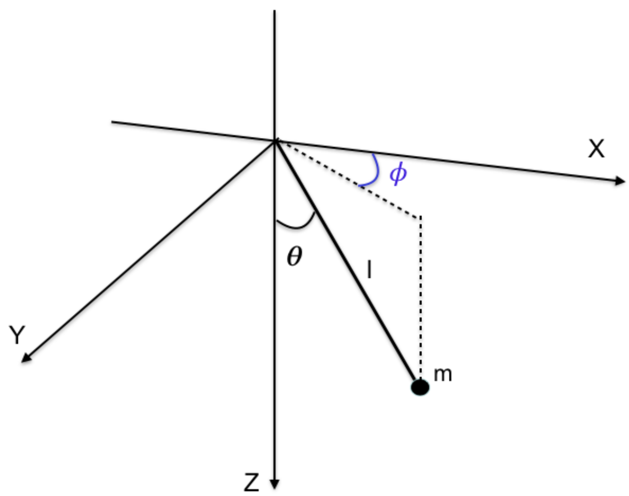


Figure 15: The spherical pendulum.

$$x = l \sin\theta \cos\phi, \quad y = l \sin\theta \sin\phi, \quad z = l \cos\theta \tag{4.99}$$

So, the components of particle's velocity are

$$\begin{aligned}
\dot{x} &= l \left[\dot{\theta} \cos\theta \cos\phi - \dot{\phi} \sin\theta \sin\phi\right], \\
\dot{y} &= l \left[\dot{\theta} \cos\theta \sin\phi - +\dot{\phi} \sin\theta \cos\phi\right], \\
z &= -l\dot{\theta} \sin\theta
\end{aligned} \tag{4.100}$$

Then, the kinetic energy of the particle is

$$\begin{aligned}
K &= \frac{m}{2}\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right) \\
&= \frac{m}{2} l^2 \dot{\theta}^2 + \frac{m}{2} l^2 \sin^2\theta \dot{\phi}^2
\end{aligned} \tag{4.101}$$

Taking the zero of the potential energy at the plan $z = 0$, the gravitational potential energy depends only on the angle θ and reads

$$U = -mgl\cos\theta \tag{4.102}$$

Thus, the Lagrangian of a spherical pendulum is given by

$$L = \frac{m}{2}l^2\dot{\theta}^2 + \frac{m}{2}l^2\sin^2\theta\dot{\phi}^2 + mgl\cos\theta \tag{4.103}$$

We see that the Lagrangian does not depend explicitly on ϕ, and so it is a cyclic coordinate. Thus, the conjugate momentum associated with ϕ is constant, which implies that $\dot{\phi} = \text{constant} = \dot{\phi}_0$, with $\dot{\phi}_0$ determined by the initial condition. Euler-Lagrange equation for θ yields

$$\ddot{\theta} = -\left(\frac{g}{l} - \dot{\phi}^2\right)\sin\theta = -\left(\frac{g}{l} - \dot{\phi}_0^2\right)\sin\theta$$

Note that if the initial condition are such that there is no motion along the azimuthal direction, then we obtain the result of the simple pendulum that we discussed in the previous example.

For non zero initial angular velocity, the motion along the polar direction is equivalent to the motion of a simple pendulum but with an effective gravitational acceleration, i.e.

$$\ddot{\theta} = -\frac{g_{\text{eff}}}{l}\,\sin\theta \tag{4.104}$$

with

$$g_{\text{eff}} = g - -l\dot{\phi}^2\cos\theta \tag{4.105}$$

Hence, considering that the pendulum is constrained to move within $0 < \theta < \pi/2$, the effective gravitational acceleration is maximal at $\theta = 0$, and vanished at $\theta = \pi/2$.

For small oscillation amplitude, the pendulum oscillates with frequency $\omega = \sqrt{\omega_0^2 - \dot{\phi}_0^2}$, with $\omega_0 = \sqrt{g/l}$ is the frequency of simple pendulum.

Example 4.3: Simple pendulum on a rotating rim

Consider a simple pendulum of length l and mass m moves on a massless rim of radius R rotating with constant angular velocity ω (see the figure below).
This system system is subject to one holonomic constraint:

$$(x - R\cos\omega t)^2 + (y - R\sin\omega t)^2 = l^2 \tag{4.106}$$

This system has only one degree of freedom, and we can chose the generalized coordinate to be the angle $\theta(t)$. So, the coordinates x and y can be expressed as

$$x = R\cos\omega t + l\sin\theta, \quad y = R\sin\omega t - l\cos\theta \tag{4.107}$$

Thus, the kinetic energy reads

$$T = \frac{1}{2}m(\dot{x}^2 + \dot{y}) = \frac{1}{2}m\left[l^2\dot{\theta}^2 + 2Rl\omega\dot{\theta}\sin(\theta - \omega t)\right] + \frac{1}{2}mR^2\omega^2 \tag{4.108}$$

By choosing the gravitational potential energy to be zero at $y = 0$, we have

$$V = -mgy = -mg\left[R\sin\omega t - l\cos\theta\right] \tag{4.109}$$

Thus, the Lagrangian of this system is given

$$L = \frac{1}{2}m\left[l^2\dot{\theta}^2 + 2Rl\omega\dot{\theta}\sin(\theta - \omega t)\right] - mg\left[R\sin\omega t - l\cos\theta\right] + \text{constant} \tag{4.110}$$

which yields the equation of motion:

$$\ddot{\theta} - \tfrac{R}{l}\omega^2\cos(\theta - \omega t) + \tfrac{g}{l}\sin(\theta - \omega t) = 0 \tag{4.111}$$

$$\ddot{\theta} - \frac{R}{l}\omega^2\cos(\theta - \omega t) + \frac{g}{l}\sin(\theta - \omega t) = 0 \tag{4.112}$$

Example 4.4: Variable mass: The Fall of a Folded U-Chain

Inn this example we will use the Lagrangian formalism to study the falling of a chain of length L and uniform mass density λ with its two ends suspended close together and at the same elevation, making a U shape. Then, one end is released free.

We choose an axis of coordinates, y-axis, and its origin, O, at the suspension point at the fixed end of the chain with the axis of coordinates oriented downward, so that $y(t)$ is the position of the free end of the falling chain (see figure below). We will also assume that the horizontal segment of the chain is very small that we can neglect it. At an instant t, the segment of the chain with fixed end has length $l_1 = (L + x)/2$, mass $m_1 = \lambda l_1$, and the falling segment has length $l_2 = (\mathrm{L} - x)/2$, mass $m_2 = \lambda l_2$, and velocity $v(t) = \dot{y}(t)$. So, the kinetic energy of the system is given by the kinetic energy of the falling segment, i.e.

$$K = \frac{1}{2}m_2\dot{y}^2 = \frac{1}{4}\lambda(\mathrm{L} - y)\,\dot{y}^2 \tag{4.113}$$

and the potential energy is purely gravitational given by

$$\begin{aligned} V(y) &= -- m_1 g\frac{l_1}{2} - m_2 g\left(y + \frac{l_2}{2}\right) \\ &= \frac{1}{4}\lambda g(\mathrm{L} - y)^2 - \frac{\lambda}{2}g\mathrm{L}^2 \end{aligned} \tag{4.114}$$

where we have chosen the horizontal plane that passed by the point O to be the zero of the gravitational potential energy. Therefore, the Lagrangian of this system reads

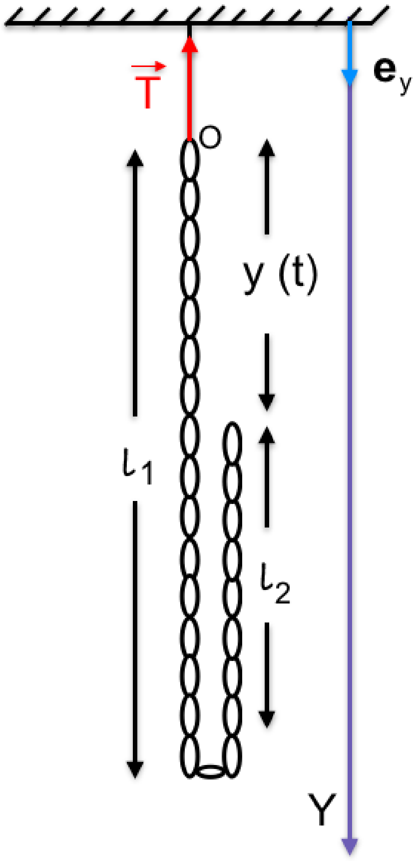


Figure 16: The falling of a folded chain.

$$\boxed{\mathcal{L} = \tfrac{\lambda}{4}\left(\mathrm{L} - y\right)\dot{y}^2 - \tfrac{1\lambda}{4}g\left(\mathrm{L} - y\right)^2 + \text{constant}} \tag{4.115}$$

from which it follows that the Euler-Lagrange equation for the distance $y(t)$ is

$$(L - y)\,\ddot{y} - \frac{1}{2}\dot{y}^2 - g\,(L - y) = 0 \tag{4.116}$$

This is a non-linear second order differential equation which is not easy to solve. However, we can use a trick to transform it into a first order differential equation by multiplying both sides by $\dot{y}$, and write

$$\frac{1}{2}\,(L - y)\,\frac{d\dot{y}^2}{dt} - \frac{1}{2}\dot{y}^3 - g\frac{d}{dt}\,(L - y)^2 = 0 \tag{4.117}$$

or, equivalently,

$$\frac{d}{dt}\left[(L - y)\,\dot{y}^2 - g\,(L - y)^2\right] = 0 \Longrightarrow \;(L - y)\,\dot{y}^2 - g\,(L - y)^2 = \text{constant} \tag{4.118}$$

A quick inspection of the above expression reveals[66] that is just the statement that the total energy of the chain is conserved. Since the chain was initially at rest, its total energy is $V(y=0) = -\lambda g L^2/4$, and so get the above equation reads

$$\boxed{(L-y)\,\dot{y}^2 - g\,(L-y)^2 = -g\mathrm{L}^2} \tag{4.119}$$

Thus, we can solve for the velocity of the falling chain, $\upsilon(y) = \dot{y}$, and obtain

$$\boxed{\upsilon(y) = \sqrt{2gy}\sqrt{\frac{1-y/2\mathrm{L}}{1-y/\mathrm{L}}}} \tag{4.120}$$

which shows that the motion of the falling segment is not a free fall. Moreover, the acceleration of the falling segment is given by

$$a(y) := \frac{d\upsilon}{dt} = \left[1 + \frac{1}{(1-y/L)^2}\right] g \tag{4.121}$$

We see that the acceleration of the free end of the chain is larger than g which indicate that in addition to its own weight the falling chain is subject to a tension directed downward.

The total time, τ, it takes the free end of the chain to fall can be obtained by integrating the inverse of the expression in Eq (4.120)of the velocity with respect the distance y from $y=0$ to $y=L$, and which after a change of variable can be written as[67]

$$\boxed{\tau = \frac{1}{2}\sqrt{\frac{2L}{g}}\int_0^1 \sqrt{\frac{1-x}{x(1-x/2)}}dx} \tag{4.122}$$

This is a finite integral which can be evaluated numerically. We find that

$$\boxed{\tau \simeq 0.85\sqrt{\frac{2L}{g}}} \tag{4.123}$$

Now let us find the tension T at the fixed end of the of the chain. For that we note that the segment fixed at the point O is acted upon by the following forces: a) the weight $m_1 g\ \mathrm{e}_y$ acting at its center of mass, b) the tension $\mathbf{T} = -\mathrm{T}\ \mathbf{e}_y$ at the point O, and c) the tension $\mathrm{T}_c = \mathrm{T}_c\ \mathrm{e}_y$ by at the lower end of the fixed segment which is the same tension that the falling segment is subjected to. Thus,

$$0 = \mathrm{T} - m_1 g - \mathrm{T}_c \tag{4.124}$$

By multiplying Eq (4.116) by $\lambda/2$, we get

$$m_2\frac{d\upsilon}{dt} = m_2 g + \frac{\lambda}{4}\upsilon^2 \tag{4.125}$$

[66] after multiplying both sides with $\lambda/4$

[67] The overall factor $\sqrt{L/g}$ is expected from dimension analysis.

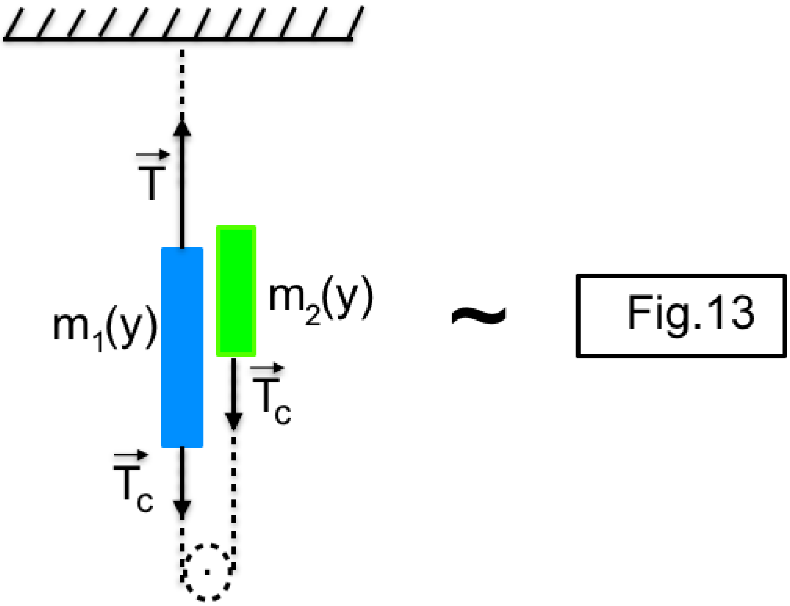


Figure 17: A simple model for the U-chain system.

We note that equations (4.124) and (4.125) have exactly the same form of the ones describing the motion of the system of two masses as in Fig. 17 shown below. This means that we can model the motion of the U-chain in Fig.16 by a system of two masses $m_1(z)$ and $m_2(z)$ that are subject to their weights and the tensions in the rope. Hence, the tension acting at the lower end of the fixed chain reads

$$\mathbf{T}_c = \frac{\lambda}{4} \upsilon^2 \mathbf{e}_y \tag{4.126}$$

Therefore, the tension at the fixed end is given by

$$\mathrm{T} = \frac{\lambda}{2} (L + y)\, g + \frac{\lambda}{4} \upsilon^2 \tag{4.127}$$

which after substituting the expression of υ yields

$$\mathrm{T} = \frac{M g}{2} \left[\frac{1 + y/L - 3y^2/L^2}{(1 - y/L)} \right] \tag{4.128}$$

where $M = \lambda L$ is the total mass of the chain.

4.4.2 Hamilton's principle for nonholonomic systems

So far we have assumed that the system is holonomic. Now, let us consider a system with n generalized coordinates and subject to a set of m nonholonomic constraint equations of the form

$$\sum_{a=1}^{n} A_{\alpha a}(q,t)\dot{q}_a + A_{\alpha t}(q,t) \tag{4.129}$$

We consider a varied path $q_a = \tilde{q}_a + \delta q_a$ around the actual path $\tilde{q}_a$, and we assume that both paths satisfy the constraint equation. We can represent the functions $A_{\alpha a}(q,t)$ by a Taylor expansion about the actual path at each instant. To first order in δq, we have

$$A_{\alpha a}(q,t) = A_{\alpha a}(q,t) + \sum_{b=1}^{n} \left(\frac{\partial A_{\alpha b}}{\partial q_b}\right)_0 \delta q_b \tag{4.130}$$

where the subscript "0" indicates that the quantity is evaluated at the actual path. By substituting the above expression of $A_{\alpha a}(q,t)$ in Eq (4.129) and using the fact that $A_{\alpha a}(q,t)$ satisfy the constraint equation, we obtain

$$\sum_{a=1}^{n} A_{\alpha a}(\tilde{q},t)\ \delta\dot{q}_a + \sum_{a,b=1}^{n} \left(\frac{\partial A_{\alpha a}}{\partial q_b}\right)_0 \dot{\tilde{q}}_a \delta q_b + \sum_{a=1}^{n} \left(\frac{\partial A_{\alpha a}}{\partial q_a}\right)_0 \delta q_a \tag{4.131}$$

Now, suppose we make the additional assumption that the $\delta q' s$ satisfy the instantaneous constraint conditions, namely

$$\sum_{a=1}^{n} A_{\alpha a}(\tilde{q},t)\ \delta q_a = 0, \qquad \alpha = 1,2,...,m \tag{4.132}$$

By differentiating it with respect to time, gives

$$\sum_{a=1}^{n} \dot{A}_{\alpha a}(\tilde{q},t)\ \delta q_a + \sum_{a=1}^{n} A_{\alpha a}(\tilde{q},t)\ \delta\dot{q}_a = 0 \tag{4.133}$$

where,

$$\dot{A}_{\alpha a}(\tilde{q},t) = \sum_{b=1}^{n} \left(\frac{\partial A_{\alpha a}}{\partial q_b}\right)_0 \dot{q}_b + \left(\frac{\partial A_{\alpha a}}{\partial t}\right)_0 \tag{4.134}$$

By subtracting Eq(4.133) from Eq(4.131), yields

$$\sum_{a,b=1}^{n} \left[\frac{\partial A_{\alpha a}}{\partial q_b} - \frac{\partial A_{\alpha b}}{\partial q_a}\right]_0 \tilde{q}_a \delta q_b + \left[\frac{\partial A_{\alpha t}}{\partial q_b} - \frac{\partial A_{\alpha b}}{\partial t}\right]_0 \delta q_b = 0 \tag{4.135}$$

If the above equation os to be valid for any set of $\delta q' s$ which conform to the constraint equations (4.132), then we must have

$$\left[\frac{\partial A_{\alpha a}}{\partial q_b} - \frac{\partial A_{\alpha b}}{\partial q_a}\right]_0 = 0, \tag{4.136}$$

$$\left[\frac{\partial A_{\alpha t}}{\partial q_b} - \frac{\partial A_{\alpha b}}{\partial t}\right]_0 = 0 \tag{4.137}$$

These equations represent the **exactness conditions** for the integrability of Eq(4.129). In other words, if the varied paths conform to the actual constraints, and the $\delta q' s$ are consistent with the instantaneous constraints, then the system must be holonomic. Whereas, for non-holonomic system, a varied path in which the $\delta q' s$ are constrained by Eq(4.131) will not be a possible path. Thus, the operation of variation and integration can be interchanged for holonomic systems, but this is not possible for the case of non-holonomic systems. This means that the Hamilton's principle, as stated in (4.86) is valid for holonomic systems only.

4.4.3 The multiplier method and the Hamilton's principle

In this subsection, we would like to apply the method of multiplier to a system that is subject to m independent constraints (holonomic or nonholonomic) of the general form:

$$\Phi_\alpha(\mathbf{q}, \dot{\mathbf{q}}, t) = 0; \qquad \alpha = 1, ...m \tag{4.138}$$

The rule of the multipliers states that the actual configuration of the system corresponds to finding the stationary value of the action:

$$\mathcal{S}_c = \int_{t_1}^{t_2} L_c(\mathbf{q}, \dot{\mathbf{q}}, \{\lambda_\alpha\}, t)dt \tag{4.139}$$

with

$$L_c(\mathbf{q}, \dot{\mathbf{q}}, \{\lambda_\alpha\}, t) = L(\mathbf{q}, \dot{\mathbf{q}}, t) + \textstyle\sum_{\alpha=1}^{m} \lambda_\alpha \Phi_\alpha(\mathbf{q}, \dot{\mathbf{q}}, t) \tag{4.140}$$

where the Lagrange multipliers $\{\lambda_\alpha, \ \alpha = 1, ...m\}$ are treated as additional variables to be determined. Demanding that $\delta\mathcal{S}_c = 0$, yields

$$\frac{\partial L_c}{\partial q_a} - \frac{d}{dt}\left(\frac{\partial L_c}{\partial \dot{q}_a}\right) = 0, \qquad a = 1, ...n \tag{4.141}$$

$$\frac{\partial L_c}{\partial \lambda_\alpha} = \Phi_\alpha(\mathbf{q}, \dot{\mathbf{q}}, t) = 0, \qquad \alpha = 1, ...m \tag{4.142}$$

Note that the $\delta\lambda_\alpha' s$ are not restricted to vanish at the end points since L_c is not an explicit function of the $\lambda' s$. Equation (4.141) can be re-written as

$$\frac{\partial L}{\partial q_a} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_a}\right) = -\sum_{\alpha=1}^{n} \dot{\lambda}_\alpha \frac{\partial \Phi_\alpha}{\partial \dot{q}_\alpha} - \sum_{\alpha=1}^{n} \lambda_\alpha G_{\alpha a} \tag{4.143}$$

where

$$G_{\alpha a} = \left[\frac{d}{dt}\left(\frac{\partial \Phi_\alpha}{\partial \dot{q}_a}\right) - \frac{\partial \Phi_\alpha}{\partial q_a}\right] \tag{4.144}$$

Now, for a system for which the constraints are linear in generalized velocities, i.e.

$$\Phi_\alpha(\mathbf{q}, \dot{\mathbf{q}}, t) = \sum_{a=1}^{n} A_{\alpha a}\dot{q}_a + A_{\alpha t} = 0 \tag{4.145}$$

and so we can write $G_{\alpha a}$ as

$$G_{\alpha a} = \left(\frac{\partial A_{\alpha a}}{\partial q_b} - \frac{\partial A_{\alpha b}}{\partial q_a}\right)\dot{q}_b + \left(\frac{\partial A_{\alpha a}}{\partial t} - \frac{\partial A_{\alpha t}}{\partial q_a}\right)\dot{q}_b = 0 \tag{4.146}$$

By comparing Eq (4.143) with the known form of the Lagrange equations for such type of constrained system, namely (see Eq (3.84) in chapter 2)

$$\frac{\partial L}{\partial q_a} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_a}\right) = \sum_{\alpha=1}^{m} \lambda_\alpha A_{\alpha a}, \qquad a = 1, ...n \tag{4.147}$$

we notice that in general the multiplier method leads to the "incorrect" dynamical equations. In order for it to reproduce (4.147), the following conditions must be satisfied

$$\frac{\partial A_{\alpha a}}{\partial q_b} - \frac{\partial A_{\alpha b}}{\partial q_a} = 0, \quad \forall \alpha \text{ and } \forall a \tag{4.148}$$
$$\frac{\partial A_{\alpha a}}{\partial t} - \frac{\partial A_{\alpha t}}{\partial q_a} = 0, \quad \forall \alpha \text{ and } \forall a$$

However, these equations are just the conditions for the integrability of the constraints, i.e. the constraints are holonomic. In such case one can **apply the Hamilton's principle** to the action $\mathcal{S}_c$. For such holonomic system, the generalized constraint force is given by

$$\mathcal{C}_a = \sum_\alpha \lambda_\alpha \frac{\partial \Phi_\alpha}{\partial q_a} \tag{4.149}$$

Example 4.5: Atwood Machine: Lagrange Multiplier Method

Consider the Atwood machine shown in Figure below. We assume that the string has constant length l, and that the pulley is massless and frictionless. There are two coordinates, y_1 and y_2, that can describe the motion of the two masses. Choosing the horizontal plane that

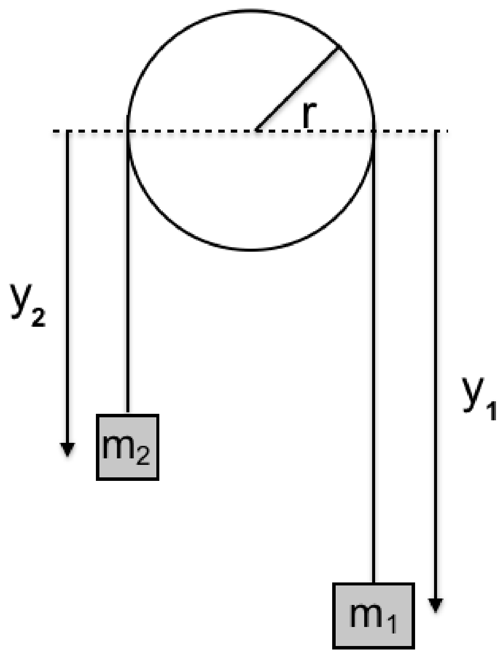


Figure 18: Atwood Machine

passes by the center of the pulley to be the zero of the potential energy, the gravitational potential energy of this system is

$$U = -m_1 g y_1 - m_2 g y_2 \tag{4.150}$$

The kinetic energy of the two masses[68]

$$T = \frac{m_1}{2}\dot{y}_1^2 + \frac{m_1 2}{2}\dot{y}_2^2 \tag{4.151}$$

However this system is subject to the holonomic constraint

$$\Phi(y_1, y_2) = y_1 + y_2 - \pi r + l \tag{4.152}$$

where r is the radius of the pulley. We can solve this system by eliminating one of the coordinates, say y_2, with the use of the above constraint, and then find the equation of motion of the other coordinates. Instead, we will use the method of Lagrange multiplier and determine the equations of motions as well the force of constraint. In this case, including the constraint in the Lagrangian via the Lagrange multiplier λ, we have

$$L = \frac{m_1}{2}\dot{y}_1^2 + \frac{m_1 2}{2}\dot{y}_2^2 + m_1 g y_1 + m_2 g y_2 + \lambda \Phi(y_1, y_2) \tag{4.153}$$

This leads to the equations of motions

$$m_1 \ddot{y}_1 = m_1 g + \lambda \tag{4.154}$$

$$m_2 \ddot{y}_2 = m_2 g + \lambda \tag{4.155}$$

Using the constraint (4.152), we have $\ddot{y}_2 = -\ddot{y}_1$, and hence the second equation reads

$$m_2 \ddot{y}_1 = -m_2 g - \lambda \tag{4.156}$$

By adding Eq (4.154) to Eq (4.156) we can eliminate the λ and find that the (components) of the accelerations of m_1 and m_2 are given by

$$\boxed{\ddot{y}_1 = -\ddot{y}_2 = \left(\frac{m_1 - m_2}{m_1 + m_2}\right) \, g} \tag{4.157}$$

The generalized forces of constraint are

$$\mathcal{C}_1 = \lambda \frac{\partial \left(y_1 + y_2 + \pi r - l\right)}{\partial y_1} = \lambda \quad \text{and} \quad \mathcal{C}_2 = \lambda \frac{\partial \left(y_1 + y_2 + \pi r - l\right)}{\partial y_2} = \lambda \tag{4.158}$$

By substituting the expression of the acceleration obtained above into equation in (4.156), we can solve for λ, and we obtain

$$\mathcal{C}_1 = \mathcal{C}_2 = \lambda = -\frac{2 m_1 m_2}{m_1 + m_2}\, g \tag{4.159}$$

[68]If the mass of the pulley is not zero, then one should include its rotational kinetic energy given by

$$K_{\text{Pulley}} = \frac{1}{2} I \dot{\phi}^2$$

where the angle ϕ describes the rotational motion of the pulley, and I is its moment of inertia with respect to the axis of rotation which perpendicular to the figure (see chapter 4).

To determine the physical significance of the generalized force of constraint above, we write Newton's equation of motion by using the free body diagram for each mass:

$$m_1\ddot{y}_1 = m_1 g - T_1 m_2\ddot{y}_1 = -m_2 g + T_2 \tag{4.160}$$

where T is the magnitude of the tension in the string. Comparing these equations with the equations (4.154) and (4.156), implies that

$$\boxed{C_1 = C_2 = -T = -\frac{2m_1m_2}{m_1+m_2}\,g} \tag{4.161}$$

Thus, the constrain force is just the tension in the string. The minus sign is due to our choice of the direction of the y-axis being directed downward.

Example 4.6: Double Atwood's Machine: Multiplier Method

The double Atwood's machine is depicted in the figure below. We assume that the pulleys are massless, the two strings that pass over the pulleys are inextensible and of lengths l_1 and l_2, respectively. From the figure, the coordinates of m_1, m_2, and m_3, are x_1, $(x_2 + y_1)$, and $(x_2 + y_2)$, respectively. So, the kinetic energy of this system reads

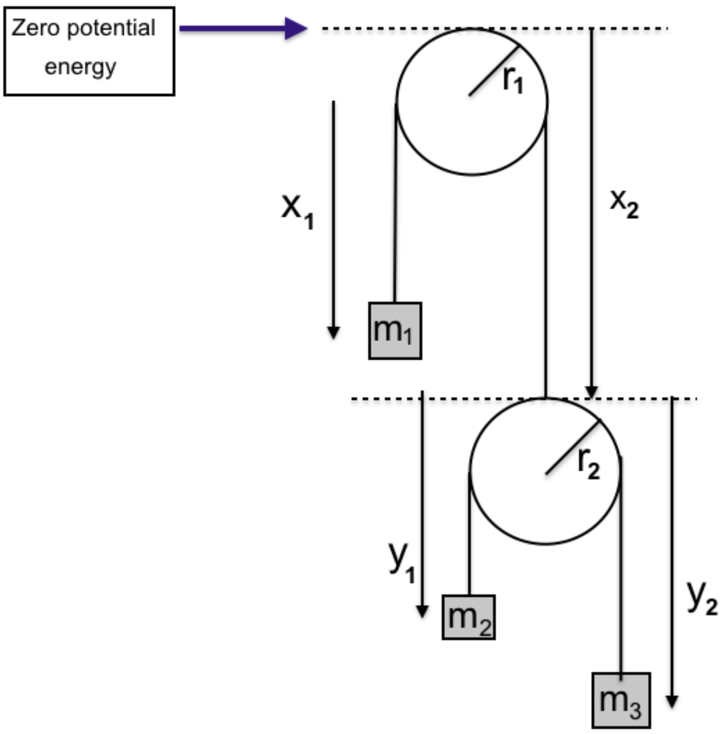


Figure 19: The double Atwood's machine.

$$K = \frac{m_1}{2}\dot{x}_1^2 + \frac{m_2}{2}(\dot{x}_2 + \dot{y}_1)^2 + \frac{m_3}{2}(\dot{x}_2 + \dot{y}_2)^2 \tag{4.162}$$

With the choice of the zero potential energy as shown in the figure, the gravitational potential energy of the system is

$$U = -m_1 g x_1 - m_2 g(x_2 + y_1) - m_3 g(x_2 + y_2) \tag{4.163}$$

However, the coordinates x_1, x_2, y_1, and y_2 are subject to the following holonomic constraints

$$\Phi_1(x_1, x_2) = x_1 + x_2 + \pi r_1 - l_1 \tag{4.164}$$
$$\Phi_1(y_1, y_2) = y_1 + y_2 + \pi r_2 - l_2 \tag{4.165}$$

Thus, using the method of Lagrange multipliers, the modified Lagrangian is given by

$$\begin{aligned} L_c = {} & \frac{m_1}{2}\dot{x}_1^2 + \frac{m_2}{2}(\dot{x}_2 + \dot{y}_1)^2 + \frac{m_3}{2}(\dot{x}_2 + \dot{y}_2)^2 \\ & + m_1 g x_1 + m_2 g(x_2 + y_1) + m_3 g(x_2 + y_2) + \lambda_1 \Phi_1 + \lambda_2 \Phi_2 \end{aligned} \tag{4.166}$$

Then, applying Euler-Lagrange equations (4.143), we obtain

$$\begin{aligned} &\text{Equation of } x_1: \quad \ddot{x}_1 = g + \frac{\lambda_1}{m_1}, \\ &\text{Equation of } y_1: \quad \ddot{x}_2 + \ddot{y}_1 = g + \frac{\lambda_2}{m_2}, \\ &\text{Equation of } y_2: \quad \ddot{x}_2 + \ddot{y}_2 = g + \frac{\lambda_2}{m_3}, \\ &\text{Equation of } x_2: \quad \lambda_1 = m_2(\ddot{x}_2 + \ddot{y}_1) + m_3(\ddot{x}_2 + \ddot{y}_2) - (m_2 + m_3)g \end{aligned} \tag{4.167}$$

From the constraint equations we deduce that $\ddot{x}_2 = -\ddot{x}_1$, and $\ddot{y}_2 = -\ddot{y}_1$, and by substituting them into the above equations we get

$$\ddot{x}_1 = g + \frac{\lambda_1}{m_1}, \tag{4.168}$$
$$-\ddot{x}_1 + \ddot{y}_1 = g + \frac{\lambda_2}{m_2}, \tag{4.169}$$
$$\ddot{x}_1 + \ddot{y}_1 = -g - \frac{\lambda_2}{m_3}, \tag{4.170}$$
$$\lambda_1 = -m_2(\ddot{x}_1 + \ddot{y}_1) - m_3(\ddot{x}_1 + \ddot{y}_1) - (m_2 + m_3)g \tag{4.171}$$
$$\tag{4.172}$$

We have four equations with four unknowns which is straight forward to solve, and we find

$$\begin{aligned}
\ddot{x}_1 &= -\ddot{x}_2 = g\left(\frac{m_1m_2 + m_1m_3}{4m_2m_3 + m_1m_2 + m_1m_3}\right) && (4.173)\\
\ddot{y}_1 &= -\ddot{y}_2 = 2g\left(\frac{m_1m_2 - m_1m_3}{4m_2m_3 + m_1m_2 + m_1m_3}\right) && (4.174)\\
\lambda_1 &= -\frac{8gm_1m_2m_3}{4m_2m_3 + m_1m_2 + m_1m_3} && (4.175)\\
\lambda_2 &= -\frac{4gm_1m_2m_3}{4m_2m_3 + m_1m_2 + m_1m_3} && (4.176)
\end{aligned}$$

Now that we found the expression of the Lagrange multipliers, the generalized forces of constraints are

$$\begin{aligned}
\mathcal{C}_1 &= \lambda_1\frac{\partial\Phi_1}{\partial x_1} + \lambda_2\frac{\partial\Phi_2}{\partial x_1} = \lambda_1 && (4.177)\\
\mathcal{C}_2 &= \lambda_1\frac{\partial\Phi_1}{\partial x_2} + \lambda_2\frac{\partial\Phi_2}{\partial x_2} = \lambda_1 && (4.178)\\
\mathcal{C}_3 &= \lambda_1\frac{\partial\Phi_1}{\partial y_1} + \lambda_2\frac{\partial\Phi_2}{\partial y_1} = \lambda_2 && (4.179)\\
\mathcal{C}_4 &= \lambda_1\frac{\partial\Phi_1}{\partial y_2} + \lambda_2\frac{\partial\Phi_2}{\partial y_2} = \lambda_2 && (4.180)
\end{aligned}$$

Thus, we have

$$\begin{aligned}
\mathcal{C}_1 &= \mathcal{C}_2 = -\frac{8gm_1m_2m_3}{4m_2m_3 + m_1m_2 + m_1m_3} && (4.181)\\
\mathcal{C}_3 &= \mathcal{C}_4 = \frac{4gm_1m_2m_3}{4m_2m_3 + m_1m_2 + m_1m_3} && (4.182)
\end{aligned}$$

To determine the relation between the constraint forces and the tensions in the strings, we need to write down Newton's second law for each mass. Using the free body diagrams, we obtain

$$\begin{aligned}
m_1\ddot{x}_1 &= m_1g - T_1 && (4.183)\\
m_2(\ddot{x}_2 + \ddot{y}_1) &= m_2g - T_2 && (4.184)\\
m_3(\ddot{x}_2 + \ddot{y}_2) &= m_3g - T_3 && (4.185)
\end{aligned}$$

where T_1, T_2, and T_3 are the tensions in the rope on m_1, m_2, and m_3, respectively. Since the pulley are massless we have $T_2 = T_3 = T_1/2$. Comparing the above equations with the expressions in (4.167), we see that

$$\begin{aligned}
\mathcal{C}_1 &= \mathcal{C}_2 = -T_1 && (4.186)\\
\mathcal{C}_3 &= \mathcal{C}_4 = -\frac{T_1}{2} && (4.187)
\end{aligned}$$

Thus, the generalized forces of constraints are just the tensions in each rope. Again the minus is just due to our choice of the axis of coordinates being directed downward.

4.5 Final comments

Before closing this chapter I would like to make a number of comments about the Lagrangian approach:

- The Lagrangian formalism uses only the kinetic and potential energy of the systems without invoking the applied forces. This makes the study of the dynamics of the system simpler than applying Newton's second law.

- The choice of Lagrangian is not unique:
  (a) Under the re-scaling of Lagrangian, i.e.

$$L \;\to\; L' = AL, \quad \forall A \in \mathbb{R} \tag{4.188}$$

  the equation of motions remain the same.

  (b) Shifting the Lagrangian by a total time derivative, i.e.

$$L \;\to\; L' = L + \frac{df(\mathbf{q}, t)}{dt} \tag{4.189}$$

  which is sometimes termed "gauge transformation ", leaves the equation of of motion invariant. This can be proved as follow. The action integral corresponding to the transformed Lagrangian reads

$$\begin{aligned} \mathcal{S}'[\mathbf{q}] &= \int_{t_1}^{t_2} L'(\mathbf{q}, \dot{\mathbf{q}}, t) = \int_{t_1}^{t_2} L(\mathbf{q}, \dot{\mathbf{q}}, t) + f(\mathbf{q}, t)\Big|_{t_1}^{t_2} \\ &= \mathcal{S}[\mathbf{q}] + f(\mathbf{q}, t)\Big|_{t_1}^{t_2} \end{aligned} \tag{4.190}$$

  So, the gauge transformation amounts to adding a constant to the action, which does vanishes when taking the variation. Therefore, the extremal of $\mathcal{S}'[\mathbf{q}]$ is also the extremal of original action.

- The Lagrange's equations hold in any system of coordinates. This follows from the fact that Hamilton's principle is about paths and not coordinates.

5 Rotating Frames

5.1 Rotating coordinate system

Consider two system of coordinates: one is fixed, which we denote by $(x'y'z')$ and we will refer to it by "Inertial", and the other is rotating, i.e. non-inertial frame, and denoted by (xyz) and we will refer to it by "Rot". Let P a point material moving in the three dimensional space. It can be represented by the vector position $\vec{r}'$ from the viewpoint of an observer in FF, and by $\vec{r}$ for an observer in the frame RF.

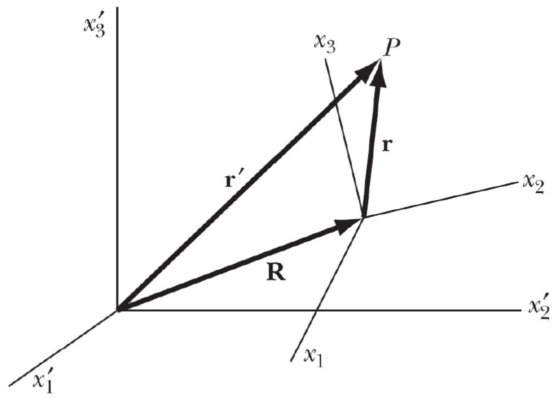


Figure 20: Two system of coordinates. One (x'y'z') is fixed and the other (xyz) is rotating.

As it is shown in the figure above, we have

$$\boldsymbol{r}' = \boldsymbol{r} + \boldsymbol{R} \tag{5.1}$$

If the point P is at rest with respect to an observer in the rotating frame RF, from the view point of the reference frame FF it is moving. If RF rotates by an infinitesimal angle $d\theta$ during a period of time dt then the position of P with respect to the fixed fame FF changes as

$$(d\boldsymbol{r})_{\text{Inertlal}} = d\boldsymbol{\theta} \times \boldsymbol{r} \;\Rightarrow\; \left(\frac{d\boldsymbol{r}}{dt}\right)_{\text{Inertial}} = \boldsymbol{\omega} \times \boldsymbol{r} \tag{5.2}$$

Here $\boldsymbol{\omega} = d\boldsymbol{\theta}/dt$ is the instantaneous angular velocity. Now if the point P is not at rest in the RF, then the above expression of velocity becomes

$$\left(\frac{d\boldsymbol{r}}{dt}\right)_{\text{Inertial}} = \left(\frac{d\boldsymbol{r}}{dt}\right)_{\text{Rot}} + \boldsymbol{\omega} \times \vec{r} \tag{5.3}$$

where $\left(\frac{d\boldsymbol{r}}{dt}\right)_{\text{Rot}}$ is the velocity of P as seen by an observer in the rotating frame.

The above relation was derived for the vector position, however it is valid for any vector quantity. To show this we consider a vector $\boldsymbol{Q}$ and let $\hat{\boldsymbol{e}}_i'$ and $\hat{\boldsymbol{e}}_i$ be the unit vectors fixed at the coordinate axis of FF and RF, respectively, then

$$\boldsymbol{Q} = \sum_i q_i\, \hat{\boldsymbol{e}}_i' = \sum_i q_i'\, \hat{\boldsymbol{e}}_i \tag{5.4}$$

where q_i and q'_i are the components of the $\vec{Q}$ as seen in the inertial and rotating frame, respectively. Taking the derivative of both sides with respect to time of the above expression, we obtain

$$\left(\frac{d\boldsymbol{Q}}{dt}\right)_{\text{Inertlal}} = \sum_i \frac{dq'_i}{dt}\, \hat{\boldsymbol{e}}_i + \sum_i q'_i\, \frac{d\hat{\boldsymbol{e}}_i}{dt} \tag{5.5}$$

Using the fact that $\hat{e}'_i$ form a basis of RF, we can expand the vector $d\hat{e}'_i$ as

$$d\hat{\boldsymbol{e}}_i = \sum_j d\theta_{ij}(t)\, \hat{\boldsymbol{e}}_j \tag{5.6}$$

The expansion parameters $d\theta_{ij}$ are just the infinitesimal angles of rotations in the plane (ij). Note that $d\theta_{ij}$ is antisymmetric under the exchange of i and j indices[69], and so we can define $d\theta_{12} = d\theta_3, d\theta_{31} = d\theta_2$, and $d\theta_{23} = d\theta_1$, which we can write in a compact form as

$$d\theta_{ij} = \sum_k \epsilon_{ijk}\, d\theta_k \;\Rightarrow\; d\hat{\boldsymbol{e}} = d\boldsymbol{\theta} \times \hat{\boldsymbol{e}} \tag{5.7}$$

with $d\boldsymbol{\theta} = d\theta\, \hat{n}$ is the infinitesimal angle of rotation about some arbitrary direction $\hat{n}$ under which the non-inertial frame is rotating. Thus, Eq (5.5) yields

$$\frac{d\hat{\boldsymbol{e}}_i}{dt} = \boldsymbol{\omega} \times \hat{\boldsymbol{e}}_i; \qquad \boldsymbol{\omega} = \frac{d\theta}{dt}\, \hat{n} \tag{5.8}$$

Therefore, that for any vector $\boldsymbol{Q}$ we have

$$\left(\frac{d\boldsymbol{Q}}{dt}\right)_{\text{Inertlal}} = \left(\frac{d\boldsymbol{Q}}{dt}\right)_{\text{Rot}} + \boldsymbol{\omega} \times \boldsymbol{Q}, \qquad \forall \boldsymbol{Q} \tag{5.9}$$

In particular, if $\boldsymbol{Q}$ represents the angular velocity vector we have

$$\left(\frac{d\boldsymbol{\omega}}{dt}\right)_{\text{Inertlal}} = \left(\frac{d\boldsymbol{\omega}}{dt}\right)_{\text{Rot}} + \boldsymbol{\omega} \times \boldsymbol{\omega} = \left(\frac{d\boldsymbol{\omega}}{dt}\right)_{\text{Rot}} \tag{5.10}$$

Thus, the angular acceleration is the same in both reference frames. Taking the derivative with respect to time on both sides of Eq(5.1) in the fixed reference frame, yields

$$\left(\frac{d\boldsymbol{r}'}{dt}\right)_{\text{Inertial}} = \left(\frac{d\boldsymbol{R}}{dt}\right)_{\text{Inertial}} + \left(\frac{d\boldsymbol{r}}{dt}\right)_{\text{Inertial}} \tag{5.11}$$

Using Eq (5.3), we obtain

$$\left(\frac{d\boldsymbol{r}'}{dt}\right)_{\text{Inertial}} = \left(\frac{d\boldsymbol{R}}{dt}\right)_{\text{Inertial}} + \left(\frac{d\boldsymbol{r}}{dt}\right)_{\text{Rot}} + \boldsymbol{\omega} \times \boldsymbol{r} \tag{5.12}$$

or, equivalently

[69] This can be seen as follows:

$$\hat{\boldsymbol{e}}_i.\hat{\boldsymbol{e}}_j = \delta_{ij} \;\Rightarrow\; 0 = d\hat{\boldsymbol{e}}_i.\hat{\boldsymbol{e}}_j + \hat{\boldsymbol{e}}_i.d\hat{\boldsymbol{e}}_j \;\Rightarrow\; d\theta_{ij} + d\theta_{ji} = 0$$

$$[\boldsymbol{v}]_{\text{Inertial}} = \boldsymbol{V} + [\boldsymbol{v}]_{\text{Rot}} + \boldsymbol{\omega} \times \boldsymbol{r} \tag{5.13}$$

where

$$\begin{aligned} [\boldsymbol{v}]_{\text{Inertlal}} &: \text{ velocity of P in the fixed frame} \\ \boldsymbol{V} &: \text{ velocity of the origin of RF w.r.t the fixed frame} \\ [\boldsymbol{v}]_{\text{Rot}} &: \text{ velocity of P in the rotating frame} \\ \boldsymbol{\omega} \times \vec{r} &: \text{ velocity of P due to the rotating axes} \end{aligned} \tag{5.14}$$

5.2 Newton's Law in the rotating frame

As we discussed at the beginning of the first chapter, Newton's second law is applicable only in the inertial frames, so that we can write

$$[\boldsymbol{a}]_{\text{Inertlal}} := \frac{\boldsymbol{F}}{m} \tag{5.15}$$

where $\vec{F}$ is the external force acting on P in the fixed frame. We can rewrite the left hand side of the above equation as

$$\begin{aligned} [\boldsymbol{a}]_{\text{Inertial}} &= \left[\frac{d\boldsymbol{V}}{dt}\right]_{\text{inertial}} + \left[\frac{d\boldsymbol{v}}{dt}\right]_{\text{Inertial}} + \left[\frac{d\boldsymbol{\omega}}{dt}\right]_{\text{Inertial}} \times \boldsymbol{r} + \ \boldsymbol{\omega} \times \left[\frac{d\boldsymbol{r}}{dt}\right]_{\text{Inertial}} \\ &= \left[\frac{d\boldsymbol{V}}{dt}\right]_{\text{Rot}} + \left[\frac{d\boldsymbol{v}}{dt}\right]_{\text{Rot}} + 2\boldsymbol{\omega} \times [\boldsymbol{v}]_{\text{Rot}} + \dot{\boldsymbol{\omega}} \times \boldsymbol{r} + \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{r}) \end{aligned} \tag{5.16}$$

where we made use of Eq(5.9) to obtain the second equation above. Since $\left[\frac{d\boldsymbol{v}}{dt}\right]_{\text{Rot}}$ is the acceleration of P in the rotating frame, we can write

$$[\boldsymbol{a}]_{\text{Rot}} = [\boldsymbol{a}]_{\text{Inertial}} - \left[\frac{d\boldsymbol{V}}{dt}\right]_{\text{Inertial}} - 2\boldsymbol{\omega} \times [\boldsymbol{v}]_{\text{Rot}} - \dot{\boldsymbol{\omega}} \times \boldsymbol{r} - \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{r}) \tag{5.17}$$

Note that if the reference frame (xyz) is not rotating and has no linear acceleration with respect to the fixed frame, i.e. $\boldsymbol{\omega} = 0$ and $\boldsymbol{V} = \text{const.}$, then the two accelerations are equal, which is expected since in this case we will be dealing with inertial frame. Multiplying both sides of the equation by the mass m of the point particle P, we obtain the effective force that an observer in the rotating frame would determine, i.e.

$$\boldsymbol{F}_{\text{eff}} = m\ [\boldsymbol{a}]_{\text{Inertial}} - m\left[\frac{d\boldsymbol{V}}{dt}\right]_{\text{Inertial}} + \boldsymbol{F}_{\text{Euler}} + \boldsymbol{F}_{\text{Coriolis}} + \boldsymbol{F}_{\text{Centrifugal}} \tag{5.18}$$

where

- $\underline{\boldsymbol{F}_{\text{Euler}} = -m\ \dot{\boldsymbol{\omega}} \times \boldsymbol{r}}$ $\equiv$ **Euler force**

  It is perpendicular to the position vector $\vec{r}$ and its effect is to resist the direction of motion. It is present only for the case where ω depends on time.

- $\underline{\boldsymbol{F}_{\text{Coriolis}} = -2m\,\boldsymbol{\omega} \times [\boldsymbol{v}]_{\text{Rot}}} \equiv$ **Coriolis force**

  It only applies to a particle that is moving in a rotating reference frame and is greatest if the velocity is perpendicular to the axis of rotation. It has the effect of moving the particle farther or closer to the axis depending if it is moving counter-clockwise or clockwise around the direction of the axis of rotation.

- $\underline{\boldsymbol{F}_{\text{Centrifugal}} = -m\,\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{r})} \equiv$ **Centrifugal force**

  Its direction is always away from the axis of rotation.

Example 4.1: Effect of centrifugal force - and apparent gravity

The Earth is an example of a non-inertial frame since it is rotating around itself and also orbiting the sun. However, we usually assume it is inertial frame. Is that justifiable? For that we will estimate the effect of the Earth rotation on the acceleration of gravity as measured in an inertial frame. First, of all since Earth rotates at constant angular velocity, the Euler acceleration vanishes, and so for a particle of mass m at or above the surface of the Earth subject to a force $\vec{F}$ in addition to its weight we have[70]

$$m\left(\frac{d^2\boldsymbol{r}}{dt^2}\right)_{\text{Rot}} = \mathbf{F}_{\text{other}} + m\mathbf{g} - 2m\boldsymbol{\omega} \times \boldsymbol{v}_{\text{Rot}} - m\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{r}) \tag{5.20}$$

where v_{Rot} is the velocity of the particle in the Earth frame, and $\mathbf{F}_{\text{other}}$. all forces that act on m other than the force of gravity. Since we are interested in the effect of the centrifugal force, let us consider the mass m to be suspended above a general point on the Earth's surface at latitude ψ and in equilibrium with respect to an observer on that location. Hence, the Coriolis force vanishes since the particle is static with respect to an observer on the Earth's surface, and $\mathbf{F}_{\text{other}} = \boldsymbol{T}$ is the tension of the string. In this case the above equation can be written as

$$\boldsymbol{T} + m\,\mathbf{g}_{\text{eff}} = 0 \tag{5.21}$$

where $\mathbf{g}_{\text{eff}}$is the effective, or apparent, acceleration of gravity given by

$$\begin{aligned} \mathbf{g}_{eff} &= -g_0\,\hat{\boldsymbol{r}} - \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{r}) = 0 \\ &= -\left(g_0 - R_\oplus\,\omega_\oplus^2\,\cos^2\psi\right)\hat{\mathbf{r}} - R_\oplus\,\omega_\oplus^2\,\sin\psi\cos\psi\,\hat{\boldsymbol{\psi}} \end{aligned} \tag{5.22}$$

where $\hat{\boldsymbol{\psi}}$ is a unit vector directed north, $g_0 = GM_\oplus/R_\oplus^2 = 9.806$ m$/s^2$, with $M_\oplus$ the mass of the Earth, $R_\oplus \simeq 6380$ km iits radius, and $\omega_\oplus = 1/$day $\simeq 7.3 \times 10^{-5}$ rad$/s$ its angular frequency. Thus, at the equator we get

$$\mathbf{g}_{eff} = -\left(g_0 - R_\oplus\,\omega_\oplus^2\right)\hat{\mathbf{r}} = \left(g_0 - 3.2 \times 10^{-2}\ \text{m}/s^2\right)\hat{\boldsymbol{r}} \tag{5.23}$$

[70]We neglect the gravitational pull of the Sun on the the mass m, $F_\odot$, as compared to the gravitational force of gravity since

$$\frac{F_\odot}{F_\oplus}\frac{M_\odot}{M_\oplus}\left(\frac{R_\oplus}{R_{\text{E-S}}}\right)^2 \simeq 6 \times 10^{-4} \tag{5.19}$$

Thus, the acceleration of gravity at the equator is about 0.3% less than at the poles. In fact, the observed difference is 0.52%. The discrepancy with the prediction above is due to the fact that the Earth is not exactly spherical; it is oblate with greater radius at the equator which is caused by the centrifugal force of its rotation.

Note that $\mathbf{g}_{eff}$ has non zero component along $\hat{\boldsymbol{\psi}}$ except at the equator and the poles. Hence, in general a plumb bob will not point exactly toward the Earth's center. The deflection angle θ that the string makes with the vertical can be found by solving Eq (5.21) along the $\hat{\mathbf{r}}$ and $\hat{\boldsymbol{\psi}}$, i.e.

$$T\cos\theta = -\left(g_0 - R_\oplus\, \omega_\oplus^2\, \cos^2\psi\right) \tag{5.24}$$
$$T\sin\theta = R_\oplus\, \omega_\oplus^2\, \sin\psi\cos\psi \tag{5.25}$$

from which it follows

$$\tan\theta = \frac{R_\oplus\, \omega_\oplus^2\, \sin\psi\cos\psi}{\left(g_0 - R_\oplus\, \omega_\oplus^2\, \cos^2\psi\right)} \tag{5.26}$$

Since $R_\oplus\, \omega_\oplus^2 \ll g_0$, we can approximate the above expression as

$$\theta \simeq \frac{R_\oplus\, \omega_\oplus^2\, \sin 2\psi}{2g_0} \simeq 1.6\times 10^{-3}\text{rad}\left(\frac{\sin 2\psi}{1}\right) \tag{5.27}$$

Thus, the maximum deflection is about 0.1° which occurs at latitude $\psi = \pi/4$. Note that the direction of the deflection is southward in the northern hemisphere (i.e. $\psi > 0$) and northward in the southern hemisphere (i.e. $\psi < 0$).

Example 4.2: Effect of Coriolis force -Free-Fall revisited

To illustrate the effect of the Coriolis force in a rotating frame, consider a body mass m initially at rest dropped from some height h above the surface of the Earth. Since the angular speed of Earth's rotation is small, we neglect terms of order ω^2 as compared to terms of lower order. So, the acceleration of a falling mass reads

$$\frac{d^2\boldsymbol{r}}{dt^2} = \mathbf{g} - 2\boldsymbol{\omega}\times\frac{d\boldsymbol{r}}{dt} \tag{5.28}$$

where $\boldsymbol{r}$ is the vector position of the mass in a reference frame on the Earth's surface with its origin at the point below the starting point of the body. Integrating the above equation with respect to time, and using the initial condition $\dot{bmr}(0) = 0$, we find

$$\frac{d\boldsymbol{r}}{dt} = \boldsymbol{g}\ t - 2\boldsymbol{\omega}\times(\boldsymbol{r} - \boldsymbol{r}(0)) \tag{5.29}$$

We substitute the above expression of the velocity in (5.28) and get

$$\frac{d^2\boldsymbol{r}}{dt^2} \simeq \boldsymbol{g} - 2\boldsymbol{\omega}\times\boldsymbol{g}\ t \tag{5.30}$$

where a term of order ω^2 has been neglected which is consistent with the approximation we made above. We now easily integrate (5.30) and obtain

$$\boldsymbol{r} = \boldsymbol{r}(0) + \frac{1}{2}\boldsymbol{g}\ t^2 - \frac{1}{3}\boldsymbol{\omega} \times \boldsymbol{g}\ t^3 \quad (+\text{terms of order } \omega^2) \tag{5.31}$$

If we chose the x-axis pointing south, y-axis pointing east and z-axis pointing vertically up and consider the mass m at a latitude ψ North, then Eq (5.32) can be written as

$$\boldsymbol{r}(t) = \frac{1}{3}\omega\ g\ t^3\ \hat{\boldsymbol{e}}_y + \left(h - -\frac{1}{2}g\ t^2\right)\ \hat{\boldsymbol{e}}_z \quad (+\text{terms of order } \omega^2) \tag{5.32}$$

or, equivalently, in components form as

$$\begin{aligned} \ddot{x} &= 0 \\ \ddot{y} &= 2gt\omega\cos\psi \\ \ddot{z} &= -g \end{aligned} \tag{5.33}$$

which upon integration (and the use of initial condition) give

$$\begin{aligned} \Delta x &= 0 \\ \Delta y &= \frac{1}{3}gt^3\omega\cos\psi \\ \Delta z &= h - \frac{1}{2}gt^2 \end{aligned} \tag{5.34}$$

The object reaches the surface a

$$t_s = \sqrt{\frac{2h}{g}} \tag{5.35}$$

and the this instant the lateral displacement of the object is

$$\Delta y = \frac{1}{3}g\left(\frac{2h}{g}\right)^{3/2}\omega\ \cos\psi \tag{5.36}$$

For example, for $h = 100\ m$ and $\psi = 0$, i.e. at the equator, we get

$$\Delta y \simeq 2.3 \text{ cm} \tag{5.37}$$

Although such displacement is very small that it can be neglected in most practical applications, it is measurable[71].

[71] If the experiment is performed on the Eiffel tower where $h = 275$ m and $\psi \simeq 49°$, we get a lateral displacement toward the east by

$$\Delta y^{(\text{Eiffel Tower})} = 6.5 \text{ cm}$$

Example 4.3: Effect of centrifugal force-Rotating bucket of fluid

Consider a bucket of a cylindrical shape filled with some fluid in a steady state, rotating with constant angular frequency around its symmetry axis which we take to be the z-axis. We assume that the fluid is stationary in the reference frame, and so for an infinitesimal element of fluid of mass δm, we have

$$\delta \boldsymbol{F}_p + \delta m \boldsymbol{g} - \delta m \omega^2 \hat{\boldsymbol{z}} \times (\hat{\boldsymbol{z}} \times \boldsymbol{r}) = 0 \tag{5.38}$$

where $\delta \boldsymbol{F}_b$ is the buoyancy force that acts on on the element of the fluid surface of mass δm and it is perpendicular to it[72]. Dividing the equation above by δm, and defining $\hat{\boldsymbol{n}}$ as the unit vector normal to the fluid surface element at the point $\boldsymbol{r} = \rho\hat{\boldsymbol{\rho}} + z(\rho)\hat{\boldsymbol{z}}$ defined by its cylindrical coordinates (ρ, z), we obtain[73]

$$\left|\frac{\delta F_b}{\delta m}\right| \ \hat{\boldsymbol{n}} = g \ \hat{\boldsymbol{z}} - \omega^2 \rho \ \hat{\boldsymbol{\rho}} \tag{5.39}$$

Since the infinitesimal vector position $d\boldsymbol{r}(P)$ is tangent to the surface of the fluid at point P, we have[74]

$$d\boldsymbol{r}(P).\hat{\boldsymbol{n}} = 0 \implies \left(d\rho \ \hat{\boldsymbol{\rho}} + \frac{dz}{d\rho} d\rho \ \hat{\boldsymbol{z}} + \rho d\phi \ \hat{\boldsymbol{\phi}}\right) . \left(g \ \hat{\boldsymbol{z}} - \omega^2 \rho \ \hat{\boldsymbol{\rho}}\right) = 0 \tag{5.40}$$

from which it follows that

[72] If $\delta \boldsymbol{F}_p$ has a component tangent to the surface of the fluid, it will cause it to flow along that direction which will be in contradiction with our assumption that the fluid is in steady state.

[73] The buoyancy force per unit mass of the fluid is just the gradient of the pressure over the mass density of the fluid, i.e.

$$\frac{\delta \boldsymbol{F}_b}{\delta m} = -\frac{\boldsymbol{\nabla} p}{\mathcal{D}}$$

where p is the pressure, and $\mathcal{D}$ denotes the mass density of the fluid (so not be confused with the cylindrical coordinate ρ used above). Since the gradient of a function $f(\boldsymbol{r})$ at arbitrary point $P = (\boldsymbol{r}_0)$, i.e.$\boldsymbol{\nabla} f|_P$, is perpendicular to the tangent at that point on the surface $f(\boldsymbol{r})$ = constant (for the case of fluid, the function $f(\boldsymbol{r})$ corresponds to the pressure)

$$\frac{\delta \boldsymbol{F}_b}{\delta m} = \left|\frac{\boldsymbol{\nabla} p}{\mathcal{D}}\right| \hat{\boldsymbol{n}}$$

[74] In obtaining the second the third terms in the first bracket of Eq (5.40) we used where we used the fact that

$$\begin{aligned} d\hat{\boldsymbol{\rho}} &= \hat{\boldsymbol{\rho}} \times d\phi\hat{\boldsymbol{z}} = d\phi \ \hat{\boldsymbol{\phi}} \\ d\hat{\boldsymbol{z}} &= \hat{\boldsymbol{z}} \times d\phi\hat{\boldsymbol{z}} = 0 \end{aligned}$$

$$g\ dz(\rho) = \omega^2 \rho\ d\rho \tag{5.41}$$

The integration of the above equation yields

$$z(\rho) = \frac{\omega^2 \rho^2}{2g} + \text{constant} \tag{5.42}$$

which implies that the shape of the surface of the rotating fluid is a paraboloid.

Example 4.4: Effect of Coriolis force - Foucault's Pendulum

In this example we will study the effect of the rotation of the Earth on a simple pendulum. Or said differently, how do we know that is rotating around its axis?[75]. We will neglect air resistance so that the pendulum can swing freely in any direction. Due to the rotation of the Earth, the plan of oscillations will rotate with respect to the surface beneath it[76]. For the inertial frame, there are just two forces on the bob, the tension T in the string and the force of gravity $m\vec{g}$. whereas in the rotating frame of the earth, there are also the centrifugal and Coriolis forces, so the equation of motion in the earth's frame is

$$\begin{aligned} m\ddot{\boldsymbol{r}} &= m\boldsymbol{g} + \boldsymbol{T} - 2m\ \boldsymbol{\omega} \times \boldsymbol{r} \qquad (5.43) \\ &= -mg\hat{\boldsymbol{z}} + T\left(\hat{\boldsymbol{z}}\ \cos\theta - \hat{\boldsymbol{x}}\ \sin\theta\cos\phi - \hat{\boldsymbol{y}}\ \sin\theta\sin\phi\right) - 2m\ \boldsymbol{\omega} \times \boldsymbol{r} \end{aligned}$$

Here ϕ is the azimuthal angle and θ is the angle between the $-\hat{z}$ and the vector from from the point of suspension of the pendulum to the mass m, and the angular frequency of the Earth $\boldsymbol{\omega}$ can be written in components as

$$\boldsymbol{\omega} = -\omega\cos\psi\ \hat{\boldsymbol{x}} + \omega\sin\psi\ \hat{\boldsymbol{z}} \tag{5.44}$$

where ψ is the latitude of the location on the Earth's surface where the pendulum is hanged. If we make the approximation that the oscillations are with small angles, then we can consider $\theta \simeq 0$ and os the tension on the string cancels the weight. Now, in terms of components, Eq (5.43) reads

$$\ddot{x} - 2\omega_z\dot{y} + \omega_0^2 x = 0 \tag{5.45}$$

$$\ddot{y} - 2\omega_z\dot{x} + \omega_0^2 y = 0 \tag{5.46}$$

where $\omega_z = \omega\sin\psi$, and $\omega_0^2 = g/L$ is the natural frequency of simple pendulum. By multiplying Eq (5.46) by i and add it to Eq (5.45) the two equations combine into one equation in terms of the complex variable $\eta = x + iy$ and we obtain:

$$\ddot{\eta} + 2i\omega_z\dot{\eta} + \omega_0^2 = 0 \tag{5.47}$$

[75] On March of 1851, at the Pantheon of Paris, Leon Foucault (1819 - 1868) used a pendulum of length $l = 67$ m, with mass of 28kg at its end, to demonstrate the rotation of the Earth.

[76] In reality, to suppress the effect of air resistance one consider a very long and heavy pendulum.

To find the solution to this is a second order linear-homogeneous differential equation, we try the following form for η:

$$\eta = e^{-i\alpha} \tag{5.48}$$

Substituting into Eq (5.47) yields

$$\alpha = \omega_z \pm \sqrt{\omega_z^2 + \omega_0^2} \tag{5.49}$$

As the natural frequency is much larger than angular frequency of the rotation of the Earth, we can expand the square root in powers of ω_z/ω_0 and keep only the leading order term. We find

$$\alpha \simeq \omega_z \pm \omega_0 \tag{5.50}$$

Hence,

$$\eta = e^{-i\omega_z t}\left[A_+ e^{i\omega_0 t} + A_- e^{-i\omega_0 t}\right] \tag{5.51}$$

where A_+ and A_- are constants to be determined from the initial conditions. For instance, if we consider $x(0) = x_0$, $y(0) = 0$, $\dot{x}(0) = 0$, and $\dot{y}(0) = 0$, we get

$$\begin{aligned} x(t) &= \frac{x_0}{\omega_0}\left(\omega_z \sin\omega_z t \sin\omega_0 t + \omega_0 \cos\omega_z t \cos\omega_0 t\right) \simeq x_0 \cos\omega_z t \cos\omega_0 t \\ y(t) &= \frac{x_0}{\omega_0}\left(\omega_z \cos\omega_z t \sin\omega_0 t - \omega_0 \sin\omega_z t \cos\omega_0 t\right) \simeq -x_0 \sin\omega_z t \cos\omega_0 t \end{aligned} \tag{5.52}$$

The above solution shows that due to the Coriolis force the plane of oscillation of Foucault's pendulum precess at constant angular velocity $\Omega = -\omega_z = -\omega\ \sin\psi$ with respect to the Earth, and hence it makes a rotation of 360° in a time period

$$T = \frac{2\pi}{\omega|\sin\psi|} = \frac{T_\oplus}{|\sin\psi|} \tag{5.53}$$

where $T_\oplus$ is the sidereal period of the Earth, i.e. with respect to fixed stars.

For Foucault's pendulum at the museum the Pantheon in Paris, the period is

$$T(\text{Paris}) = \frac{23\text{h}\ 56'\ 4''}{\sin 49^\circ} \simeq 31\text{h}\ 48' \tag{5.54}$$

6 Rotating Frames and Rigid Body

By definition a rigid body is a system of N point particles in which the distance between them does not change, i.e.

$$|\boldsymbol{r}_i - \boldsymbol{r}_j| = \text{constant}, \quad \forall\, i \neq j = 1, 2..., N \tag{6.1}$$

In the limit where $N \to \infty$, we have a continuous rigid body. In this case, the object is characterized by a mass density, $\rho(\boldsymbol{r})$, so that its total mass is given by

$$M = \int \rho(\boldsymbol{r})\; d^3\boldsymbol{r} \tag{6.2}$$

In general, the motion of a rigid body can be decomposed into a translational motion of the center of mass as seen by an observer in some inertial frame which we can take fixed[77], and a rotational motion around an axis passing through the center of mass. Below we will describe the kinematics as seen by a rotating observer and how it is related to the ones in an inertial reference frame.

6.1 The inertia tensor

Consider a rigid body made of a collection of N particles with masses $\{m_\alpha : \alpha = 1, 2, ...N\}$, located at the positions $\vec{r}_\alpha$ with respect to a reference frame attached to the body. If this system rotates at angular velocity $\vec{\omega}$ around a fixed axis, and moves at velocity $\vec{V}$ with respect to a fixed inertial reference frame, then the kinetic energy of the α^{th} particle is given by

$$\begin{aligned} T_\alpha &= \frac{1}{2} m_\alpha v_\alpha^2 = \frac{1}{2} m_\alpha \left(\boldsymbol{V} + \boldsymbol{\omega} \times \boldsymbol{r}_\alpha\right)^2 \\ &= \frac{1}{2} m_\alpha V^2 + m_\alpha \boldsymbol{V} . \left(\boldsymbol{\omega} \times \boldsymbol{r}_\alpha\right) + \frac{1}{2} m_\alpha \left(\boldsymbol{\omega} \times \boldsymbol{r}_\alpha\right)^2 \end{aligned} \tag{6.3}$$

Thus, the total kinetic energy of the whole system reads

$$T = \frac{1}{2} M V^2 + \boldsymbol{V} . \left[\boldsymbol{\omega} \times \sum_\alpha m_\alpha \boldsymbol{r}_\alpha\right] + \frac{1}{2} \sum_\alpha m_\alpha \left(\boldsymbol{\omega} \times \boldsymbol{r}_\alpha\right)^2 \tag{6.4}$$

where $M = \sum_\alpha m_\alpha$ is the mass of the rigid body. Let us now chose the origin of the frame attached to the body to be its center of mass (CM). In this case we have

$$\boldsymbol{R}_{\text{CM}} := \frac{\sum_\alpha m_\alpha \boldsymbol{r}_\alpha}{M} = 0 \tag{6.5}$$

Then, the kinetic energy of the rigid body is given by

[77] We can always make such a choice since the laws of motion remain unchanged in any inertial reference frame.

$$T = \frac{1}{2}MV^2 + \frac{1}{2}\sum_\alpha m_\alpha \left(\boldsymbol{\omega} \times \boldsymbol{r}_\alpha\right)^2 \tag{6.6}$$

The first term on the RHS of (6.6) is the translation kinetic energy of the body's CM, and the second term represents the rotational kinetic energy about the CM. Now using the vector identity[78]

$$(\boldsymbol{A} \times \boldsymbol{B})^2 = A^2B^2 - (\boldsymbol{A}.\boldsymbol{B})^2 \tag{6.7}$$

we can write the rotational kinetic energy of the body as

$$\begin{aligned} T &= \frac{1}{2}\sum_\alpha m_\alpha \left[\omega^2 r_\alpha^2 - (\boldsymbol{\omega}.\boldsymbol{r}_\alpha)^2\right] \\ &= \frac{1}{2}\sum_\alpha m_\alpha \Big\{ \sum_{i,j} \omega_i \delta_{ij} \omega_j r_\alpha^2 - \sum_{i,j} \omega_i x_{\alpha,i} x_{\alpha,j} \omega_j \Big\} \\ &= \frac{1}{2}\sum_{i,j} \omega_i \, \mathcal{I}_{ij} \, \omega_j \end{aligned} \tag{6.8}$$

where

$$\mathcal{I}_{ij} = \textstyle\sum_\alpha m_\alpha \ \left[\delta_{ij} \ r_\alpha^2 - x_{\alpha,i} x_{\alpha,j}\right] \tag{6.9}$$

The elements of $\{\mathcal{I}_{ij} : i, j = 1, 2, 3\}$, are the components of the **tensor of inertia** measured in the body frame of reference. Note that from its definition given above we see that $\mathcal{I}_{ij} = \mathcal{I}_{ji}$, and so the tensor of inertia is symmetric.

For a continuous distribution of mass of density ρ, the expression (6.9) becomes

$$\mathcal{I}_{ij} = \textstyle\int_\mathcal{V} \rho(\vec{r}) \ \left[\delta_{ij} \ r^2 - x_i x_j\right] \tag{6.10}$$

or more explicitly, we have

$$\mathcal{I}_{ij} = \int_\mathcal{V} \rho(\vec{r}) \begin{bmatrix} y^2 + z^2 & -xy & -xz \\ -xy & x^2 + z^2 & -yz \\ -xz & -yz & x^2 + y^2 \end{bmatrix} \tag{6.11}$$

where $\mathcal{V}$ is the volume containing the mass distribution.

Another related quantity is **the scalar moment of inertia** of a rigid body about some axis Δ which is defined by the matrix multiplication

[78] sss

$$\mathcal{I} = \hat{\boldsymbol{n}}\ \hat{\mathcal{I}}\ \hat{\boldsymbol{n}} \tag{6.12}$$

where $\hat{\mathcal{I}}$ is the moment inertia matrix with the elements $\mathcal{I}_{ij}$ and $\hat{n}$ is unit vector along the axis Δ. Note that the scalar moment of inertia is always positive.

The angular momentum of a rigid body made of N particles is given by

$$\boldsymbol{L} = \sum_\alpha \boldsymbol{r}_\alpha \times \boldsymbol{p}_\alpha = \sum_\alpha m_\alpha \boldsymbol{r}_\alpha \times (\boldsymbol{\omega} \times \boldsymbol{r}_\alpha) \tag{6.13}$$

With the help of the vector identity

$$\boldsymbol{A} \times (\boldsymbol{B} \times \boldsymbol{C}) = A^2 \boldsymbol{B} - (\boldsymbol{A}.\boldsymbol{B})\, \boldsymbol{A} \tag{6.14}$$

yields

$$\vec{L} = \sum_\alpha \left[\boldsymbol{\omega} r_\alpha^2 - (\boldsymbol{\omega}.\boldsymbol{r}_\alpha)\, \boldsymbol{r}_\alpha\right] \tag{6.15}$$

and which in component form reads

$$\begin{aligned} L_i &= \sum_j \omega_j \left[\sum_\alpha \omega_\alpha \left(\delta_{ij} r_\alpha^2 - x_{\alpha,i} x_{\alpha,j}\right)\right] \\ &= \sum_j \mathcal{I}_{ij}\ \omega_j \end{aligned} \tag{6.16}$$

Thus, we can write

$$\boldsymbol{L} = \hat{\mathcal{I}}.\boldsymbol{\omega} \equiv \begin{bmatrix} \mathcal{I}_{11} & \mathcal{I}_{12} & \mathcal{I}_{13} \\ \mathcal{I}_{12} & \mathcal{I}_{12} & \mathcal{I}_{23} \\ \mathcal{I}_{31} & \mathcal{I}_{32} & \mathcal{I}_{33} \end{bmatrix} \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix} \tag{6.17}$$

This implies that the rotational kinetic energy can be written as

$$T_{\text{rot}} = \tfrac{1}{2}\boldsymbol{\omega}\ \hat{\mathcal{I}}\ \boldsymbol{\omega} = \tfrac{1}{2}\vec{L}.\boldsymbol{\omega} = \tfrac{1}{2}\boldsymbol{\omega}.\boldsymbol{L} \tag{6.18}$$

6.2 Principal axes of inertia

In general, the inertia tensor is non-diagonal. To seek a diagonal form of $\mathcal{I}$ amounts to finding a new system of three axis for which the kinetic energy and the angular momentum take the form

$$\begin{aligned} T_{\text{rot}} &= \frac{1}{2}\sum_{i=1}^{3} \tilde{\mathcal{I}}_i\ \omega_i^2 \\ L_i &= \tilde{\mathcal{I}}_i\ \omega_i \qquad\qquad \text{(no summation over i)} \end{aligned} \tag{6.19}$$

In this case, these new axes are called **principal axes of inertia**. This means that given an inertial reference frame system within the rigid body, we can make a coordinate transformation $\boldsymbol{r}_\alpha \ \longrightarrow\ \mathcal{R}\ \boldsymbol{r}_\alpha$, such that

$$\mathcal{I}\ \longrightarrow\ \mathcal{R}^{-1}\ \mathcal{I}\ \mathcal{R} = \tilde{\mathcal{I}} = \text{diag}\left(\tilde{\mathcal{I}}_1, \tilde{\mathcal{I}}_2, \tilde{\mathcal{I}}_3\right) \tag{6.20}$$

where $\tilde{\mathcal{I}}_i$ are the eigenvalues of the moment of inertia tensor and they called **principal moments of inertia**. For an arbitrary vector $\boldsymbol{v}$, we have

$$\sum_{i,j} v_i \mathcal{I}_{ij} v_j = \sum_\alpha m_\alpha\ \left[|\boldsymbol{v}|^2\ r_\alpha^2 - (\boldsymbol{x}_\alpha.\boldsymbol{v})^2\right] \geq 0 \tag{6.21}$$

Hence, in the coordinate system where the moment of inertia is diagonal, the principal moment of inertia are always greater or equal to zero. Moreover, In this basis, we can write

$$\begin{aligned} \tilde{\mathcal{I}}_1 + \tilde{\mathcal{I}}_2 &= \sum_\alpha m_\alpha\ \left[2r_\alpha^2 - x_{\alpha,1}^2 - x_{\alpha,2}^2\right] \\ &= \sum_\alpha m_\alpha\ \left[x_{\alpha,1}^2 + x_{\alpha,2}^2 + 2x_{\alpha,3}^2\right] \geq \sum_\alpha m_\alpha\ \left[x_{\alpha,1}^2 + x_{\alpha,2}^2\right] = \tilde{\mathcal{I}}_3 \end{aligned} \tag{6.22}$$

So no moment of inertia can be larger than the sum of the other two. The **equality** occurs only if $x_3 = 0$ for all the point particles that constitute the rigid body, which means that it is a **planar body**.

Since the angular frequency $\boldsymbol{\omega}$ is a vector, it transforms as

$$\boldsymbol{\omega}\ \longrightarrow\ \mathcal{R}\ \boldsymbol{\omega} \tag{6.23}$$

Thus, the moment of inertia tensor of a rotating rigid body is diagonal when $\boldsymbol{\omega}$ aligns long a principal axis of the body.

Example 5.1: Principal axis of cube

Consider a cube of uniform density of side a, mass M, and with one corner placed at the origin. It is straight forward to show that its tensor of inertia in the (xyz) reference frame attached to the cube is given by

$$\hat{\mathcal{I}} = \mu \begin{bmatrix} 8 & -3 & -3 \\ -3 & 8 & -3 \\ -3 & -3 & 8 \end{bmatrix}, \qquad \mu = \frac{1}{12} M a^2 \tag{6.24}$$

which has the eigenvalues $\tilde{\mathcal{I}}_1 = 2\mu$, and $\tilde{\mathcal{I}}_2 = \tilde{\mathcal{I}}_3 = 11\mu$. To find the principal axis of this cube, we solve the equation

$$\left(\hat{\mathcal{I}}_{ab} - \hat{\mathcal{I}}_i \delta_{ab} \right) \omega_b = 0; \qquad i = 1, 2, 3 \tag{6.25}$$

Hence,

1. **principal axe 1**:

$$\begin{bmatrix} 6 & -3 & -3 \\ -3 & 8 & -3 \\ -3 & -3 & 8 \end{bmatrix} \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix} = 0 \;\rightarrow\; \omega_1 = \omega_2 = \omega_3 \;\rightarrow\; \hat{e}_1 = \frac{1}{\sqrt{3}} (1, 1, 1) \tag{6.26}$$

2. **principal axes 2 and 3**:

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix} = 0 \;\Rightarrow\; \omega_1 + \omega_2 + \omega_3 = 0 \tag{6.27}$$

Here we have the freedom to chose $\hat{e}_2$ and $\hat{e}_3$ as long as they are linearly independent with their components satisfy the above relation and such that $\{\hat{e}_1, \hat{e}_2, \hat{e}_3\}$ form orthonormal basis. For instance the principal axis can be defined as

$$\hat{\boldsymbol{e}}_1 = \frac{1}{\sqrt{3}} (1, 1, 1)\,; \qquad \hat{\boldsymbol{e}}_2 = \frac{1}{\sqrt{2}} (1, -1, 0); \quad \hat{\boldsymbol{e}}_3 = \frac{1}{\sqrt{6}} (1, 1, -2) \tag{6.28}$$

6.3 The theorem of parallel axis

Let $\mathcal{I}_{ij}$ is the tensor of inertia in the body-fixed frame. In another body-fixed frame displaced by vector distance $\boldsymbol{d}$, the position of a point-like mass can be written as

$$\boldsymbol{R}_\alpha = \boldsymbol{r}_\alpha + \boldsymbol{d} \tag{6.29}$$

where $\boldsymbol{r}_\alpha$ is the position of the point like mass with respect to the first frame of reference, Then, the tensor of inertia in the displaced body-fixed frame is given by

$$\begin{aligned} \mathcal{I}_{ij}(\boldsymbol{d}) &= \sum_\alpha m_\alpha \left[\delta_{ij} \left(\boldsymbol{r}_\alpha + \boldsymbol{a} \right)^2 - (x_{\alpha,i} + d_i)(x_{\alpha,j} + d_j) \right] \\ &= \sum_\alpha m_\alpha \left(\delta_{ij} r_\alpha^2 - x_{\alpha,i} x_{\alpha,j} \right) + \sum_\alpha m_\alpha \left(\delta_{ij} d^2 - d_i d_j \right) \\ &+ 2 \sum_{k=1}^{3} d_k \delta_{ij} \left(\sum_\alpha m_\alpha x_{\alpha,k} \right) - d_i \left(\sum_\alpha m_\alpha x_{\alpha,j} \right) - d_j \left(\sum_\alpha m_\alpha x_{\alpha,i} \right) \end{aligned} \tag{6.30}$$

If $\vec{r}_\alpha$ is measured with respect to the CM, then

$$\sum_\alpha m_\alpha \boldsymbol{r}_\alpha = 0 \tag{6.31}$$

which yields

$$\mathcal{I}_{ij}(\boldsymbol{d}) = \mathcal{I}_{ij}(\text{CM}) + M\left(\delta_{ij}d^2 - d_i d_j\right) \tag{6.32}$$

The above result is known as **the parallel axis theorem**[79].

The scalar moment of inertia with respect to an axis Δ that passes defined with a unit vector $\hat{n}$ is given by

$$\mathcal{I}(\boldsymbol{d}) = \mathcal{I}(\text{CM}) + MR_\perp^2 \tag{6.33}$$

where $R_\perp = \sqrt{d^2 - (\boldsymbol{d}.\hat{\boldsymbol{n}})^2}$ is the perpendicular distance from the Δ axis[80].

Example 5.2: Tensor of inertia of some systems with respect to their CM

Below we compute the tensor of inertia for some systems of total mass M and constant mass density in reference frame with the origin at their CM.

- **Rod of length** l
  Let the the CM of the rod be the origin the reference frame (xyz) with the $x-$ axis along the rod. Then, we have

$$\hat{\mathcal{I}}^{(\text{Rod})}(\text{CM}) = \frac{M}{\pi R^2}\begin{bmatrix} 0 & 0 & 0 \\ 0 & \int_{-l/2}^{l/2} x^2\ dx & 0 \\ 0 & 0 & \int_{-l/2}^{l/2} x^2\ dx \end{bmatrix} \tag{6.34}$$

  Thus,

$$\hat{\mathcal{I}}^{(\text{Rod})}(\text{CM}) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{Ml^2}{12} & 0 \\ 0 & 0 & \frac{Ml^2}{12} \end{bmatrix} \tag{6.35}$$

[79]It is sometimes known as Huygens-Steiner theorem, after **Christiaan Huygens** a Dutch mathematician, physicist and astronomer, and the Swiss mathematician **Jakob Steiner** (1796- 1863).

[80]Note that $R_\perp$ is independent of $\vec{d}$, the position of arbitrary point on the Δ axis.

- **Disk of radius R**
  This is a planar body, and so we are have

$$\hat{\mathcal{I}}^{(\text{Disk})}(\text{CM}) = \frac{M}{\pi R^2}\begin{bmatrix} \int y^2\ dxdy & -\int xy\ dxdy & 0 \\ -\int xy\ dxdy & \int x^2\ dxdy & 0 \\ 0 & 0 & \int\left(x^2+y^2\right)\ dxdy \end{bmatrix} \tag{6.36}$$

where the factor $M/\pi R^2$ is the mass density of the disk. invariance of the disk under the inversion of x or y implies $I_{xy} = 0$[81]. Moreover, the disk is invariant under the transformation $x \to y$, which yields $I_{xx} = I_{yy}$, and so $I_{zz} = 2I_{xx}$. With the use of the polar coordinates $x = r\cos\theta$ and $r = \sin\theta$, we have

$$\int x^2\ dxdy = \int_0^R r^3\ dr \int_0^{2\pi} \cos^2\theta\ d\theta = \frac{\pi R^4}{4} \tag{6.37}$$

Thus,

$$\boxed{\hat{\mathcal{I}}^{(\text{Disk})}(\text{CM}) = \begin{bmatrix} \frac{MR^2}{4} & 0 & 0 \\ 0 & \frac{MR^2}{4} & 0 \\ 0 & 0 & \frac{MR^2}{2} \end{bmatrix}} \tag{6.38}$$

- **Sphere of radius R**
  The sphere is invariant under rotation around arbitrary axis that pass by its center (which is its CM since we are assuming constant mass density), and hence

$$\begin{aligned} \mathrm{I}_{xy} &= \mathrm{I}_{xz} = \mathrm{I}_{yz} = 0 \\ \mathrm{I}_{xx} &= \mathrm{I}_{yy} = \mathrm{I}_{zz} \equiv \mathrm{I} \end{aligned} \tag{6.39}$$

To compute I, we slice the sphere into discs of mass dm parallel to the (xy) plane, each has a moment of inertia $d\mathrm{I}_{zz}$ with respect to the axis to the z-axis. Then, according to the result of Eq (6.35) we can write

$$\begin{aligned} \mathrm{I} &= \int d\ \mathrm{I}_{zz} = \int \frac{1}{2} \times dm \times \left(R^2 - z^2\right) \\ &= \int_{-R}^{R} \frac{1}{2} \times \left[\rho\ \pi\left(R^2 - z^2\right)\ dz\right] \times \left(R^2 - z^2\right) = \frac{8}{15}\pi\rho R^5 \\ &= \frac{2MR^2}{5} \end{aligned}$$

where we used the fact that $M = 4\pi/3\rho R^3$ to obtain the the last equation. Thus,

[81] This can be also shown explicitly as follow:

$$\int xy\ dxdy = \int_0^R r^3\ dr \int_0^{2\pi} \sin\theta\cos\theta\ d\theta = 0$$

$$\hat{\mathcal{I}}^{(\text{Disk})}(\text{CM}) = \begin{bmatrix} \frac{2MR^2}{5} & 0 & 0 \\ 0 & \frac{2MR^2}{5} & 0 \\ 0 & 0 & \frac{2MR^2}{5} \end{bmatrix} \tag{6.40}$$

- **Solid cone of height H and base radius R**
We will first compute the moment of inertia tensor with respect to the reference frame (x y z) that has as origin the tip of the cone. Due to the symmetry of the cone under rotation around the z-axis, we have

$$\begin{aligned} \mathrm{I}_{xy} &= \mathrm{I}_{xz} = \mathrm{I}_{yz} = 0 \\ \mathrm{I}_{xx} &= \mathrm{I}_{yy} \end{aligned} \tag{6.41}$$

Using the cylindrical coordinates (r, ϕ, z)[82], with $0 \leq \theta \leq 2\pi$, $0 < \rho < zR/H$, and $0 < z < H$, the infinitesimal volume element of the cone reads

$$dV = r\ dr\ d\phi\ dz \implies V = \int_0^H dz \int_0^{2\pi} d\phi \int_0^{\frac{zR}{H}} r\ dr = \frac{\pi}{3} R^2 H \tag{6.42}$$

Thus, we can write

$$\begin{aligned} I_{xx} = I_{yy} &= \frac{M}{\frac{\pi}{3}R^2 H} \int \left(z^2 + r^2 \sin^2\theta\right) \times r dr\ d\phi\ dz \\ &= \frac{3M}{R^2 H} \int_0^H dz \int_0^{\frac{zR}{H}} r dr \left(2z^2 + r^2\right) \\ &= \frac{3}{5} M \left(H^2 + \frac{R^2}{4}\right) \end{aligned} \tag{6.43}$$

For the component I_{zz} we have

$$\begin{aligned} I_{zz} &= \frac{M}{\frac{\pi}{3}R^2 H} \int r^2 \times r dr\ d\phi\ dz \\ &= \frac{6M}{R^2 H} \int_0^H dz \int_0^{\frac{zR}{H}} r^2\ r dr \\ &= \frac{3}{10} M R^2 \end{aligned} \tag{6.44}$$

So the moment of inertia tensor with respect to the tip of the cone reads

$$\hat{\mathcal{I}}^{(\text{cone})}(\mathbf{0}) = \begin{bmatrix} \frac{3}{5}M\left(H^2 + \frac{R^2}{4}\right) & 0 & 0 \\ 0 & \frac{3}{5}M\left(H^2 + \frac{R^2}{4}\right) & 0 \\ 0 & 0 & \frac{3}{10}MR^2 \end{bmatrix} \tag{6.45}$$

[82]Note that to avoid confusion with the mass density, we are denoting the distance in the spherical coordinate by r instead of ρ.

To calculate $\hat{\mathcal{I}}(\mathrm{CM})$, we need first to find the coordinates of the CM of the cone. Based on the invariance of the cone under rotation around the z axis, we deduce that $x_{CM} = y_{CM} = 0$, and

$$\begin{aligned} z_{\mathrm{CM}} &= \frac{M}{\frac{\pi}{3}R^2H}\int z \times dV = \frac{M}{\frac{\pi}{3}R^2H}\int_0^H zdz \int_0^{2\pi} d\phi \int_0^{\frac{zR}{H}} r\,dr \\ &= \frac{3M}{R^2H}\int_0^H z^3 \frac{R^2}{H^2}\,dz = \frac{3H}{4} \end{aligned} \tag{6.46}$$

Thus, according to the parallel axis theorem (6.32), and with $\boldsymbol{d} = 3H/4\ \hat{\boldsymbol{e}}_z$, the the components of the moment of inertia with respect to the CM are given by

$$\hat{\mathcal{I}}(\mathrm{CM}) = \hat{\mathcal{I}}_{ij}\mathbf{0} - M\frac{9}{16}MH^2\left(\delta_{ij} - \delta_{i3}\delta_{j3}\right) \tag{6.47}$$

and so

$$\boxed{\hat{\mathcal{I}}^{(\mathrm{cone})}(\mathrm{CM}) = \begin{bmatrix} \frac{3}{20}M\left(H^2+\frac{R^2}{4}\right) & 0 & 0 \\ 0 & \frac{3}{20}M\left(H^2+\frac{R^2}{4}\right) & 0 \\ 0 & 0 & \frac{3}{10}MR^2 \end{bmatrix}} \tag{6.48}$$

- **Moment of inertia tensor of ellipsoid**

An ellipsoid is defined by the equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \tag{6.49}$$

By rescaling the coordinates as follows

$$x = a\ \xi_1, \quad y = b\ \xi_2, \quad z = c\ \xi_3 \tag{6.50}$$

we can re-write (6.49) in the form

$$\xi_1^2 + \xi_2^2 + \xi_3^2 = 1 \tag{6.51}$$

which is the equation of a sphere of unit radius. Then, a volume element of the ellipsoid is given by

$$dV = dxdydz = abc\ d^3\boldsymbol{\xi} \tag{6.52}$$

where $d^3\boldsymbol{\xi}$ denotes the element of volume of a sphere. Thus, we have

$$V = abc\ d^3\boldsymbol{\xi} = \frac{4\pi}{3}abc \tag{6.53}$$

where I used the fact that the integral is the volume of a unit radius sphere.

Now since the ellipsoid is invariant under reflection symmetry, all the off diagonal elements of the moment of inertia vanish. Thus, the axes of the coordinate system (x y z) are principal axis. It convenient to use spherical coordinate

$$\xi_1 = \xi \sin\theta \cos\phi, \quad \xi_2 = \xi \sin\theta \sin\phi, \quad \xi_3 = \xi \cos\theta \tag{6.54}$$

where ξ is the distance from the center of the sphere, which is the origin of the coordinate system we are using, to the point inside a sphere. Then the diagonal elements of $\hat{\mathcal{I}}$ for an ellipsoid are given by[83]

<u>**$\mathbf{I_{xx} := I_1}$**</u>

$$\begin{aligned}
I_1 &= \frac{M}{V} \int \left(y^2 + z^2\right) dxdydz = abc\frac{M}{V} \int \left(b^2\xi_2^2 + c^2\xi_3^2\right)\ d^2\boldsymbol{\xi} \\
&= \frac{3}{4\pi} M \left[\int_0^1 \xi^4\ d\xi\right] \times \left\{a^2 \int_0^\pi \sin^3\theta\ d\theta \int_0^{2\pi} \sin^2\phi\ d\phi + b^2 \int \cos^2\theta \sin\theta\ d\theta \int_0^{2\pi} d\phi\right\} \\
&= \frac{1}{5}\left(b^2 + c^2\right) \qquad (6.55)
\end{aligned}$$

<u>**$\mathbf{I_{yy} := I_2}$**</u>

$$\begin{aligned}
I_2 &= \frac{M}{V} \int \left(x^2 + z^2\right) dxdydz = abc\frac{M}{V} \int \left(a^2\xi_1^2 + c^2\xi_3^2\right)\ d^2\boldsymbol{\xi} \\
&= \frac{3}{4\pi} M \left[\int_0^1 \xi^4\ d\xi\right] \times \left\{a^2 \int_0^\pi \sin^3\theta\ d\theta \int_0^{2\pi} \cos^2\phi\ d\phi + b^2 \int \cos^2\theta \sin\theta\ d\theta \int_0^{2\pi} d\phi\right\} \\
&= \frac{1}{5}\left(a^2 + c^2\right) \qquad (6.56)
\end{aligned}$$

<u>**$\mathbf{I_{zz} := I_3}$**</u>

$$\begin{aligned}
I_3 &= \frac{M}{V} \int \left(x^2 + y^2\right) dxdydz = abc\frac{M}{V} \int \left(a^2\xi_1^2 + b^2\xi_2^2\right)\ d^2\boldsymbol{\xi} \\
&= \frac{3}{4\pi} M \left[\int_0^1 \xi^4\ d\xi\right] \times \left[\int_0^\pi \sin^3\theta\ d\theta\right] \times \left\{a^2 \int_0^{2\pi} \cos^2\phi\ d\phi + b^2 \int_0^{2\pi} \sin^2\phi\ d\phi\right\} \\
&= \frac{1}{5}\left(a^2 + b^2\right) \qquad (6.57)
\end{aligned}$$

Thus,

[83] I evaluating I_1, I_2, and I_3we have used the following integrals of trigonometric functions

$$\begin{aligned}
\int_0^\pi \sin\theta\ d\theta &= 2; \quad \int_0^\pi \sin^2\theta\ d\theta = \int_0^\pi \cos^2\theta\ d\theta = \frac{1}{2}\int_0^{2\pi} d\phi = \pi \\
\int_0^\pi \cos^2\theta \sin\theta &= \frac{2}{3}; \quad \int_0^\pi \sin^3\theta\ d\theta = \int_0^\pi \left(1 - \cos^2\theta\right) \sin\theta = \frac{4}{3}
\end{aligned}$$

$$\hat{\mathcal{I}}^{(\text{ellipsoid})}(\text{CM}) = \begin{bmatrix} \frac{1}{5}M\left(b^2+c^2\right) & 0 & 0 \\ 0 & \frac{1}{5}M\left(a^2+c^2\right) & 0 \\ 0 & 0 & \frac{1}{5}M\left(a^2+b^2\right) \end{bmatrix} \tag{6.58}$$

- **Moment of inertia tensor of a right triangle**
Consider a right triangle on the (x-y) plane, with one side of length a along the x-axis, and the perpendicular has length b on the y-axis. We assume that it has a total mass M that is uniformly distributed in its surface. The x and y coordinates of a point on the hypotenuse of this triangle are related by

$$y = b\left(1 - \frac{x}{a}\right) \tag{6.59}$$

The center of mass of this triangle lies on the x-y plane, with coordinates

$$X_{\text{CM}} = \frac{1}{M}\int \rho\ x\ dx\ dy = \frac{\rho}{M}\int_0^a x\ dx \int_0^{b(1-x/a)} dy = \frac{1}{3}a \tag{6.60}$$

$$Y_{\text{CM}} = \frac{1}{M}\int \rho\ dx\ y\ dy = \frac{\rho}{M}\int_0^a dx \int_0^{b(1-x/a)} ydy = \frac{1}{3}b \tag{6.61}$$

where we used the fact that $M = \rho \times ab/2$.

Now let us compute the moment of inertia with respect to the reference frame (xyz) with the origin at the intersection point on of the two perpendicular sides of the triangle. Since this triangle is on the (x-y) plane, we have

$$I_{xz} = I_{yz} = 0, \quad I_{zz} = I_{xx} + I_{yy} \tag{6.62}$$

The remaining components are given by

$$\begin{aligned} I_{xx} &= \int \rho\ y^2\ dx\ dy = \rho\int_0^a dx \int_0^{b(1-x/a)} y^2\ dy = \frac{1}{6}Mb^2 \\ I_{yy} &= \int \rho\ x^2\ dx\ dy = \rho\int_0^a x^2\ dx \int_0^{b(1-x/a)} dy = \frac{1}{6}Ma^2 \\ I_{xy} &= \int \rho\ xy\ dx\ dy = \rho\int_0^a x\ dx \int_0^{b(1-x/a)} y\ dy = -\frac{1}{12}Mab \end{aligned} \tag{6.63}$$

Hence,

$$\hat{\mathcal{I}} = \frac{M}{6}\begin{bmatrix} b^2 & -\frac{1}{2}ab & 0 \\ -\frac{1}{2}ab & a^2 & 0 \\ 0 & 0 & (a^2+b^2) \end{bmatrix} \tag{6.64}$$

The moment of inertia tensor with respect to the the CM as can be obtained using the parallel axis theorem given in Eq (6.32) as

$$\hat{\mathcal{I}}(\text{CM}) = \mathcal{I}_{ij}(\vec{d}) - M\left(\delta_{ij}d^2 - d_i d_j\right) \tag{6.65}$$

with

$$\vec{d} = -\frac{a}{3}\hat{e}_x - \frac{b}{3}\hat{e}_y \;\Rightarrow\; M\left(\delta_{ij}d^2 - d_i d_j\right) = \frac{1}{9}\begin{bmatrix} b^2 & -ab & 0 \\ -ab & a^2 & 0 \\ 0 & 0 & (a^2+b^2) \end{bmatrix} \tag{6.66}$$

Thus,

$$\boxed{\hat{\mathcal{I}}(\mathrm{CM}) = M\begin{bmatrix} b^2 & \frac{1}{2}ab & 0 \\ \frac{1}{2}ab & a^2 & 0 \\ 0 & 0 & (a^2+b^2) \end{bmatrix}} \tag{6.67}$$

- **Rotating rectangle**
We consider a rectangle on the $x-z$ plane which has a width a in the $\hat{x}$ direction, and height b on the $\hat{z}$ direction. We assume that the rectangle has has mass M that is uniformly distributed over its surface. By symmetry the principal are the xyz axis with the origon at the center of the rectangle. In this reference frame the moment of inertia tensor is diagonal with elements given by

$$\begin{aligned} I_{xx} &= \int \rho\; dxdz\; \left(y^2+z^2\right) = \frac{M}{ab}\int_{-a/2}^{a/2} dx \int_{-b/2}^{b/2} dz\; z^2 = \frac{Mb^2}{12} \\ I_{yy} &= \int \rho\; dxdz\; \left(x^2+y^2\right) = \frac{M}{ab}\int_{-a/2}^{a/2} dx\; x^2 \int_{-b/2}^{b/2} dz = \frac{Ma^2}{12} \\ I_{zz} &= \int \rho\; dxdz\; \left(x^2+z^2\right) = I_1 + I_2 = \frac{M\left(a^2+b^2\right)}{12} \end{aligned}$$

Example 5.3: A System of Pulley and two masses on an Incline

A block of mass m_1 that can slide along an incline at an angle θ is attached to a mass m_2 hanging down the right vertical side via a rope which passes around a pulley in the shape of a uniform disk of radius R and mass m_p, as shown in Fig.21. Suppose that $m_2 > m_1$ so that when the system is released from rest the block of mass m_2 starts moving upward on the incline. Let x_i, y_i are the initial coordinates of the masses m_1, m_2 with respect to the X and Y axis, respectively, and x_f, y_f the corresponding coordinates after the masses have moved a distance d. Assuming that the rope does not slip on the pulley and that there is no friction in the system, we would like to determine the speed of m_2 (which is the same for m_1) after it has moved a distance d along the plane of the incline.

Let us chose the zero of the gravitational potential energy to be the horizontal plane at the same heigh as the center of the pulley. Then the initial mechanical energy of the system is just a gravitational potential energy, i.e.

$$E_{\text{initial}} = V_i = -m_2 g y_i - m_1 g x_i \sin\theta \tag{6.68}$$

and after the masses moved a distance, the energy is

$$E_{\text{final}} = V_f + T_f = \left[-m_2 g y_f - m_1 g x_f \sin\theta\right] + \left[\frac{1}{2}m_1 v_f^2 + \frac{1}{2}m_2 v_f^2 + \frac{1}{2}\mathrm{I}\omega^2\right] \tag{6.69}$$

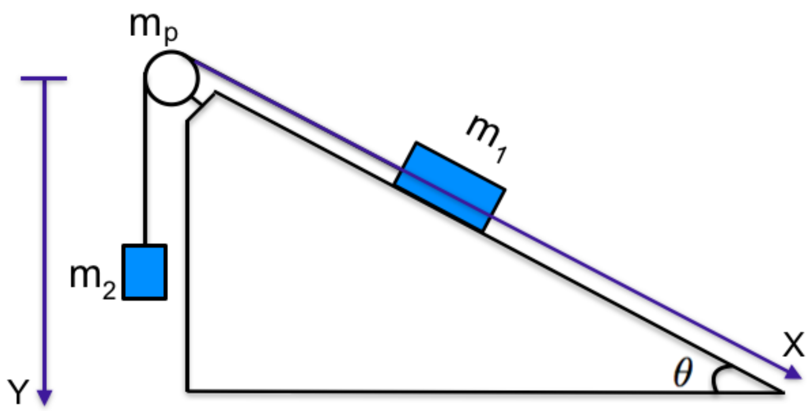


Figure 21: A system of a pulley and two masses on incline.

where v_f is the moving blocks, ω is the angular velocity of the pulley, and I its moment of inertia with respect to the axis passing through its center and perpendicular to its surface. Since we are assuming that the rope does not slip on the pulley, we have

$$\omega = \frac{v}{R} \tag{6.70}$$

and now the final energy reads

$$E_{\text{final}} = -m_2 g y_f - m_1 g x_f \sin\theta + \frac{1}{2}\left(m_1 + m_2 + \frac{\mathrm{I}}{R^2}\right) v_f^2 \tag{6.71}$$

Since there is no friction in the system, the total energy is conserved, and hence

$$-m_2 g y_i - m_1 g x_i = -m_2 g y_f - m_1 g x_f \sin\theta + \frac{1}{2}\left(m_1 + m_2 + \frac{\mathrm{I}}{R^2}\right) v_f^2 \tag{6.72}$$

After re-arranging the above equation, and using the fact that $(x_f - x_i) = (y_f - y_i) = d$, we can solve for the final speed as function of the distance and obtain

$$v_f = \sqrt{2gd}\sqrt{\frac{-m_2+m_1\sin\theta}{m_1+m_2+m_p/2}} \tag{6.73}$$

Example 5.4: Physical Pendulum

Fig. shows a physical pendulum which consists of a uniform road of length l and mass M pivoted about the point O at one of its end, and a disk of radius r mass m, attached at a distance R from the pivot point. We displace the pendulum to an angle θ_i and then its is released from rest.

We would like to determine the angular velocity of the pendulum when it passes by the lowest point of its swing. For that we will use the conservation of the total mechanical energy

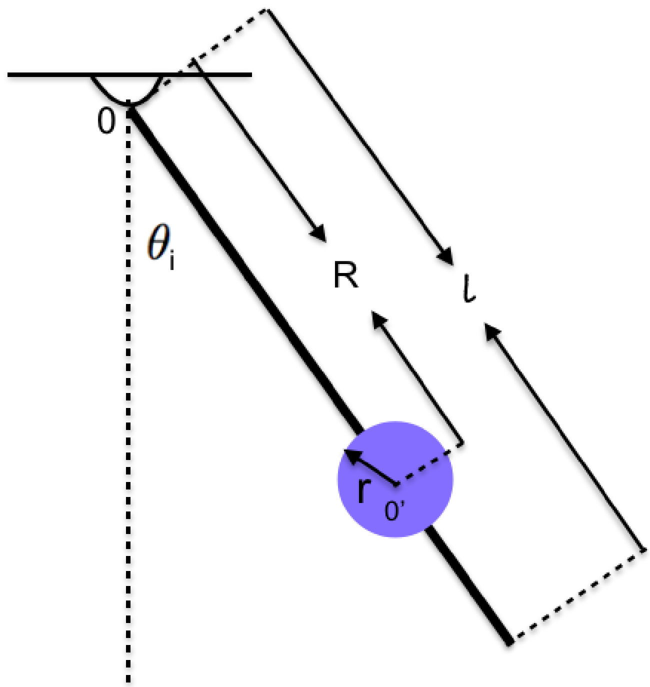


Figure 22: Physical pendulum consisting of rod and a disk.

of the system. By taking the zero point of the gravitational potential to be at 0, the initial and final energy of the physical pendulum are given by

$$E_{\text{initial}} = V_i = -Mg\frac{l}{2}\cos\theta_i - mgR\cos\theta_i \tag{6.74}$$

$$E_{\text{finall}} = V_f + T_f = -Mg\frac{l}{2} - mgR + \frac{1}{2}I^{(\text{system})}_{/\text{O}}\omega_f^2 \tag{6.75}$$

where $I^{(\text{system})}_{/\text{O}}$ is the moment of inertia with respect to the axis that passes through the point O and is perpendicular to the plane of the pendulum. It is given by

$$I^{(\text{system})}_{/\text{O}} = I^{(\text{rod})}_{/\text{O}} + I^{(\text{disk})}_{/\text{O}} \tag{6.76}$$

Using the theorem of parallel axis and the expressions of the moment inertia with respect to the center of mass derived in (6.35) and (6.38) for the a rod and the disk, respectively, we have

$$I^{(\text{system})}_{/\text{O}} = \left[M\frac{l^2}{12} + M\left(\frac{l}{2}\right)^2\right] + \left[m\frac{r^2}{2} + mR^2\right] \tag{6.77}$$

Substituting the above expression into Eq.(6.75), the conservation of mechanical energy read

$$-Mg\frac{l}{2}\cos\theta_i - mgR\cos\theta_i = -Mg\frac{l}{2} - mgR + \frac{1}{2}\left[M\frac{l^2}{3} + m\left(\frac{r^2}{2} + R^2\right)\right]\omega_f^2 \tag{6.78}$$

Thus, the angular velocity of the pendulum at the bottom of its swing is

$$\omega_f = \sqrt{\frac{g\,(Ml + 2mR)\,(1 - \cos\theta_i)}{M\frac{l^2}{3} + m\left(\frac{r^2}{2} + R^2\right)}} \tag{6.79}$$

Example 5.5: Equation of motion of a rolling ring on incline

Consider an inclined plane of angle θ and of mass M that is free to slide along a frictionless horizontal surface. A ring[84] of mass m and radius R is given an initial speed v_0 up the incline and starts rolling up until it gets to some point where it stops and rolls back down the incline. We would like to determine the maximum height the ring can rise before it starts rolling down the plane of the incline.

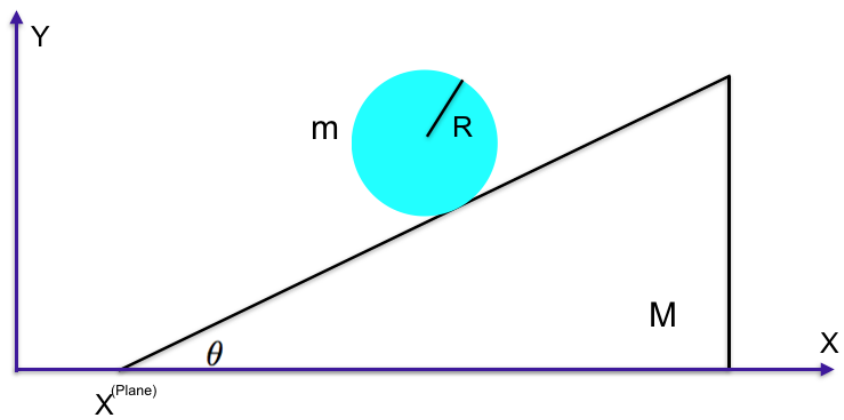


Figure 23: A ring rolling down the incline.

Let $(x^{(\text{ring})}, y^{(\text{ring})})$ be the coordinates of the ring with respect to the XY axes, and denote by h its y coordinate at the highest point on the incline where it stops rolling up. The position of the incline is given by $X^{(\text{plane})}$, the x-coordinate of the point at the bottom of the ramp . The only external force acting on the system ring-incline as a whole is the total weight $(M+m)g$, which is a conservative force, and hence the total energy of this system is conserve. Moreover, since there are no external forces along the x-axis, the total momentum of the system is also conserved. Thus, we have

$$\text{conservation of momentum}: mv_0\cos\theta = (M+m)\dot{X}^{(\text{plane})}_{\text{final}} \tag{6.80}$$

$$\text{conservation of total energy}: \frac{1}{2}mv_0^2 + \frac{1}{2}\omega_0^2 = \frac{1}{2}(M+m)\left(\dot{X}^{(\text{plane})}_{\text{final}}\right)^2 + mgh \tag{6.81}$$

where $X^{(\text{plane})}_{\text{final}}$ and $\dot{X}^{(\text{plane})}_{\text{final}}$ denote the position and the velocity of the incline when the ring stops rolling up, respectively, I is the moment of inertia of the ring about its axis of rotation, and ω_0 its initial angular velocity. Since the ring rolls without slipping, then its angular

[84]We could consider any object of mass m and moment of inertia I, sliding or rolling up the incline.

velocity $\omega = \upsilon/R$, with υ is the speed of its center of mass. Hence, the conservation of energy reads

$$\frac{1}{2}m\upsilon_0^2 + \frac{1}{2}\left(\frac{I}{R^2}\right)^2 \upsilon_0^2 = \frac{1}{2}(M+m)\left(\dot{X}_{\text{final}}^{(\text{plane})}\right)^2 + mgh \tag{6.82}$$

Now Eq.(6.80) yields

$$\dot{X}_{\text{final}}^{(\text{plane})} = \frac{\cos\theta}{1+M/m}\upsilon_0 \tag{6.83}$$

and after substituting it into Eq.(6.82) yields

$$h = \left(1 + \frac{I}{R^2} - \frac{\cos^2\theta}{1+M/m}\right)\frac{\upsilon_0^2}{2g} \tag{6.84}$$

Example 5.6: Falling Ladder: Lagrange Multiplier Method

Consider a ladder of length l and mass m with one end leaning on a smooth wall and the other end on a frictionless horizontal floor. The ladder is initially at rest and makes an angle θ_0 with the floor, as shown in Fig. 24.

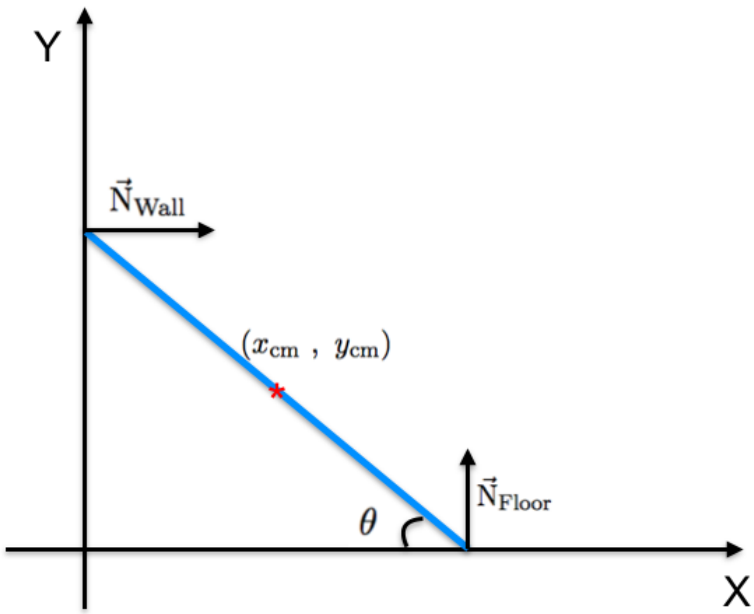


Figure 24: Sliding ladder.

In this example we would like to determine when the ladder lose contact with the wall. For that we will use the method of Lagrange multiplier where the motion of the ladder is described with three generalized coordinates: the angle θ, and the cartesian coordinates (x, y)

of the ladder's center of mass. By choosing the gravitational potential on the floor to be zero the Lagrangian of the system is given by

$$L = \frac{1}{2}m\left(\dot{x}^2 + \dot{y}^2\right) + \frac{1}{2}\mathrm{I}\dot{\theta}^2 + mgy \tag{6.85}$$

where I is the moment of inertia of the ladder about its end point. Using the parallel axis theorem and the expression of the moment of inertia of rod of mass m and length l, derived in Example 5.2, we find [85]

$$\mathrm{I} = \frac{1}{12}ml^2 + m\left(\frac{l}{2}\right)^2 = \frac{1}{3}ml^2 \tag{6.86}$$

The ladder is subjected to the following holonomic constraints:

$$\Phi_1 = x - \frac{l}{2}\cos\theta = 0 \tag{6.87}$$

$$\Phi_2 = y - \frac{\mathrm{l}}{2}\sin\theta = 0 \tag{6.88}$$

Thus the equation of motion reads

$$\frac{d}{dt}\left(\frac{\partial l}{\partial \dot{x}}\right) - \frac{\partial l}{\partial x} = \lambda_1\frac{\partial \Phi_1}{\partial x} + \lambda_2\frac{\partial \Phi_2}{\partial x} \tag{6.89}$$

$$\frac{d}{dt}\left(\frac{\partial l}{\partial \dot{y}}\right) - \frac{\partial l}{\partial y} = \lambda_1\frac{\partial \Phi_1}{\partial y} + \lambda_2\frac{\partial \Phi_2}{\partial y} \tag{6.90}$$

$$\frac{d}{dt}\left(\frac{\partial l}{\partial \dot{\theta}}\right) - \frac{\partial l}{\partial \theta} = \lambda_1\frac{\partial \Phi_1}{\partial \theta} + \lambda_2\frac{\partial \Phi_2}{\partial \theta} \tag{6.91}$$

or, equivalently,

$$m\ddot{x} = \lambda_1 \equiv \mathcal{C}_x \tag{6.92}$$

$$m\ddot{y} + mg = \lambda_2 \equiv \mathcal{C}_y \tag{6.93}$$

$$I\ddot{\theta} = \frac{\mathrm{L}}{2}\left(\lambda_1\sin\theta - \lambda_2\cos\theta\right) \equiv \mathcal{C}_\theta \tag{6.94}$$

Differentiating the constraints equations in (6.87) twice with respect to time, we get

$$\ddot{x} = -\frac{1}{2}l\,cos\theta\;\dot{\theta}^2 - \frac{l}{2}\sin\theta\;\ddot{\theta} \tag{6.95}$$

$$\ddot{y} = -\frac{1}{2}l\,sin\theta\;\dot{\theta}^2 + \frac{l}{2}\cos\theta\;\ddot{\theta} \tag{6.96}$$

A comparison with the equations of motion yields

$$\lambda_1 = -\frac{ml}{2}\left(\sin\theta\ddot{\theta} + \cos\theta\dot{\theta}^2\right) \tag{6.97}$$

$$\lambda_2 = \frac{ml}{2}\left(\cos\theta\ddot{\theta} - \sin\theta\dot{\theta}^2\right) \tag{6.98}$$

[85] We could also evaluate it directly as:

$$\mathrm{I} = \int_0^{\mathrm{l}} r^2\frac{m}{l}dr = \frac{1}{3}ml^2$$

where r is the length variable of the ladder.

Substituting the above expressions of λ_1 and λ_2 into (6.96) we obtain

$$\ddot{\theta} = -2\omega_0^2 \frac{\cos\theta}{\left(1 + \frac{4I}{ml^2}\right)}, \qquad \omega_0 = \sqrt{\frac{g}{l}} \tag{6.99}$$

The $\dot{\theta}^2$ in the expression of the Lagrange multipliers can be written as a function of θ solely by multiplying both sides of the above differential equation by $\ddot{\theta}$ and then integrate over θ, which gives

$$\dot{\theta}^2 = \left(\frac{4\omega_0^2}{1 + \frac{4I}{ml^2}}\right)(\sin\theta_0 - \sin\theta) \tag{6.100}$$

where θ_0 is the initial angle that the ladder makes with the floor. Thus, we have

$$\lambda_1(\theta) = -\frac{mg}{1 + \frac{4I}{ml^2}}(3\sin\theta - 2\sin\theta_0) \tag{6.101}$$

$$\lambda_2(\theta) = \frac{mg}{1 + \frac{4I}{ml^2}}\left[(3\sin\theta - 2\sin\theta_0) + \frac{4I}{ml^2}\right] \tag{6.102}$$

Thus, we deduced that the generalized forces of constraint, $\mathcal{C}_x$ and $\mathcal{C}_y$, which correspond to the normal forces exerted by the wall and the floor (see Fig.24), are

$$\mathrm{N}_{\mathrm{Wall}} = \sum_{\alpha=1}^{2} \lambda_\alpha \frac{\Phi_\alpha}{\partial x} = \lambda_1(\theta) \tag{6.103}$$

$$\mathrm{N}_{\mathrm{Floor}} = \sum_{\alpha=1}^{2} \lambda_\alpha \frac{\Phi_\alpha}{\partial y} = \lambda_2(\theta) \tag{6.104}$$

We can determine the angle θ_c at which the ladder lose contact with the wall by setting $\lambda_1(\theta)$ to zero. We get

$$\theta_c = \sin^{-1}\left[\tfrac{2}{3}\sin\theta_0\right] \tag{6.105}$$

and the time it takes for that to occur is

$$T_c = \int_{\theta_0}^{\theta_c} \frac{d\theta}{\dot{\theta}} = \frac{\sqrt{1+\frac{4I}{ml^2}}}{2\omega_0} \int \frac{d\theta}{\sqrt{\sin\theta_0 - \sin\theta}} \tag{6.106}$$

6.4 Euler's angles

Let us change the notation and denote by $(\hat{\boldsymbol{e}}_x, \hat{\boldsymbol{e}}_y, \hat{\boldsymbol{e}}_z)$ to be the unit vectors of the coordinate system associated with the laboratory frame (i.e. inertial), and by $(\hat{\boldsymbol{e}}_1, \hat{\boldsymbol{e}}_2, \hat{\boldsymbol{e}}_3)$ the unit vectors attached to to the rigid body after it has been rotated around some axis (see the figure below). The aim is to relate the two system of unit vectors.
Euler, noticed noted that a general rotation of the rigid body can be described by a successive three planar rotations as follows[86]:

[86] A general general rotation in 3 dimensions can be parameterized with three represented by 3 parameters, which we are the three angles of rotation. In d dimensional space, a general rotation can be parametrized with $d(d-1)/2$ parameters.

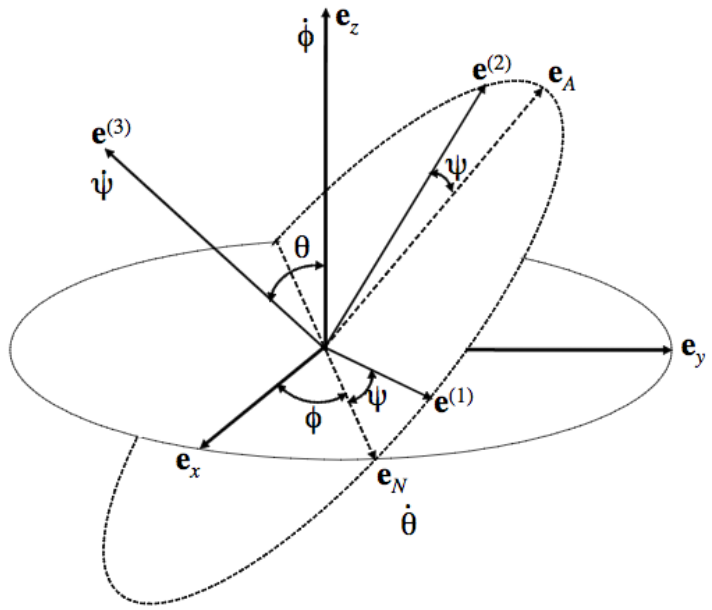


Figure 25: The three successive rotations with the the Euler's angles ϕ, θ, and ψ.

- **First rotation**:
A counter-clockwise rotation through an angle ϕ about z-axis, with angle ϕ can be within the interval range $[0, 2\pi]$. This transform the inertial axis into a new set of axis defines by the unit vector $(\hat{\boldsymbol{e}}_N, \hat{\boldsymbol{e}}'_y, \hat{\boldsymbol{e}}'_z)$, where $\hat{\boldsymbol{e}}'_z = \hat{\boldsymbol{e}}_z$, and $\hat{\boldsymbol{e}}'_y = \hat{\boldsymbol{e}}'_z \times \hat{\boldsymbol{e}}_N$.

- **Second rotation**:
A counter-clockwise rotation through an angle θ about direction of the the unit vector $\hat{\boldsymbol{e}}_N$, where θ can be within the interval range $[0, \pi]$. The result of this rotation is axis defined with the unit vectors $(\hat{\boldsymbol{e}}_N, \hat{\boldsymbol{e}}_A, \hat{\boldsymbol{e}}_3)$. The intersection of the planes spanned by $(\hat{\boldsymbol{e}}_x, \hat{\boldsymbol{e}}_y)$ and $(\hat{\boldsymbol{e}}_N, \hat{\boldsymbol{e}}_A)$ is the axis defined by $\hat{\boldsymbol{e}}_N$, called the **line of nodes**.

- **Third rotation**:
A counter-clockwise rotation through an angle ψ about direction of the the unit vector $\hat{\boldsymbol{e}}_3$, where ψ can be within the interval range $[0, \pi]$. The outcome of the final transformation is the axis defined by $(\hat{\boldsymbol{e}}_1, \hat{\boldsymbol{e}}_2, \hat{\boldsymbol{e}}_3)$.

Thus, the unit vectors of the rotating frame can be obtained from the ones of the laboratory frame by applying the three successive rotations described above. In equation form we have

$$\begin{bmatrix} \hat{\boldsymbol{e}}_1 \\ \hat{\boldsymbol{e}}_2 \\ \hat{\boldsymbol{e}}_3 \end{bmatrix} = \mathcal{R}_{\hat{\mathbf{e}}_3} \mathcal{R}_{\hat{\mathbf{e}}_N} \mathcal{R}_{\hat{\mathbf{e}}_z} \begin{bmatrix} \hat{\boldsymbol{e}}_x \\ \hat{\boldsymbol{e}}_y \\ \hat{\boldsymbol{e}}_z \end{bmatrix} \equiv \mathcal{R}_{\text{passive}} \begin{bmatrix} \hat{\boldsymbol{e}}_x \\ \hat{\boldsymbol{e}}_y \\ \hat{\boldsymbol{e}}_z \end{bmatrix} \tag{6.107}$$

where $\mathcal{R}_{\hat{\boldsymbol{e}}_i}$ is a planar rotation about the axis defined by the unit vector $\hat{\boldsymbol{e}}_i$, which according the above discussion are given by[87]

$$\mathcal{R}_{\hat{\boldsymbol{e}}_z} = \begin{bmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathcal{R}_{\hat{\boldsymbol{e}}_N} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{bmatrix}, \quad \mathcal{R}_{\hat{\boldsymbol{e}}_N} = \begin{bmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{6.108}$$

The transformation matrix $\mathcal{R}_{\text{passive}}$ is called passive rotation, where one rotate the unit vectors where as the vector position of point in space remains unchanged[88]. Taking the product of the above matrices yields

$$\mathcal{R}_{\text{passive}} = \begin{bmatrix} \cos\phi\cos\psi - \cos\theta\sin\phi\sin\psi & \sin\phi\cos\psi + \cos\theta\cos\phi\sin\psi & \sin\theta\sin\psi \\ -\cos\phi\sin\psi - \cos\theta\sin\phi\cos\psi & -\sin\phi\sin\psi + \cos\theta\cos\phi\cos\psi & \sin\theta\cos\psi \\ \sin\theta\sin\phi & -\sin\theta\cos\phi & \cos\theta \end{bmatrix} \tag{6.109}$$

The angular velocity vector rigid body angular velocity vector can be written

$$\boldsymbol{\omega} = \dot{\phi}\ \hat{\boldsymbol{e}}_z + \dot{\theta}\ \hat{\boldsymbol{e}}_N + \dot{\psi}\ \hat{\boldsymbol{e}}_3 \tag{6.110}$$

However, it is not convenient to use the above components for the angular velocity. Instead, we would like to write $\boldsymbol{\omega}$ in the laboratory or the rotating frame of reference. For that, we note from the figure that

$$\hat{\boldsymbol{e}}_z = \sin\theta\ \hat{\boldsymbol{e}}_A - \cos\theta\ \hat{\boldsymbol{e}}_3 \tag{6.111}$$
$$\hat{\boldsymbol{e}}_A = \sin\psi\ \hat{\boldsymbol{e}}_1 + \cos\psi\ \hat{\boldsymbol{e}}_2 \tag{6.112}$$
$$\hat{\boldsymbol{e}}_N = \cos\psi\ \hat{\boldsymbol{e}}_1 - \sin\psi\ \hat{\boldsymbol{e}}_2 \tag{6.113}$$

which we substitute into (6.110), to obtain the component of the angular velocity vector in the body frame (i.e. rotating frame)

$$\omega_1 = \dot{\phi}\sin\theta\sin\psi + \dot{\theta}\cos\psi \tag{6.114}$$
$$\omega_2 = \dot{\phi}\sin\theta\cos\psi - \dot{\theta}\sin\psi \tag{6.115}$$
$$\omega_3 = \dot{\phi}\cos\theta + \dot{\psi} \tag{6.116}$$

The motion associated with $\dot{\theta}$ is known as a **nutation**, $\dot{\phi}$ as a **precession** (for prolate bodies) or a **wobble** (for oblate bodies), and $\dot{\psi}$ as a **spin**.

[87] A rotation in a plane about an axis Δ with angle α is given by

$$\mathcal{R}_\Delta = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix} =$$

[88] An active rotation is where the frame axis remain unchanged where as the vector position is rotated. It is easy to show that the relation between passive rotation, $\mathcal{R}_{\text{passive}}$, and the active rotation, $\mathcal{R}_{\text{active}}$, is that one is the transpose of the other.

The component of the angular velocity in the laboratory frame (i.e. inertial frame), can be determined by writing $\hat{\boldsymbol{e}}_N$ and $\hat{\boldsymbol{e}}_3$ in the basis $(\hat{\boldsymbol{e}}_x, \hat{\boldsymbol{e}}_y, \hat{\boldsymbol{e}}_z)$ as

$$\hat{\boldsymbol{e}}_N = \cos\phi\ \hat{\boldsymbol{e}}_x + \sin\phi\ \hat{\boldsymbol{e}}_y \tag{6.117}$$

$$\hat{\boldsymbol{e}}_3 = -\sin\theta\sin\phi\ \hat{\boldsymbol{e}}_x - \sin\theta\cos\phi\ \hat{\boldsymbol{e}}_y + \cos\theta\ \hat{\boldsymbol{e}}_z \tag{6.118}$$

where the above expression of $\hat{\boldsymbol{e}}_N$ was deduced from the figure above, and for $\hat{\boldsymbol{e}}_3$ we used the matrix $\mathcal{R}_{\text{passive}}$ that relate the rotating basis to the laboratory one. Replacing the above relations into (6.110), we obtain

$$\omega_x = \dot{\psi}\sin\theta\sin\phi + \dot{\theta}\cos\phi \tag{6.119}$$

$$\omega_y = -\dot{\phi}\sin\theta\cos\phi - \dot{\theta}\sin\phi \tag{6.120}$$

$$\omega_z = \dot{\psi}\cos\theta + \dot{\phi} \tag{6.121}$$

6.5 Euler equations

In the previous chapter we have seen that the torque on a system of particles in an inertial frame is equal to the time derivative of the angular momentum:

$$\boldsymbol{\tau}^{(\text{ext})} = \left[\frac{d\boldsymbol{L}}{dt}\right]_{\text{inertial}} \tag{6.122}$$

which according to (5.9) can written as

$$\vec{\tau}^{(\text{ext})} = \left[\frac{d\vec{L}}{dt}\right]_{\text{rotating}} + \vec{\omega}\times\vec{L} \tag{6.123}$$

By choosing the fixed-body frame to be the reference frame in which its coordinates axis are the principal axes of the rigid body, we obtain **Euler's equations**

$$\begin{aligned} I_1\ \dot{\omega}_1 &= (I_2 - I_3)\,\omega_2\omega_3 + \tau_1^{(\text{ext})} \\ I_2\ \dot{\omega}_2 &= (I_3 - I_1)\,\omega_3\omega_1 + \tau_2^{(\text{ext})} \\ I_3\ \dot{\omega}_3 &= (I_1 - I_2)\,\omega_1\omega_2 + \tau_3^{(\text{ext})} \end{aligned} \tag{6.124}$$

For a free torque motion, Euler's equations become

$$I_1\ \dot{\omega}_1 = (I_2 - I_3)\,\omega_2\omega_3 \tag{6.125}$$

$$I_2\ \dot{\omega}_2 = (I_3 - I_1)\,\omega_3\omega_1 \tag{6.126}$$

$$I_3\ \dot{\omega}_3 = (I_1 - I_2)\,\omega_1\omega_2$$

The first two equations can be re-written as

$$\begin{aligned} I_1\,(I_1 - I_3)\,\omega_1\dot{\omega}_1 &= \omega_1\omega_2\omega_3\,(I_1 - I_3)\,(I_2 - I_3) \\ I_2\,(I_2 - I_3)\,\omega_2\dot{\omega}_2 &= -\omega_1\omega_2\omega_3\,(I_1 - I_3)\,(I_2 - I_3) \end{aligned} \tag{6.127}$$

Adding these two equations gives

$$\begin{aligned} 0 &= I_1\,(I_1 - I_3)\,\omega_1\dot{\omega}_1 + I_2\,(I_2 - I_3)\,\omega_2\dot{\omega}_2 \\ &= \frac{d}{dt}\Big\{I_1\,(I_1 - I_3)\,\omega_1^2 + I_2\,(I_2 - I_3)\,\omega_2^2\Big\} \end{aligned} \tag{6.128}$$

which has the solution

$$I_1\left(I_1 - I_3\right)\omega_1^2 + I_2\left(I_2 - I_3\right)\omega_2^2 = C_1 \tag{6.129}$$

where C_1 is a constant. Similarly, by combining the two last equations in (6.125) we get

$$I_2\left(I_1 - I_2\right)\omega_2^2 + I_3\left(I_1 - I_3\right)\omega_3^2 = C_2 \tag{6.130}$$

Using the two equations (6.129) and (6.130) we can solve for two angular frequencies in terms of a third one. Without loss of generality we may assume that $I_1 < I_2 < I_3$, so that we can express ω_1 and ω_3 in terms of ω_2 as

$$\omega_1 = \sqrt{\frac{1}{I_1\left(I_1 - I_3\right)}\left[C_1 - I_2\left(I_2 - I_3\right)\omega_2^2\right]} \tag{6.131}$$

$$\omega_3 = \sqrt{\frac{1}{I_3\left(I_1 - I_3\right)}\left[C_2 - I_2\left(I_1 - I_2\right)\omega_2^2\right]} \tag{6.132}$$

The energy of this system, which in the case of free-torque is the kinetic energy, reads

$$\begin{aligned} E &= \frac{1}{2}\left[I_1\omega_1^2 + I_2\omega_2^2 + I_3\omega_3^2\right] \\ &= \frac{1}{I_1 - I_2}\left(C_1 + C_2\right) \end{aligned} \tag{6.133}$$

Thus, the sum of the constants C_1 and C_2 is related to the energy of the rigid body.

Now, let us substitute the expressions of ω_1 and ω_3 in (6.131) into the differential equation for ω_2, gives

$$-\sqrt{\frac{C_1 C_2}{I_1 I_2^2 I_3}}\, t = \frac{d\omega_2}{\sqrt{\left(1 - \frac{I_2(I_2 - I_3)}{C_1}\omega_2^2\right)\left(1 - \frac{I_2(I_1 - I_2)}{C_1}\omega_2^2\right)}} \tag{6.134}$$

Making the change of variable

$$\xi = \sqrt{\frac{I_2\left(I_2 - I_3\right)}{C_1}}\,\omega_2 \tag{6.135}$$

we obtain

$$\mathcal{F}\left(\sqrt{I_2\left(I_2 - I_3\right)C_1}\,\omega_2,\ \frac{C_1\left(I_1 - I_2\right)}{C_2\left(I_2 - I_3\right)}\right) = -\sqrt{\frac{I_2\left(I_2 - I_3\right)}{C_1}}\sqrt{\frac{C_1 C_2}{I_1 I_2^2 I_3}}\, t \tag{6.136}$$

with

$$\mathcal{F}\left(t_*, k\right) = \int_0^{t_*} \frac{dt}{\sqrt{\left(1 - t^2\right)\left(1 - k^2 t^2\right)}} \tag{6.137}$$

which is **the elliptic integral of the first kind**.

6.6 Torque-free Symmetric Tops

The above solution for free-torque involve a complicated hyper-geometric function, however if the rigid body possess an axis of symmetry so that two of the three principal moments of inertia are equal. Let us chose the axis of symmetry to be the z-axis, then

$$I_1 = I_2 := I \tag{6.138}$$

In this case, Euler's equations read

$$\begin{aligned} I\,\dot{\omega}_1 &= (I - I_3)\,\omega_2\omega_3 && (6.139) \\ I\,\dot{\omega}_2 &= (I_3 - I)\,\omega_3\omega_1 && (6.140) \\ I_3\,\dot{\omega}_3 &= 0 \end{aligned}$$

which shows that ω_3 is constant. We define

$$\Omega = \left(\frac{I - I_3}{I_1}\right)\omega_3 \tag{6.141}$$

so that Euler equations for ω_1 and ω_2 read

$$\begin{aligned} \dot{\omega}_1 &= \Omega\dot{\omega}_2 && (6.142) \\ \dot{\omega}_2 &= -\Omega\dot{\omega}_1 && (6.143) \end{aligned}$$

These two coupled first order differential equations can brought into two decoupled second order differential equations given by

$$\begin{aligned} \ddot{\omega}_1 + \Omega^2\omega_1 &= 0 && (6.144) \\ \ddot{\omega}_2 + \Omega^2\omega_2 &= 0 && (6.145) \end{aligned}$$

With the use of Eq (6.143), the solution to the above equations reads

$$\begin{aligned} \omega_1 &= \omega_\perp \cos(\Omega t + \delta) && (6.146) \\ \omega_2 &= \omega_\perp \sin(\Omega t + \delta) \end{aligned}$$

where $\omega_\perp$ and δ are constants given by the initial conditions. Hence, the frequency vector $\boldsymbol{\omega}$ has a constant length $|\vec{\omega}| = \left(\omega_\perp^2 + \omega_3^2\right)^{1/2}$ that rotates around the body symmetry axis (i.e. along $\hat{\boldsymbol{e}}_3$) with frequency Ω which makes a constant angle $\theta = \tan^{-1}(\omega_\perp/\omega_3)$ with respect to it. This rotation of the $\boldsymbol{\omega}$ around one of the principal axis of the body is called **spin**. The cone traced around the body symmetry axis is called the **body cone** and the direction of the spin depends on the sign of Ω; that is whether $I > I_3$ (cigar-shaped, i.e. oblate) or $I < I_3$ (i.e. pancake-shaped, i.e. prolate)[89].

The above results were derived in the rotating frame which is fixed to the rigid body. To translate the above results in an inertial lab frame, we find the motion with respect to the axis of the total angular momentum $\boldsymbol{L}$ which is is constant of motion. We will show that $\hat{\boldsymbol{e}}_3$ rotates with frequency $\boldsymbol{\omega}_{\text{pre}} \neq \Omega$ about the axis of the angular momentum. The rotation of the symmetry axis about the the axis of $\boldsymbol{L}$ is called **precession** which looks like a **wobble**

[89]It is important to note that the axis $(\hat{\boldsymbol{e}}_1, \hat{\boldsymbol{e}}_2, \hat{\boldsymbol{e}}_3)$ are rotating in space with frequency $\boldsymbol{\omega}$.

to an observer in an inertial reference frame.

An example of rotating system that can be approximated to a symmetric top is provided by the Earth[90], which is approximately an oblate spheroid (i.e. special case of an ellipsoid where $a = b$) with moments of inertia

$$I_1 = I_2 =: I = \frac{1}{5}M\left(a^2 + c^2\right); \qquad I_3 = \frac{2}{5}Ma^2 \tag{6.147}$$

where $2c = 12714$ km is the pole-to-pole distance, and $2a = 12756$ km is the equatorial diameter[91]. So, we have

$$\left(\frac{I_3 - I}{I}\right)_{\text{Earth}} = \frac{a^2 - c^2}{a^2 + c^2} \simeq \frac{1}{300} \tag{6.148}$$

With the above value the north pole spins around the symmetry axis with a period[92],

$$T_s = \frac{2\pi}{\Omega} \simeq \frac{2\pi}{2\pi/\text{day}} \times 300 = 300 \text{ days} \tag{6.149}$$

A tiny wobbling of the Earth polar axis with a period of $T_s = 427$ days was detected in 1891 by the American astronomer **Seth Chandler** (1846- 1913) and now this effect is known as **Chandler wobble**. The recent measurement of the period of wobbling gives a value of 440 days, almost 50% larger than the predicted value. This discrepancy is attributed to the fact that Earth is not perfectly rigid body due to tidal effects[93].

6.7 Heavy Symmetric Top

Example 5.7: The Rolling Disk

Let a disk of mass M and radius R rolling with slipping on horizontal surface.

[90] We are neglecting the small torque caused by the Moon and the Sun on the equatorial bulge.
[91] The bulge at the equator is caused by the rotation of the Earth.
[92] This calculation was first performed by **Leonard Euler** in **1749**.
[93] Recent discoveries show that the molten core of the Earth rotates with a speed different than the the mantle. However many researchers don't think this is the full explanation.

7 Small Oscillations

7.1 Harmonic Oscillator

In general, the equations of motions for system that undergoes oscillations are non linear, and not easy to solve analytically. Thus, as a first step to solve these equations, we linearize them assuming that the amplitudes of the oscillations to be small. To start with, consider a particle moving in one dimension under the effect of a potential $V(x)$, such that

$$\left(\frac{dV}{dx}\right)(x_0) = 0 \tag{7.1}$$

which means that there exist a point x_0 at which the system is in equilibrium. Expanding the force, $V(x)$, around x_0, we get

$$V(x) = \frac{1}{2}\frac{d^2V}{dx^2}|_{x_0}(x-x_0)^2 + \frac{1}{6}\frac{d^3V}{dx^3}|_{x_0}(x-x_0)^3 + ... + \frac{1}{n!}\frac{d^nV}{dx^n}|_{x_0}(x-x_0)^n + ... \tag{7.2}$$

where we used the fact that x_0 is an equilibrium point. Furthermore, if $|x - x_0| << 1$, we can approximate the potential by

$$V(x) \simeq \tfrac{1}{2}k(x-x_0)^2 \tag{7.3}$$

where

$$k := \frac{d^2V}{dx^2}|_{x_0} \tag{7.4}$$

Thus the force implies applied to this oscillating particle is

$$F(x) \simeq -k(x-x_0) \tag{7.5}$$

If $k is positive, i.e. \frac{d^2V}{dx^2}|_{x_0} > 0$, then x_0 is a stable equilibrium point. In this case $F(x)$ is a restoring force pointing toward the point x_0, called **Hooke's law**, and which applies to most systems when displaced small distance from equilibrium. The Lagrangian describing this system is given by

$$L = \tfrac{1}{2}m\dot{\eta}^2 - \tfrac{1}{2}k\eta^2 \tag{7.6}$$

where $\eta = (x - x_0)$ is the small displacement around the stable equilibrium point. The Euler-Lagrange equations read

$$\ddot{\eta} + \omega_0^2\eta = 0; \qquad \omega_0 = \sqrt{\frac{k}{m}} \tag{7.7}$$

which is known as the equation of motion of **harmonic oscillator**. The solution of this differential equation can be written in the form

$$\eta(t) = A \sin \omega_0 (t - t_0) \tag{7.8}$$

where A and t_0 are real constants that can be determined from the initial conditions. The frequency ω_0 is called the **natural frequency** of the harmonic oscillator.

Example 6.1: Simple pendulum

Consider a simple pendulum[94] of mass m and length l, displaced by a small distance along its arc from its equilibrium position. The potential energy is

$$V = mgl(1 - \cos\theta) \tag{7.9}$$

which for small displacement it can be approximated as

$$\begin{aligned} V &\simeq mgl\left(\frac{\theta^2}{2} - \frac{\theta^4}{4!} + ...\right) \\ &\simeq \frac{mg}{l}\left(\frac{s^2}{2} - \frac{1}{l^2}\frac{s^4}{4!}\right) \end{aligned} \tag{7.10}$$

where $s = l\theta$ is the length of the arc enclosing an angle θ. Thus, for $s << l$, by neglecting terms of order s^4 and higher, the Lagrangian reads

$$L = \frac{1}{2}m\dot{s}^2 - \frac{1}{2}ks^2; \quad k = \frac{mg}{l} \tag{7.11}$$

which has the same form as the Lagrangian of harmonic oscillator, an so the motion is oscillatory with a period

$$T_0 = 2\pi\sqrt{\frac{l}{g}} \tag{7.12}$$

So for $\theta << 1$, the period is independent of the amplitude of the oscillations.

If the displacement s is not too small compared to the length of the pendulum, then one should include higher order corrections in the expansion series of the potential. In this case, to estimate the period of oscillations, we make use of the conservation of the mechanical energy of the pendulum, i.e.

$$E := \frac{1}{2}ml^2\dot{\theta}^2 + mgl(1 - \cos\theta) = mgl(1 - \cos\theta_{\max}) \tag{7.13}$$

where $\theta_{\max}$ is angle at which the kinetic energy vanishes. Solving for the angular velocity $\dot{\theta}$, gives

$$\dot{\theta} = \pm\sqrt{\frac{2g}{l}(\cos\theta - \cos\theta_{\max})} \tag{7.14}$$

[94]The pendulum clock was invented by Christiaan Huygens in the winter of 1656-1657.

where the $+(-)$ sign corresponds to the motion in counter-clockwise (clockwise). So, by integrating $(\dot{\theta})^{-1}$ with respect to the angle θ from max to 0 (i.e. counter-clockwise), we get a quarter of the period of oscillation:

$$T = 4\int_0^{\theta_{\text{max}}} \frac{d\theta}{\dot{\theta}} = \sqrt{\frac{8l}{g}} \int_0^{\theta_{\text{max}}} \frac{d\theta}{\sqrt{\cos\theta - \cos\theta_{\text{max}}}} \tag{7.15}$$

Using the identity $\cos\theta = 1 - \sin^2\theta/2$, and making a change of variable $\sin\phi = \sin\theta/2 \sin\theta_{\text{max}}/2$, we can re-write T in the following form

$$T = T_0\left[\frac{2}{\pi}\mathcal{K}(\delta)\right] \tag{7.16}$$

where

$$\mathcal{K} = \int_0^{\pi/2} \frac{d\phi}{\sqrt{1-\delta^2\sin^2\phi}}, \qquad \delta = \sin\theta_{\text{max}} \tag{7.17}$$

The above integral is called **elliptic integral of the first kind**, which for $k \simeq 0$, approximates to $\pi/2$. In our case of θ_{max} being small (compared to $\pi/2$), but not very small, we Taylor expand the above integrant in powers of δ^2, and get

$$\begin{aligned} \mathcal{K}(\delta) &\simeq \int_0^{\pi/2}\left[1+\frac{\delta^2}{2}\sin^2\phi + \frac{3\delta^4}{8}\sin^4\phi + ...\right] d\phi \\ &\simeq T_0\left[1+\left(\frac{1}{2}\right)^2\delta^2 + \left(\frac{3}{8}\right)^2\delta^4 + ..\right] \end{aligned} \tag{7.18}$$

Keeping the lowest order term in δ^2, we can approximate the period by

$$\boxed{T \simeq T_0\left(1+\frac{\theta^2_{\text{max}}}{16}\right) \;\; > \; T_0} \tag{7.19}$$

which shows that the period varies with the amplitude.
Recently, other approximation formulas for the period have been proposed. The simplest of these, is the so called Kidd-Fogg formula (see reference [20]), given by

$$\boxed{T^{(\text{KF})} = T_0(1-\delta^2)^{-1/4} = T_0\frac{1}{\sqrt{\cos\theta_{\text{max}}}}} \tag{7.20}$$

This formula is in good agreement with the data with error less than 1% for angles close to $\pi/2$. For more details about other approximate expressions for the period for large amplitude oscillations of simple pendulum see the list of references [21–25].

7.2 Damped Harmonic Oscillator

In reality, most of the harmonic oscillators contain some sort of friction within them. Here we will assume a damping force linear in velocity is applied to the harmonic oscillator, i.e.

$$F^{(\mathrm{damping})} = -b\,\dot{x} \tag{7.21}$$

where b is a real positive number. Thus, the equation of motion is given by[95]

$$\boxed{\ddot{x} + 2\beta\dot{x} + \omega_0^2 x = 0} \tag{7.22}$$

with $\beta = b/2m$, which has the dimension of time inverse of . As a trial solution to the above equation, we use

$$x(t) = \mathrm{Re}\left[Ae^{\alpha t}\right] \tag{7.23}$$

Since we are dealing a "linear" differential equation, we can leave out the "Re" and simply work work with $Ae^{\alpha t}$ and only at the end of the calculations we take the real part. By substituting $x(t) = Ae^{\alpha t}$ into (7.22) we get the **characteristic equation**:

$$Ae^{\alpha t}\left[\alpha^2 + 2\beta\alpha + \omega_0^2\right] = 0 \tag{7.24}$$

Thus, in order to have non-trivial solutions requires that

$$\alpha = -\beta \pm \sqrt{\beta^2 - \omega_0^2} \tag{7.25}$$

Thus, the general solution reads

$$\boxed{x(t) = e^{-\beta t}\mathrm{Re}\left[c_1 e^{\sqrt{\beta^2-\omega_0^2}t} + c_2 e^{-\sqrt{\beta^2-\omega_0^2}t}\right]} \tag{7.26}$$

where c_1 and c_2 are, in general, complex numbers. There are three possible regimes of oscillations depending on the value of $(\beta^2 - \omega_0^2)$.

1. $\underline{\beta < \omega_0}$: **Weak damping** (Under damping)

   In this case, we can write $\alpha = -\beta \pm i\omega_1$, with

$$\omega_1 = \sqrt{\omega_0^2 - \beta^2} < \omega_0 \tag{7.27}$$

   Then, the solution is

$$x(t) = e^{-\beta t}\mathrm{Re}\left[c_1 e^{i\omega_1 t} + c_2 e^{-i\omega_1 t}\right] \tag{7.28}$$

   which yields

[95] Note that for an electrical oscillator, such as a series of LRC circuit, the charge q on the capacitor is a solution of the differential equation

$$\ddot{q} + \frac{R}{L}\dot{q} + \frac{1}{LC}q = 0$$

which has exactly the same form of a damped oscillator with $\beta = R/2L$ and $\omega_0 = 1/\sqrt{LC}$.

$$x^{(\mathrm{WD})}(t) = e^{-\beta t}\left(a\cos\omega_1 t + b\sin\omega_1 t\right) \equiv \mathcal{A}e^{-\beta t}\cos\left(\omega t + \phi\right) \tag{7.29}$$

with a, b, $\mathcal{A}$, and ϕ real constants, that can be determined from the initial conditions[96]. We see that the decay time, or the damping time, of these oscillations is

$$\tau^{(\mathrm{damping})} = \frac{1}{\beta} = \frac{2b}{m} \tag{7.31}$$

2. $\underline{\beta > \omega_0}$: **Strong damping** (Over damping)

   In this case α is real, given by $\alpha = -\beta \pm \sqrt{\beta^2 - \omega_0^2}$ which is real. So, we have

$$x(t) = e^{-\beta t}\left(C_1 e^{\sqrt{\beta^2-\omega_0^2}t} + C_2 e^{-\sqrt{\beta^2-\omega_0^2}t}\right) \tag{7.32}$$

   or, equivalently,

$$x^{(\mathrm{SD})}(t) = e^{-\beta t}\left(a\cosh\sqrt{\beta^2 - \omega_0^2}t + b\sinh\sqrt{\beta^2 - \omega_0^2}t\right) \tag{7.33}$$

   and hence the motion is non-oscillatory. For instance, if we assume the initial conditions $x(0) = x_0$ and $\dot{x}(0) = 0$, then

$$x(t) = x_0 e^{-\beta t}\left(\cosh\sqrt{\beta^2 - \omega_0^2}t + \frac{\beta}{\sqrt{\beta^2 - \omega_0^2}}\sinh\sqrt{\beta^2 - \omega_0^2}t\right) \tag{7.34}$$

   which for $t >> 1/\beta$ behaves as

$$x(t) \sim x_0 e^{-(\beta - \sqrt{\beta^2-\omega_0^2})t}\left[\frac{1 + \beta/\sqrt{\beta^2 - \omega_0^2}}{2}\right] \tag{7.35}$$

   and goes asymptotically to its equilibrium position $x = 0$.

---

[96]For example, with the initial conditions $x(0) = x_0$ and $\dot{x}(0) = 0$, we find

$$\begin{aligned}(a, b) &= (x_0,\ \frac{\beta}{\omega_1}`x_0)\\ (\mathcal{A}, \phi) &= \left(\frac{x_0}{\sqrt{1 + \beta^2/\omega_1^2}},\ -\tan^{-1}\left(\frac{\beta}{\omega_1}\right)\right)\end{aligned} \tag{7.30}$$

3. $\underline{\beta = \omega_0}$: **Critical damping**

This implies that characteristic equation has two repeated roots $\alpha = -\beta$. Substituting this root in Eq(7.26), we can easily check that there two independent solutions: $e^{-\beta t}$ and $t\, e^{-\beta t}$. Thus, the most general solution reads[97]

$$x(t) = e^{-\beta t}\,(A + Bt) \tag{7.36}$$

or, equivalently, in terms of the initial conditions $x(0) \equiv x_0$ and $\dot{x}(0) \equiv \dot{x}_0$, as

$$x^{(\mathrm{CD})}(t) = e^{-\beta t}\left[x_0(1+\beta t) + \dot{x}_0 t\right] \tag{7.37}$$

As in the strong damped case, this is a non-oscillatory solution. However note that in the critical regime, the system return to its equilibrium position faster than the case of strong damping[98].

7.3 Energy of Under-damped Harmonic Oscillator

Because damping is a friction, the mechanical energy of the system, $E = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}kx^2$, is not conserved. If we consider an under damped oscillator, the energy reads

$$\begin{aligned} E = {} & \frac{1}{2}m\left[\beta^2\mathcal{A}^2 e^{-2\beta t}\cos^2\omega_1 t + \delta + \omega_1^2\mathcal{A}^2 e^{-2\beta t} + 2\beta\omega_1\cos(\omega_1 t + \delta)\sin(\omega_1 t + \delta)\right] + \\ & \frac{1}{2}k\mathcal{A}^2 e^{-2\beta}\cos^2(\omega_1 t + \delta) \end{aligned} \tag{7.38}$$

where we used $x(t) = \mathcal{A}\cos(\omega_1 t + \delta)$ for the the expression of the displacement. By averaging the energy over a cycle of oscillation, yields[99]

$$< E >= E_0 e^{-2\beta t} \tag{7.39}$$

where $E_0 = \frac{1}{2}k\mathcal{A}^2$ is the initial energy of the oscillator. Note that the decay time is $\tau = 1/2\beta$ half as big as the decay time for the oscillation the amplitude. An infinitesimal change of energy is

$$dE = -\frac{1}{\tau}E_0 e^{-t/\tau}dt = -\frac{E}{\tau}dt \tag{7.40}$$

If the energy loss per cycle, ΔE, is small, then we can integrate the above equation over one period of oscillation and obtain

---

[97] Another way to obtain this general solution is to take the limit $\beta \to \omega_0$ in expression (7.26). As a rule of thumb, when given a differential equation of order n, for which its characteristic equation has m repeated roots, α, the general solution has the form

$$x(t) = \sum_{j=1}^{n-1} A_j t^j e^{\alpha t}$$

[98] These type of solution are usually to model the shocks in cars.

[99] The average of a quantity X over a cycle is

$$< X >= \frac{1}{T}\int_0^T X(t)\; dt$$

$$\left(\frac{|\Delta E|}{E}\right)_{\text{cycle}} = \frac{T}{\tau} = \frac{2\pi}{Q} \tag{7.41}$$

where

$$Q = \omega_0 \tau = \frac{\omega_0}{2\beta} \tag{7.42}$$

is called the **quality factor** of the damped oscillator[100].

7.4 Driven Damped Harmonic Oscillator

In many situations, the oscillator interacts with the environment; which means it is under the effect of an external driving force. If $F(t)$ denotes the external force, then the equation of motion for the displacement $x(t)$ is given by the is a non-homogeneous second order differential equation

$$\ddot{x} + 2\beta\dot{x} + \omega_0^2 x = \frac{F(t)}{m} \tag{7.44}$$

The general solution can be written as

$$x(t) = x_h(t) + x_p(t) \tag{7.45}$$

where $x_{\mathrm{h}}(t)$ is the solution of the homogeneous equation, i.e. of a free damped harmonic oscillator, which after some transitory time it dies off, and $x_{\mathrm{p}}(t)$ is the particular solution to the equation above and it is also called the steady solution. We will examine two types of driving forces:

- **Step function**

$$F(t) = F_0\ \Theta(t) \tag{7.46}$$

where $\Theta(t)$ is the Heaviside step function, defined as

$$\Theta(t) = \begin{cases} 0, & t < 0 \\ 1, & t > 0 \end{cases} \tag{7.47}$$

In this case a particular solution to the equation of motion is a constant given by

$$x_p = \frac{F_0}{m\omega_0^2} \tag{7.48}$$

---

[100] In term of the quality factor Q, the expressions of the displacement for weak, strong, and critical damping correspond to

$$\begin{aligned} \text{Weak damping}: &\quad Q > 1/2 \\ \text{Strong damping}: &\quad Q < 1/2 \\ \text{Strong damping}: &\quad Q = 1/2 \end{aligned} \tag{7.43}$$

At this value of the position the force of the spring exactly balance the external force. Then, the most general solution is

$$x^{(\mathrm{WD})}(t) = \frac{F_0}{m\omega_0^2} + e^{-\beta t}\left(a\cos\omega_1 t + b\sin\omega_1 t\right) \tag{7.49}$$

$$x^{(\mathrm{SD})}(t) = \frac{F_0}{m\omega_0^2} + e^{-\beta t}\left[x_0(1+\beta t) + \dot{x}_0 t\right] \tag{7.50}$$

$$x^{(\mathrm{CD})}(t) = \frac{F_0}{m\omega_0^2} + e^{-\beta t}\left(a\cosh\sqrt{\beta^2-\omega_0^2}t + b\sinh\sqrt{\beta^2-\omega_0^2}t\right) \tag{7.51}$$

If we define a new coordinate

$$\tilde{x}(t) = x(t) - \frac{F_0}{m\omega_0^2} \tag{7.52}$$

the motion of the system corresponds to the oscillation of a damped oscillator around $x_0 = F_0/m\omega_0^2$. So, the effect of an external constant force amounts to a shift in the equilibrium position of the system.

- **Sinusoidal external force**:

  Let us consider $F(t)$ to be a cosine function of the form

$$F(t) = F_0\cos\omega t = \left[F_0 e^{i\omega t}\right] \tag{7.53}$$

We look for the solution $x_p(t)$ of the form

$$x_p(t) = \mathrm{Re}(\mathcal{A}e^{i\omega t}) \tag{7.54}$$

After substituting it into Eq(7.44), we find

$$\mathcal{A}(\omega) = \frac{F_0/m}{(\omega_0^2-\omega^2) + 2i\beta\omega} \tag{7.55}$$

Hence, the steady solution can be written as

$$\boxed{x_p(t) = \mathcal{A}_p(\omega)\cos\left(\omega t + \delta(\omega)\right)} \tag{7.56}$$

where

$$\mathcal{A}_p(\omega) = \frac{F_0/m}{\sqrt{(\omega_0^2-\omega^2) + 4\beta^2\omega^2}} \tag{7.57}$$
$$\delta(\omega) = -\arctan\left[\frac{2\beta\omega}{\omega_0^2-\omega^2}\right]$$

Therefore, the complete solution is obtained by adding the homogeneous solution to the particular solution given above.

In Fig., I have plotted the amplitude and the phase of the oscillation as function of the frequency of the driving force. The amplitude $\mathcal{A}(\omega)$ is extremum when its first derivative with respect to ω vanishes, i.e.

$$\frac{d\mathcal{A}(\omega)}{d\omega} := \frac{2\omega}{\mathcal{A}^3(\omega)}\left(\omega^2 - \omega_0^2 + 2\beta^2\right) = 0 \tag{7.58}$$

and so

$$\omega = 0, \qquad \text{or} \quad \omega = \sqrt{\omega_0^2 - 2\beta^2} \equiv \omega_R \tag{7.59}$$

If $\omega_0^2 > 2\beta^2$, then $\omega = \omega_R$ is a local maximum, where as $\omega = 0$ is a local minimum. Whereas if $\omega_0^2 < 2\beta^2$, then there is only a local maximum at $\omega = 0$. Thus, for a weakly damped oscillator, the amplitude of driven oscillator is maximal when $\omega \simeq \omega_0$. In this case we say that a **resonance** occurs with a resonance frequency $\omega_R \simeq \omega_0$.

The energy stored in this oscillator is

$$E(t) = \frac{1}{2}m\omega^2\mathcal{A}_s^2 \sin^2(\omega t + \delta) + \frac{1}{2}k\mathcal{A}_s^2\cos^2(\omega t + \delta) \tag{7.60}$$

where $k = m\omega_0^2$. Averaging over one period of oscillations, T, gives

$$< E >= \frac{1}{4}m\mathcal{A}_s^2(\omega^2 + \omega_0^2) \tag{7.61}$$

Substituting the expression of $\mathcal{A}_s$ given in Eq (7.57), yields

$$< E >= \frac{F_0^2}{4m}\frac{(\omega^2 + \omega_0^2)}{(\omega_0^2 - \omega^2) + 4\beta^2\omega^2} \tag{7.62}$$

Thus, at the resonance, the average energy over one cycle is

$$\boxed{< E >_R = \frac{F_0^2}{8m\beta^2}} \tag{7.63}$$

Note that in the limit where $\beta \to 0$, the amplitude $\mathcal{A}_s$ diverges at the resonance, which is unphysical. The reason is that in this limit the non-linear effects become important and the so the small amplitude approximation is violated.

The work done by the driving force over one cycle is

$$W_F = \int_0^T F\frac{dx}{dt}dt = \frac{\omega\mathcal{A}_s F_0 \sin\delta}{2}T \tag{7.64}$$

which represents the energy absorbed by the oscillator from the external force $F = F_0 \cos\omega t$ in one period of oscillation. The ratio of the energy absorbed to the energy stored in the oscillator at the resonance is given by

$$\boxed{\left(\frac{W_F}{<E>}\right)|_R = 2\pi\left(\frac{2\beta}{\omega_0}\right) = \frac{2\pi}{Q}} \tag{7.65}$$

- **Periodic driving force**

  For a periodic driving force with period T, we can decompose it into cosines and sines functions, i.e. as a Fourier series, as

$$F(t) = \frac{A_0}{2} + \sum_1^\infty A_n \cos n\omega t + \sum_1^\infty A_n \sin n\omega t \tag{7.66}$$

  Here $\omega = 2\pi/T$ is the principal frequency, and A_n and B_n are real constants. Since the differential equation of $x(t)$ is linear, the solution is a superposition of the form obtained in the previous subsection, i.e.

$$x(t) = \frac{A_0}{2m\omega_0^2} + \sum_1^\infty \frac{A_n \cos n\omega t + \theta_n + B_n \sin n\omega t + \theta_n}{m\sqrt{(\omega_0^2 - n^2\omega^2) + 4\beta^2 n^2 \omega^2}} \tag{7.67}$$

  where

$$\tan\theta_n = -\frac{2n\beta\omega}{\omega_0^2 - n^2\omega^2} \tag{7.68}$$

  Due to the presence of the harmonics $2\omega, 3\omega, ...$, it is now possible to have resonant motion whenever

$$\omega = \frac{\omega_0}{n}, \qquad n = 1, 2, ... \tag{7.69}$$

  The coefficients A_n and B_n are given by

$$A_n = \frac{2}{T}\int_{-T/2}^{T/2} F(t)\cos n\omega t \; dt, \qquad B_n = \frac{2}{T}\int_{-T/2}^{T/2} F(t)\sin n\omega t \; dt \tag{7.70}$$

  where I have used the identities

$$\int_{-T/2}^{T/2} \cos(n\omega t)\cos(m\omega t) \; dt = \begin{cases} T & : \; n = m = 0 \\ \frac{T}{2}\delta_{nm} & : \; n > 0; m > 0 \end{cases} \tag{7.71}$$

$$\int_{-T/2}^{T/2} \sin(n\omega t)\sin(m\omega t) \; dt = \begin{cases} 0 & : \; n = m = 0 \\ \frac{T}{2}\delta_{nm} & : \; n > 0; m > 0 \end{cases} \tag{7.72}$$

$$\int_{-T/2}^{T/2} \cos(n\omega t)\sin(m\omega t) \; dt = 0, \quad \forall n, m \in \mathbb{N} \tag{7.73}$$

7.5 General driving force: Green's function method

Green function is widely used in mathematics and physics, and it is defined as follows. Let be a linear differential equation of the form

$$\hat{\mathcal{L}}x(t) = g(t) \tag{7.74}$$

where $\hat{\mathcal{L}}_t$ is a differential operator with respect to a variable t, $g(t)$ is some function. Then, the Green function, $G(t,t')$, associated with equation satisfies

$$\hat{\mathcal{L}}_t G(t,t') = \delta(t-t') \tag{7.75}$$

In other words, the Green function describes the respond to sharp impulse centered around $t = t'$, described by the delta-distribution above. Now, the solution $x(t)$ can be written in terms of $G(t,t')$ as

$$x(t) = \int_{-\infty}^{+\infty} dt' \; g(t') \; G(t,t') \tag{7.76}$$

To show that this is indeed a solution, we substitute Eq. (7.76) into Eq. (7.74), and get

$$\hat{\mathcal{L}}_t x(t) = \int_{-\infty}^{+\infty} dt' \; g(t') \; \hat{\mathcal{L}}_t G(t,t') = \int_{-\infty}^{+\infty} dt' \; g(t') \; \delta(t-t') = g(t)$$

Intuitively, the linearity of the differential equation of $x(t)$ allows one to use the superposition principle to the solution of Eq.(7.74) as a the sum of all the responses to the impulses centered at $t = t'$ and weighted by $g(t')$

Let us apply this method to the harmonic oscillator system. In this case, the Green function is a solution to the differential equation

$$\left[\frac{d^2}{dt^2} + 2\beta\frac{d}{dt} + \omega_0^2\right] G(t,t') = \delta(t-t') \tag{7.77}$$

Before solving for $G(t,t')$ the boundary conditions. Before that the external force was applied, the system was at rest, and hence for $t < t'$ the Green function vanishes[101]. Moreover, $G(t,t')$ is continuous at $t = t'$, because if it was not the first term in the above equation will be the derivative of a delta distribution making it not compatible with the right hide side. Thus, $G(t',t')$ vanishes. However, the derivative of $G(t,t')$ is not continuous so that its derivative would give a delta-distribution.

If we integrate both sides of Eq. (7.77) the above equation from $t'_+ = t' - \epsilon$ to $t'_- = t' + \epsilon$, with ϵ being a real number close to zero, make use of the boundary conditions $G(t'_-,t') = 0$, $\dot{G}(t'_-,t') = 0$, and $G(t',t') = 0$, we get

$$\dot{G}(t'_+,t') + 2\beta G(t'_+,t') = 1 \tag{7.78}$$

We can further Taylor-expand the second term in ϵ and then take $\epsilon \to 0$ to obtain[102]

$$\dot{G}(t'_+,t') = 1 \tag{7.79}$$

which shows that the derivative of the Green function is discontinuous at $t = t'$, but finite.

Now, for $t > t'$, the delta-distribution vanishes and $G(t,t')$ satisfy the homogeneous equation, which we already derived the form of its solutions, but now instead of the variable t we substitute $(t-t')$ in the arguments. With the boundary condition (7.79) and $G(t',t') = 0$, the solution reads

[101] This is known as the **causality condition**, i.e. the system can not be affected at any time before the disturbance is applied to it.

[102] We made use of the fact that $G(t',t') = 0$.

$$G(t,t') = e^{-\beta(t-t')} \ \mathrm{Re}\left[\frac{e^{i\sqrt{\beta^2-\omega_0^2}(t-t')} - e^{-i\sqrt{\beta^2-\omega_0^2}(t-t')}}{2i\sqrt{\beta^2-\omega_0^2}}\right] \Theta(t-t') \tag{7.80}$$

Thus, according to E q.(7.76) gives

$$x(t) = \frac{1}{2m}\int_{\infty}^{+\infty} F(t')e^{-\beta(t-t')} \ \mathrm{Re}\left[\frac{e^{i\sqrt{\beta^2-\omega_0^2}(t-t')} - e^{-i\sqrt{\beta^2-\omega_0^2}(t-t')}}{2i\sqrt{\beta^2-\omega_0^2}}\right] dt' \tag{7.81}$$

Applying this to the three regimes of damped oscillator, we have

$$G^{(\mathrm{WD})}(t-t') = \frac{1}{\omega_1}e^{-\beta(t-t')}\sin\left[\omega_1(t-t')\right] \ \Theta(t-t') \tag{7.82}$$

$$G^{(\mathrm{SD})}(t-t') = \frac{1}{\sqrt{\beta^2-\omega_0^2}}e^{-\beta(t-t')}\sinh\left[\sqrt{\beta^2-\omega_0^2}(t-t')\right] \ \Theta(t-t') \tag{7.83}$$

$$G^{(\mathrm{CD})}(t-t') = e^{-\beta(t-t')}\left(t-t'\right) \ \Theta(t-t') \tag{7.84}$$

and so

$$x^{(\mathrm{WD})}(t) = \frac{1}{m\omega_1}\int_{-\infty}^{\infty} dt' \ F(t')e^{-\beta(t-t')}\sin\left[\omega_1(t-t')\right]$$

$$x^{(\mathrm{SD})}(t) = \frac{1}{m\sqrt{\beta^2-\omega_0^2}}\int_{-\infty}^{t} dt' \ F(t')e^{-\beta(t-t')}\sinh\left[\sqrt{\beta^2-\omega_0^2}(t-t')\right] \tag{7.85}$$

$$x^{(\mathrm{CD})}(t) = \frac{1}{m}\int_{-\infty}^{t} dt' \ F(t')e^{-\beta(t-t')}\sin\left[\omega_1(t-t')\right] \tag{7.86}$$

7.6 Systems with many degrees of freedom

Now , we would like to generalize the results obtained for the case of single particle to a system of N particles, described by n generalized coordinates $(q_1, ..., q_n) \equiv \mathbf{q}$ subject to a potential $U(\mathbf{q})$. When the system is in equilibrium we have

$$\mathcal{Q}_\alpha := -\nabla_\alpha U(\mathbf{q})\Big|_{\mathbf{q}_0} = 0 \tag{7.87}$$

where $\mathbf{q}^{(0)} \equiv (q_1^{(0)}, ..., q_n^{(0)})$ is the equilibrium point. The Lagrangian of the system reads

$$L = \frac{1}{2}\sum_{\alpha,\beta=1}^{n} T_{\alpha\beta}(\mathbf{q}) \ \dot{q}_\alpha\dot{q}_\beta - V(\mathbf{q}) \tag{7.88}$$

with

$$T_{\alpha\beta}(\mathbf{q}) = \sum_{i=1}^{3N} m_i\frac{\partial x_i}{\partial q_\alpha}\frac{\partial x_i}{\partial q_\beta} \tag{7.89}$$

Let us study the motion of the system around $\mathbf{q}_0$, and define

$$q_\alpha(t) = q_\alpha^{(0)} + \eta_\alpha(t) \tag{7.90}$$

If we assume that the fluctuation η_α to be small. compared to the equilibrium positions $q_a^{(0)}$, then we can Taylor expand both the kinetic and the potential energy terms about the equilibrium point $\mathbf{q}_0$, and get

$$L = \frac{1}{2}\sum_{\alpha,\beta=1}^{n}\Big[\mathcal{M}_{\alpha\beta}\ \dot\eta_\alpha\dot\eta_\beta - \mathcal{K}_{\alpha\beta}\ \eta_\alpha\eta_\beta\Big] + U(\mathbf{q}_0) + (\text{higher orders in } \eta) \tag{7.91}$$

where

$$\mathcal{M}_{\alpha\beta} = T_{\alpha\beta}(\mathbf{q}^{(0)}), \quad \mathcal{K}_{\alpha\beta} = \frac{\partial^2 V}{\partial q_\alpha \partial q_\beta}\Big|_{\mathbf{q}^{(0)}} \tag{7.92}$$

We will assume that the matrices $\hat{\mathcal{M}}$ and $\hat{\mathcal{K}}$, whose matrix elements are $\mathcal{M}_{\alpha\beta}$ and $\mathcal{K}_{\alpha\beta}$, respectively, are positive definite. Now the Euler-Lagrange equations for η_α read

$$\sum_{\beta=1}^{n}\Big[\mathcal{M}_{\alpha\beta}\ddot\eta_\beta + \mathcal{K}_{\alpha\beta}\eta_\beta\Big] = 0 \tag{7.93}$$

which represent a set of n second order coupled differential equations, with constant coefficients. We search for solutions of the form

$$\eta_\beta = \mathrm{Re}\{C\psi_\beta e^{i\omega t}\} \tag{7.94}$$

and after substituting into Eq.(7.93), we have

$$\sum_{\beta=1}^{n}\left(\mathcal{K}_{\alpha\beta} - \omega^2\mathcal{M}_{\alpha\beta}\right)\psi_\beta = 0 \tag{7.95}$$

or, equivalently, in matrix form as

$$\begin{bmatrix} \mathcal{K}_{11} - \omega^2\mathcal{M}_{11} & \ldots\ldots & \mathcal{K}_{1n} - \omega^2\mathcal{M}_{1n} \\ \vdots & \ddots & \vdots \\ \mathcal{K}_{n1} - \omega^2\mathcal{M}_{n1} & \ldots\ldots & \mathcal{K}_{nn} - \omega^2\mathcal{M}_{nn} \end{bmatrix} \begin{bmatrix} \psi_1 \\ \vdots \\ \psi_n \end{bmatrix} = 0 \tag{7.96}$$

For the system of equation to have non trivial solutions, we need the determinant of the matrix above to vanish:

$$\det\left(\hat{\mathcal{K}} - \omega^2\hat{\mathcal{M}}\right) = 0 \tag{7.97}$$

Since this is an n^{th} order polynomial in ω^2, there are n roots $\{\omega_1^2, \omega_2^2, \ldots\omega_n^2\}$, which are called the **normal frequencies** squared[103]. The condition (7.97) for finding eigenfrequencies is called the **characteristic equation**. Now for each ω_n corresponds an eigenvector

$$\psi^{(m)} = \left(\psi_1^{(n)}, \psi_2^{(n)}, \ldots., \psi_n^{(n)}\right)^{\mathrm{T}} \tag{7.98}$$

[103]The positivity of ω_i^2 is a consequence of the fact that both $\mathcal{K}$ and V are positive definite. This can be shown as follow. Multiply the left hand side of (7.95) by a_α and sum over α from 1 to n, yields

$$\omega_m^2 = \frac{\sum_{\alpha\beta}\mathcal{K}_{\alpha\beta}\psi_\alpha\psi_\beta}{\sum_{\alpha\beta}T_{\alpha\beta}\psi_\alpha\psi_\beta} \geq 0$$

that is a solution to the eigenvalue equation

$$\left(\hat{\mathcal{K}} - \omega_m^2 \hat{\mathcal{M}}\right) \psi^{(m)} = 0 \tag{7.99}$$

so that the n^{th} normal mode can be written as

$$\omega_m^2 = \frac{\sum_{\alpha,\beta} \mathcal{K}_{\alpha\beta} \psi_\alpha^{(m)} \psi_\beta^{(m)}}{\sum_{\alpha,\beta} \mathcal{M}_{\alpha\beta} \psi_\alpha^{(m)} \psi_\beta^{(m)}} \tag{7.100}$$

Let us consider two eigenvectors $\psi^{(m)}$ and $\psi^{(n)}$ with normal frequencies ω_m and ω_n, respectively. Then,

$$\left(\hat{\mathcal{K}} - \omega_m^2 \hat{\mathcal{M}}\right) \psi^{(m)} = 0 \tag{7.101}$$
$$\left(\hat{\mathcal{K}} - \omega_n^2 \hat{\mathcal{M}}\right) \psi^{(n)} = 0$$

Now multiply the first equation from the left by $\psi^{(n)^T}$ and second one by $\psi^{(m)^T}$, then by subtracting them the potential terms cancel, and we get

$$\left(\omega_n^2 - \omega_m^2\right) \left[\psi^{(i)}\right]^T \mathcal{M} \psi^{(m)} = 0$$

and hence

$$\left[\psi^{(i)}\right]^T \mathcal{M} \psi^{(j)} = 0, \qquad \text{if } \omega_n \neq \omega_m \tag{7.102}$$

Moreover, we make the the following rescaling

$$\psi^{(m)} \rightarrow \frac{\psi^{(m)}}{\sqrt{\left[\psi^{(m)}\right]^T \mathcal{M} \psi^{(m)}}}$$

Thus, for non degenerate eigenvalues we can always chose $\{\psi^{(m)}, m = 1, 2, ..n\}$ to satisfy the orthonormality relation By subtracting these equations we obtain

$$\left[\psi^{(i)}\right]^T \hat{\mathcal{M}} \psi^{(j)} = \delta_{ij} \tag{7.103}$$

In this case, the eigenvectors $\{\psi^{(m)}, m = 1, 2, ..n\}$ are linearly independent, forming an orthonormal basis, and hence the most general solution can always be be written as

$$\boldsymbol{\eta}(t) = \sum_{m=1}^{n} u^{(m)}(t)\ \psi^{(m)} \tag{7.104}$$

with the time dependent factors $u^{(m)}(t)$ represent the coordinates in this basis, and they are called the **normal coordinates**. By substituting the above expression into Eq.(7.95), and make use of Eq. (7.99), yields

$$\sum_{m=1}^{n} \left[\ddot{u}^{(m)}(t) + \omega_m^2 u^{(m)}(t)\right] \hat{\mathcal{M}} \psi^{(m)} = 0 \tag{7.105}$$

Multiplying by $\left[\psi^{(n)}\right]^T$ from the left and using the orthonormality property (7.103), gives

$$\ddot{u}^{(n)}(t) + \omega_m^2 u^{(n)}(t) = 0 \tag{7.106}$$

This is the equation of uncoupled simple harmonic oscillator,. Therefore, we have

$$\begin{aligned}\boldsymbol{\eta}(t) &= \sum_{m=1}^{n} A^{(m)} \sin\left(\omega_m t - \delta^{(m)}\right) \psi^{(m)} \\ &\equiv \sum_{m=1}^{n} A^{(m)} \Psi^{(m)}(t)\end{aligned} \tag{7.107}$$

The constant $A^{(m)}$ is the amplitude of the normal mode $\Psi^{(m)}(t)$ and $\delta^{(m)}$ its corresponding phase which can be determined from initial conditions.

Example 5.2: Double pendulum

In figure 7 we show a double pendulum. The position of the masses m_1 and m_2 are

$$\begin{aligned} x_1 &= l_1 \sin\theta_1, \quad y_1 = l_1 \cos\theta \\ x_2 &= l_1 \sin\theta_1 + l_2 \sin\theta_2, \quad y_2 = l_1 \cos\theta_1 + l_2 \cos\theta_2 \end{aligned} \tag{7.108}$$

Choosing the zero pf the potential energy at the plane $y = 0$, the gravitational potential energy of the double pendulum is

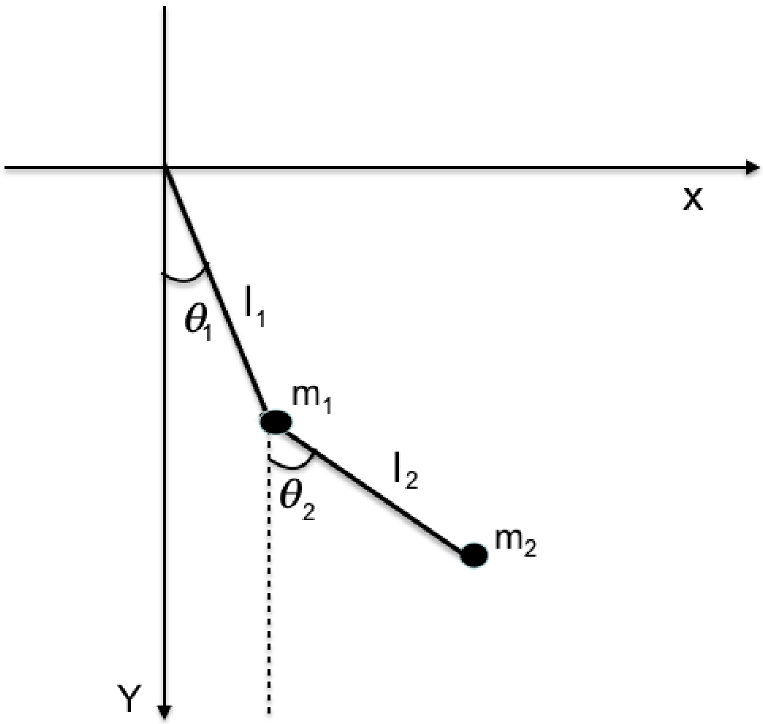


Figure 26: Double pendulum.

$$V = -m_1 g l_1 \cos\theta_1 - m_2 g \left[l_1 \cos\theta_1 + l_2 \cos\theta_2\right] \tag{7.109}$$

The kinetic energy of this system is just the sum of the kinetic energy of each mass, i.e.

$$
\begin{aligned}
T &= \frac{1}{2}m_1\left(\dot{x}_1^2+\dot{y}_1^2\right)+\frac{1}{2}m_2\left(\dot{x}_2^2+\dot{y}_2^2\right) \\
&= \frac{1}{2}\left(m_1+m_2\right)\dot{\theta}_1^2 l_1^2+\frac{1}{2}m_2\dot{\theta}_2^2 l_2^2+m_2\dot{\theta}_1\dot{\theta}_2 l_1 l_2\cos\left(\theta_1-\theta_2\right)
\end{aligned}
\tag{7.110}
$$

Thus, the Lagrangian of the system reads

$$
\begin{aligned}
L = &\frac{(m_1+m_2)}{2}\dot{\theta}_1^2 l_1^2+\frac{1}{2}m_2\dot{\theta}_2^2 l_2^2+m_2\dot{\theta}_1\dot{\theta}_2 l_1 l_2\cos\left(\theta_1-\theta_2\right) \\
&+\left(m_1+m_2\right)g l_1\cos\theta_1+m_2 g l_2\cos\theta_2
\end{aligned}
\tag{7.111}
$$

It follows that the Euler-Lagrange equations for θ_1 and θ_2 read

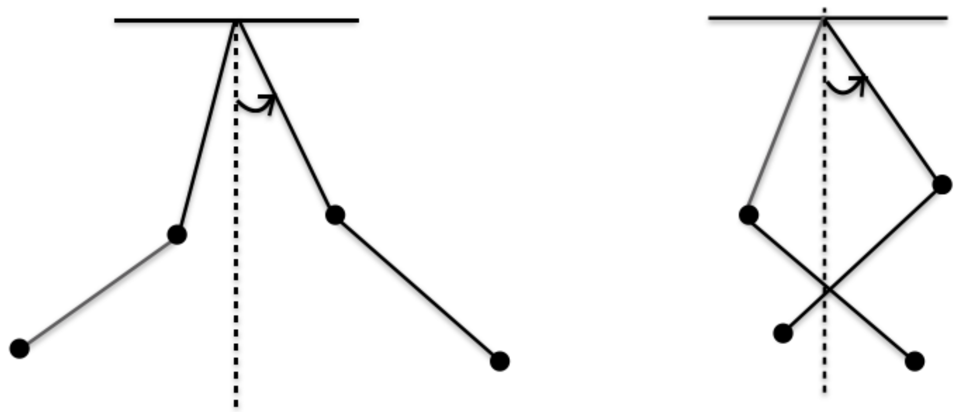
$$
\left(m_1+m_2\right)l_1\ddot{\theta}_1+m_2 l_2\ddot{\theta}_2\ \cos\left(\theta_1-\theta_2\right)+m_2 l_2\dot{\theta}_2^2\ \sin\left(\theta_1-\theta_2\right)+\left(m_1+m_2\right)g\ \sin\theta_1 = 0 \tag{7.112}
$$

$$
l_2\ddot{\theta}_2+l_1\ddot{\theta}_1\ \cos\left(\theta_1-\theta_2\right)-l_1\dot{\theta}_2^2\ \sin\left(\theta_1-\theta_2\right)+g\sin\theta_2 = 0 \tag{7.113}
$$

This nonlinear system of differential equations can not be solved analytically; instead one look for a graphical representation or numerical solutions. In fact, the system displays chaotic behavior. However, for small oscillations the above Lagrangian will have form[104]

$$
L=\frac{1}{2}[\dot{\theta}_1,\dot{\theta}_2]\hat{\mathcal{M}}\begin{bmatrix}\dot{\theta}_1\\ \dot{\theta}_2\end{bmatrix}+\frac{1}{2}[\theta_1,\theta_2]\hat{\mathcal{K}}\begin{bmatrix}\theta_1\\ \theta_2\end{bmatrix} \tag{7.114}
$$

where



(a) (b)

Figure 27: A sketch of the normal modes. Figures (a) and (b) correspond to oscillation in phase ($\Psi^{(-)}(t)$) and out-of phase ($\Psi^{(+)}(t)$), respectively.

[104] Note that for small oscillations, the system's degrees of freedom are coupled through the kinetic terms.

$$\hat{\mathcal{M}} = \begin{bmatrix} (m_1+m_2)l_1^2 & m_2 l_1 l_2 \\ \\ m_2 l_1 l_2 & m_2 l_2^2 \end{bmatrix}, \quad \hat{\mathcal{K}} = g \begin{bmatrix} (m_1+m_2)l_1 & 0 \\ \\ 0 & m_2 l_2 \end{bmatrix} \tag{7.115}$$

We are seeking nontrivial solutions for the equation of motion that are of the form $\mathrm{Re}\left(C\ e^{i\omega t}\right)$, and that is possible only if the frequencies satisfy the characteristic equation

$$\det\left(\hat{\mathcal{K}} - \omega^2 \hat{\mathcal{M}}\right) := \begin{vmatrix} \left(\omega^2 l_1^2 - g l_1\right)(m_1+m_2) & \omega^2 m_2 l_1 l_2 \\ \\ \omega^2 m_2 l_1 l_2 & \left(\omega^2 l_2^2 - g l_2\right) m_2 \end{vmatrix} = 0$$

which has the roots

$$\omega_\pm = \frac{g}{2 m_1 l_1 l_2}\left[(m_1+m_2)(l_1+l_2) \pm \sqrt{(m_1+m_2)\left[(m_1+m_2)(l_1+l_2)^2 - 4 m_1 l_1 l_2\right]}\right] \tag{7.116}$$

We can simplify greatly the above calculation by considering the special case where both masses and the lengths of the pendulum are equal, i.e $m_1 = m_2 = m$, and $l_1 = l_2$. In this case the frequencies are

$$\omega_\pm = \frac{g}{l}\left(2 \pm \sqrt{2}\right) \tag{7.117}$$

The corresponding normal modes, $\psi^{(\pm)}$, satisfy the equations

$$\hat{\mathcal{K}}\ \psi^{(+)} = \omega^2 \hat{\mathcal{M}}\ \psi^{(+)} \tag{7.118}$$

A straightforward calculation gives

$$\psi^{(\pm)} = \begin{bmatrix} 1 \\ \mp\sqrt{2} \end{bmatrix} \tag{7.119}$$

Hence,

$$\Psi^{(\pm)}(t) = \begin{bmatrix} 1 \\ \mp\sqrt{2} \end{bmatrix} \cos\left(\omega_\pm t + \phi_\pm\right) \tag{7.120}$$

In Fig. 27, we show a sketch of the normal oscillation modes for the double pendulum. For the mode $\Psi^{(-)}(t)$, the two pendulum oscillate in the same direction, where as for the mode $\Psi^{(+)}(t)$ oscillate in opposite direction.

Example 5.3: Vibrations of the carbon dioxide molecule

In this example we would like to find the normal modes of small vibration a carbon dioxide molecule. For this we model this system by considering the two oxygen atoms as two identical particles of mass m located symmetrically on either side of a carbon atom of mass M, and approximate the interatomic interactions between them by two springs of identical length l_0 and a force constant k, as shown in Fig.28. We will denote by $\{x_1^{(0)}, x_2\}$, and $x_3^{(0)}$ to be the positions of the two oxygen atoms and the carbon atom, respectively.

The kinetic and potential energy for this system are

$$T = \frac{m}{2}\left(\dot{x}_1^2 + \dot{x}_2^2\right) + \frac{M}{2}\dot{x}_3^2 \tag{7.121}$$

$$V = \frac{k}{2}(x_2 - x_1 - l_0)^2 + \frac{k}{2}(x_3 - x_2 - l_0)^2 \tag{7.122}$$

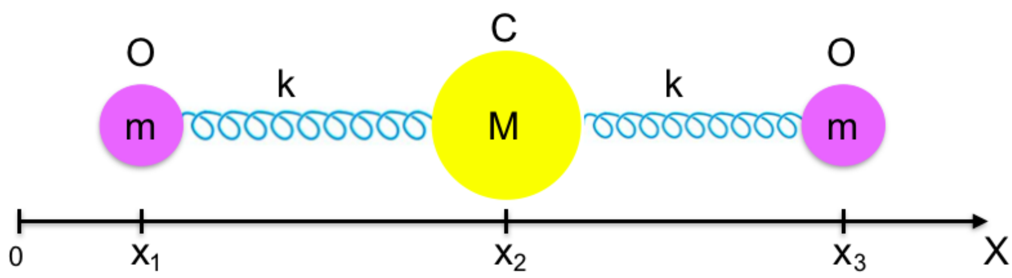


Figure 28: A model for a CO_2 molecule.

Now we introduce new coordinates $\eta_i = x_i - x_i^{(0)}$, with $i = 1, 2, 3$, which describe the motion of the atoms relative to their equilibrium positions. Then, we have

$$T = \frac{m}{2}\left(\dot{\eta}_1^2 + \dot{\eta}_2^2\right) + \frac{M}{2}\dot{x}\eta_3^2 \tag{7.123}$$

$$\tag{7.124}$$

$$V = \frac{k}{2}(\eta_2 - \eta_1)^2 + \frac{k}{2}(\eta_3 - \eta_2)^2 \tag{7.125}$$

where we have used the fact that

$$x_2^{(0)} - x_1^{(0)} = l_0, \qquad x_3^{(0)} - x_2^{(0)} = l_0$$

By expanding the square terms in Eq. (7.123), the Lagrangian can have the form

$$\mathcal{L} = \frac{1}{2}\sum_{\alpha,\beta=1}^{3}\Big[\mathcal{M}_{\alpha\beta}\ \dot{\eta}_a\dot{\eta}_b - \mathcal{K}_{\alpha\beta}\ \eta_a\eta_b\Big] \tag{7.126}$$

where the matrices $\mathcal{K}$ and $\mathcal{M}$ are given by

$$\hat{\mathcal{M}} = \begin{bmatrix} m & 0 & 0 \\ 0 & M & 0 \\ 0 & 0 & m \end{bmatrix}, \quad \hat{\mathcal{K}} = \begin{bmatrix} k & -k & 0 \\ -k & 2k & k \\ 0 & -k & k \end{bmatrix} \tag{7.127}$$

Now the normal modes of oscillation must be solutions of the characteristic equation

$$\begin{vmatrix} k - m\omega^2 & -k & 0 \\ -k & 2k - m\omega^2 & k \\ 0 & -k & k - m\omega^2 \end{vmatrix} = 0 \tag{7.128}$$

which after expanding the determinant can be put in the form

$$\omega^2\left(\omega^2 m - k^2\right)\left[\omega^2 Mm - k\,(M + 2m)\right] = 0 \tag{7.129}$$

So the normal frequencies are

$$\omega_1 = 0, \qquad \omega_2 = \sqrt{\frac{k}{m}}, \qquad \omega_3 = \sqrt{\frac{k}{m}\left(1+\frac{2m}{M}\right)} \tag{7.130}$$

The normal coordinates for the zero mode $\omega_1 = 0$ is a solution to the ordinary differential equation (see Eq. (7.106))

$$\ddot{u}^{(1)}(t) = 0 \tag{7.131}$$

That is, the three atoms either remains at rest or if it is moving it does so with uniform velocity to the right or to the left. The corresponding normal mode is a solution to the equation

$$\sum_\beta \hat{\mathcal{K}}_{\alpha\beta}\psi_\beta^{(1)} = 0 \quad \Longrightarrow \quad \begin{cases} \psi_1^{(1)}k - \psi_2^{(1)}k = 0, \\ \psi_1^{(1)}k + 2\psi_2^{(1)}k - \psi_3^{(1)}k = 0, \\ -\psi_2^{(1)}k - \psi_3^{(1)}k = 0 \end{cases} \tag{7.132}$$

and hence

$$\psi_1^{(1)} = \psi_2^{(1)} = \psi_3^{(1)} \quad \Longrightarrow \quad \psi^{(1)} = \psi_1^{(1)} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \tag{7.133}$$

For the normal modes corresponding to the eigenfrequencies ω_2 and ω_3, we have

$$\sum_\beta \left(\hat{\mathcal{K}} - \omega_2^2\hat{\mathcal{M}}\right)_{\alpha\beta}\psi_\beta^{(2)} = 0 \tag{7.134}$$

$$\sum_\beta \left(\hat{\mathcal{K}} - \omega_3^2\hat{\mathcal{M}}\right)_{\alpha\beta}\psi_\beta^{(3)} = 0 \tag{7.135}$$

which give

$$\psi_1^{(2)} = -\psi_3^{(2)}, \ \psi_2^{(2)} = 0 \quad \Longrightarrow \quad \psi^{(2)} = \psi_1^{(2)} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \tag{7.136}$$

$$\psi_1^{(3)} = -\psi_3^{(1)}, \quad \psi_2^{(3)} = -\frac{2}{M}\psi_1^{(3)} \quad \Longrightarrow \quad \psi^{(3)} = \psi_1^{(3)} \begin{bmatrix} 1 \\ -\frac{2m}{M} \\ 1 \end{bmatrix} \tag{7.137}$$

Normalizing $\psi^{(1)}, \psi^{(2)}$, and $\psi^{(3)}$ according to Eq. (7.103), yields

$$\psi^{(1)} = \frac{1}{\sqrt{M+2m}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \psi^{(2)} = \frac{1}{\sqrt{2m}} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \quad \psi^{(3)} = \frac{1}{\sqrt{2m\left(1+\frac{2m}{M}\right)}} \begin{bmatrix} 1 \\ -\frac{2m}{M} \\ 1 \end{bmatrix} \tag{7.138}$$

So, based on these results we have the following (see also Fig.):

- **For the zero mode (i.e. $\omega_1 = 0$)**: There is no oscillation. Either all three atoms remain at rest or they move as a whole in the same direction and with a constant velocity.

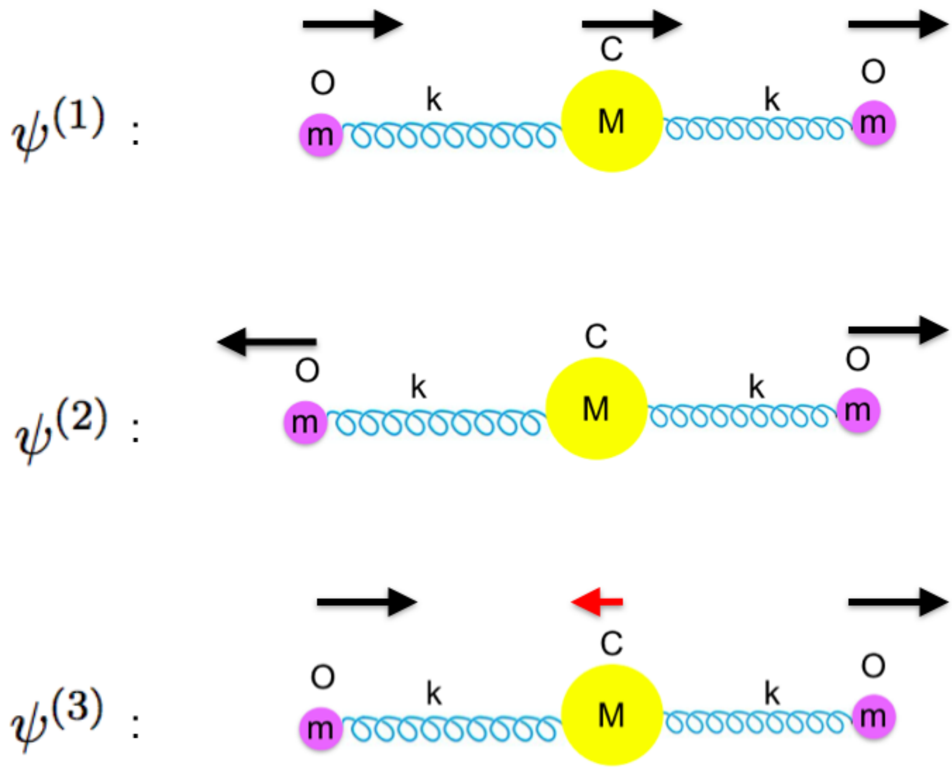


Figure 29: The normal oscillation mode of a CO_2 molecule.

- **For the mode** $\omega_2 = \sqrt{\frac{k}{m}}$: The two oxygen atoms move with equal amplitudes but in opposite directions while the carbon atom remains at rest.
- **For the mode** $\omega_3 = \sqrt{\frac{k}{m}\left(1 + \frac{2m}{M}\right)}$: The two oxygen atoms move with equal amplitude and in the same direction while the carbon atom with a fraction of their amplitude given by $2m/M$ opposite to their direction of motion.

Example 5.4: Three masses and four springs

Now let us consider three blocks with equal mass m attached with three identical springs with force constant k as shown in figure 30 below. If we denote by η_1, η_2, and η_3 the displacement of the masses relative to their equilibrium, the potential and kinetic energies for this system read

$$T = \frac{1}{2}m\left(\dot{\eta}_1^2 + \dot{\eta}_2^2 + \dot{\eta}_3^2\right) \tag{7.139}$$

$$V = \frac{1}{2}k\left(\eta_2 - \eta_1\right)\frac{1}{2}k\left(\eta_3 - \eta_2\right) + \frac{1}{2}k\eta_3^2 \tag{7.140}$$

and so the Lagrangian is

$$\mathcal{L} = \frac{1}{2}[\dot{\eta}_1, \dot{\eta}_2, \dot{\eta}_3]\hat{\mathcal{M}} \begin{bmatrix} \dot{\eta}_1 \\ \dot{\eta}_2 \\ \dot{\eta}_3 \end{bmatrix} + \frac{1}{2}[\eta_1, \eta_2, \eta_3]\hat{\mathcal{K}} \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{bmatrix} \tag{7.141}$$

where the matrices $\hat{\mathcal{M}}$ and $\hat{\mathcal{K}}$ are given by

$$\hat{\mathcal{M}} = \begin{bmatrix} m & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & m \end{bmatrix}, \quad \hat{\mathcal{K}} = \begin{bmatrix} 2k & -k & 0- \\ -K & -2k & -k \\ 0 & -k & 2k \end{bmatrix} \tag{7.142}$$

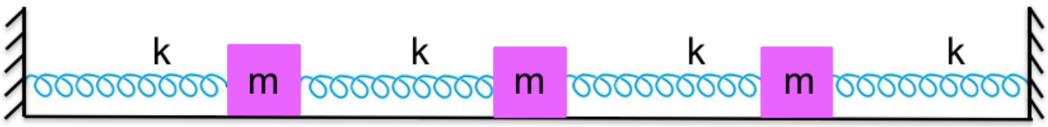


Figure 30: Three masses moving on a straight line under the action of four springs.

To find the normal frequencies of oscillation, we solve the characteristic equation

$$\begin{vmatrix} 2k - m\omega^2 & -k & 0 \\ -k & 2k - m\omega^2 & -k \\ 0 & -k & 2k - m\omega^2 \end{vmatrix} = 0 \tag{7.143}$$

Expanding the determinant, and factoring , we obtain

$$\left(2k - m\omega^2\right)\left[\left(2k - m\omega^2\right)^2 - 2k^2\right] = 0 \tag{7.144}$$

Hence, the normal frequencies are

$$\omega_1 = \sqrt{\frac{2k}{m}}, \quad \omega_2 = \sqrt{\frac{(2-\sqrt{2})\,k}{m}}, \quad \omega_2 = \sqrt{\frac{(2+\sqrt{2})\,k}{m}} \tag{7.145}$$

For the corresponding normal modes $\psi^{(\alpha)} = (\psi_1^{(\alpha)}, \psi_2^{(\alpha)}, \psi_3^{(\alpha)})$, with $\alpha = 1, 2, 3$ we solve the equations

$$\sum_{\beta=1}^{3}\left(\hat{\mathcal{K}} - \omega_1^2\hat{\mathcal{M}}\right)_{\alpha\beta}\psi_\beta^{(1)} = 0 \tag{7.146}$$

$$\sum_{\beta=1}^{3}\left(\hat{\mathcal{K}} - \omega_2^2\hat{\mathcal{M}}\right)_{\alpha\beta}\psi_\beta^{(2)} = 0 \tag{7.147}$$

$$\sum_{\beta=1}^{3}\left(\hat{\mathcal{K}} - \omega_3^2\hat{\mathcal{M}}\right)_{\alpha\beta}\psi_\beta^{(3)} = 0 \tag{7.148}$$

and a straightforward algebra gives

$$\psi^{(1)} = \psi_1^{(1)} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \quad \psi^{(2)} = \psi_1^{(2)} \begin{bmatrix} 1 \\ \sqrt{2} \\ 1 \end{bmatrix}, \quad \psi^{(3)} = \psi_1^{(3)} \begin{bmatrix} 1 \\ -\sqrt{2} \\ 1 \end{bmatrix} \tag{7.149}$$

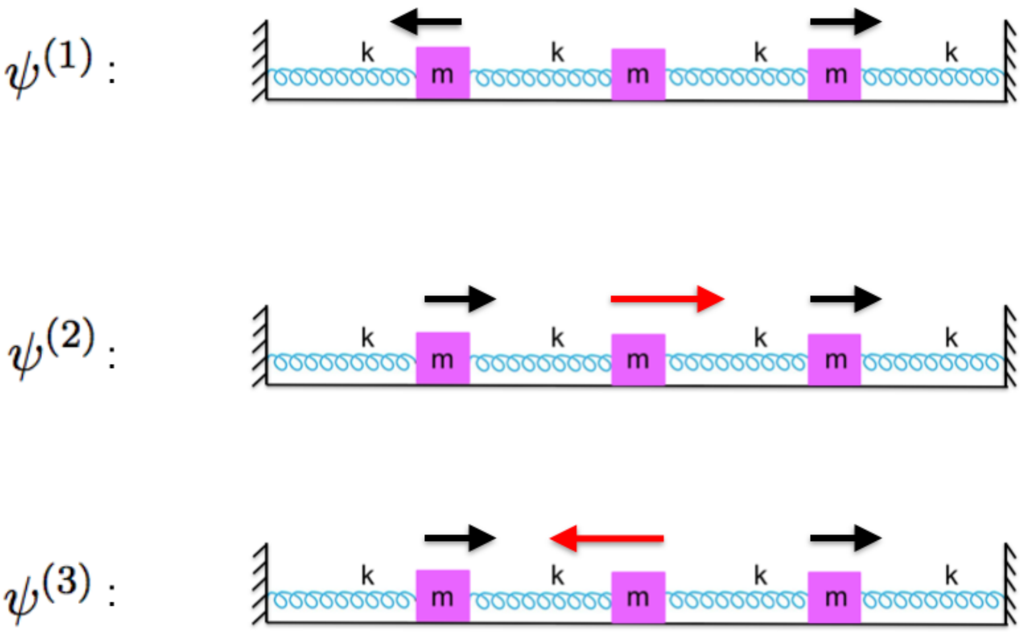


Figure 31: The normal oscillation mode of the three masses-four springs system.

Therefore, in normal mode, the masses oscillate as follows (see Fig.**??**)

- **For the zero mode** $\omega_1 = \sqrt{\frac{2k}{m}}$: The two masses at the opposite ends move with equal amplitude but opposite direction, while the mass in between remains at rest.
- **For the mode** $\omega_2 = \sqrt{\frac{(2-\sqrt{2})k}{m}}$: The two masses at the opposite ends move with equal amplitude and same direction (i.e. with the same phase), while the mass in between move also in the same direction but with an amplitude a factor of $\sqrt{2}$ larger.
- **For the mode** $\omega_3 = \sqrt{\frac{(2+\sqrt{2})k}{m}}$: The two masses at the opposite ends move with equal amplitude and same direction (i.e. with the same phase), while the mass in between move in the opposite direction but with an amplitude a factor of $\sqrt{2}$ larger.

Example 5.5: Transverse oscillations of two masses on a string

7.7 One dimensional lattice vibrations

8 The Hamiltonian Formalism

9 Nonlinear Dynamics and Chaos

- In 1666 Newton showed that when a planet is being pulled by a star's gravity, it moves in an elliptical orbit. We already discussed this in some detail in chapter 2. But, in reality the motion of the planet can be affected by the presence of other planets around and in this case one has to consider solving the many body problem. In fact, Newton's tried to solve the of three body problem of the Earth, moon, and the Sun, but he could not[105]. Many great minds after Newton, such as Euler, Gauss, also tried solving the three body problem but did not succeed. It was until the end of 19$^{\text{th}}$ century when **Henri Poincare** showed why it was not possible to solve it.

[105] He actually wrote a letter to one of his friends saying that no problem got headache as the three body problem.

10 Exam Problems

11 Answers to Exam Problems

12 Appendices

12.1 The First Measurement of the Universal Constant of Gravitation

In 1798, Henri Cavendish performed an extremely delicate experiment that allowed the measurement of the numerical value of the constant G_N in the expression of the gravitational force between two masses [106]. The setup is as follows:

A dumbbell with two masses m on its ends hang from a very thin wire. The dumbbell is free to twist, although if it twists, the wire will provide a tiny restoring torque. Two other masses M are placed at the positions shown in the figure below. Theses masses produce attractive forces on the two masses m and cause the dumbbell to twist counterclockwise. The dumbbell will oscillate back and forth before finally settling down at some tiny angle θ away from the initial position.

For small angles θ, the torque on the dumbbell that arises from the twist takes the form

$$\tau = -b\theta \tag{12.1}$$

where b is a constant that depends on the thickness and the material from which the wire is made of. The gravitational force between each pair of masses is

$$F = G_N \frac{Mm}{d^2} \tag{12.2}$$

where d is the separation between the centers of the masses in each pair. So the torque on the dumbbell due to the two gravitational forces is $2G_N Mm/d^2 l$, where l is the half length of the dumbbell. Since at equilibrium the total torque is zero, we get

$$G_N = \frac{b\ \theta_* d^2}{2Mml} \tag{12.3}$$

where θ_* is the angular position of the dumbbell at equilibrium. This angle of rotation is measured by the deflection of light beam reflected from a mirror attached to the vertical suspension. The deflection of the light beam a technique for amplifying the motion. All the parameters on the right hand side except b. Measuring b is very difficult with any reasonable accuracy, because the torque in the wire is so tiny. But we know that the rotational analogue of Newton's second law is

$$-b\ \theta = I\ddot{\theta} \tag{12.4}$$

This is just the equation of harmonic oscillator with frequency $\omega = \sqrt{b}I$. Therefore, if we measure the period, T, of the oscillation we can determine the constant b from the relation

$$b = I\left(\frac{2\pi}{T}\right) \tag{12.5}$$

Plugging this expression of b into (**??**GN) gives

$$\boxed{G_N = \frac{4\pi^2 I\theta_* d^2}{4MmlT^2}} \tag{12.6}$$

[106] Actually, Cavendish's aim was not to measure the gravitational constant, but rather to measure the Earth's density relative to water, through the precise knowledge of the gravitational interaction.

The measured value of the universal gravitational constant is

$$\boxed{G_N = 6.67 \times 10^{-11} N.m^2/kg^2} \tag{12.7}$$

12.2 Gauss's Law of Gravitation

12.3 Bertrand's Theorem

12.4 Conic Sections

12.5 Multipole Expansion of Gravitational Potential of a Mass Distribution

12.6 Finding Green Function Using Fourier Transform

13 Who is Who

13.1 Galileo Galilei (1564- 1642; Italian).

Figure 32: Leonhard E uler.

13.2 Johannes Kepler (1571-1630; German).

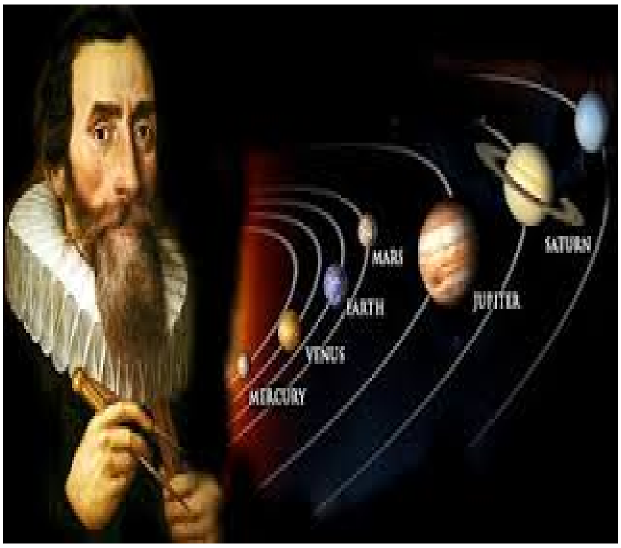


Figure 33: Leonhard E uler.

13.3 Isaac Newton (1642 -1727; British).

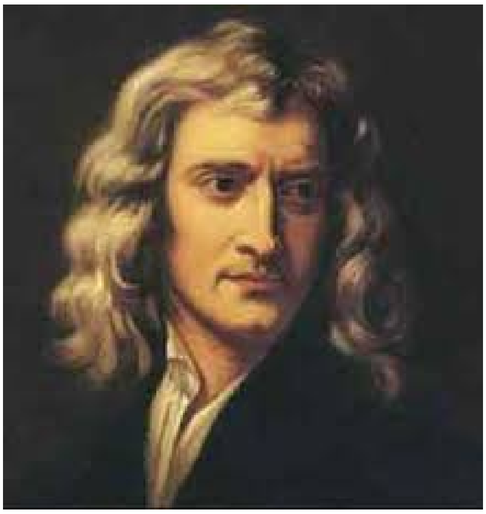


Figure 34: Isaac Newton.

13.4 Leanhard Euler (1707-1783; Swiss).

Figure 35: Leonhard Euler.

13.5 Jean le Rond D'alembert (1717-1783; French).

Figure 36: Jean D'alembert.

13.6 Joseph-Louis Lagrange (1736-1813; French).

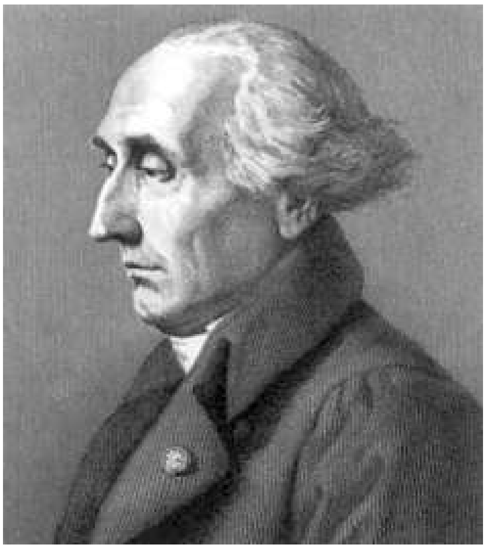


Figure 37: Joseph Lagrange.

13.7 William Rowan Hamilton (1805-1865; Irish).

Figure 38: William Hamilton.

14 References

References

[1] **P. G. Roll, R. Krotov and R. H. Dicke**. "The Equivalence of Inertial and Passive Gravitational Mass". Published in **Ann. Phys. (NY) 26, 442 (1964)**.

[2] **T. Wagner, S. Schlamminger, J. Gundlach, and E. Adelberger**. Published in the Journal **Classical and Quantum Gravity, 29, 184002 (2012)**.

[3] **S. Dimopoulos, P. W. Graham, J. M. Hogan, and M. A. Kasevich**. Published in the Journal **Phys. Rev. Lett. 98, 111102 (2007)**.

[4] **CERN Courier** *http://cerncourier.com/cws/article/cern/38299*.

[5] **NASA website**. *http://www.nasa-klass.com/*.

[6] **V. I. Arnold**: "Mathematical Methods of Classical Mechanics". Published by **Springer, 2nd edition (1997)**.

[7] **S. Ray and J. Shamanna**. Published in the Journal **European Journal of Physics, Volume 27, Number 2, (2006)**.

[8] **D. T. Greenwood**: "Classical Dynamics". Published by **Dover Publications; New edition (1997)**.

[9] **D. Morin**: "Introduction to Classical Mechanics". Published by **Cambridge University, 1st edition (2008)**.

[10] **A. Arya**: "Introduction to Classical Mechanics". Published by **Pearson; 2nd edition (1997))**.

[11] **W. Greiner**. "Classical Mechanics: System of Particles and Hamiltonian Dynamics". Published by **Springer; 2nd ed. 2010 edition (2009)**.

[12] **T. Morion**. "Classical Dynamics of Particles and Systems". Published by **Cengage Learning, 5th edition (2003)**.

[13] **G. Fowles and G. Cassiday**. "Analytical Mechanics". Published by **Thomson Brooks/Cole, 7th edition (2005)**.

[14] **A. Assis**. "Relational Mechanics and Implementation of Mach's Principle with Weber's Gravitational Force". Published by **Apeiron, (2014)**.

[15] **D. Tong**. "Lecture Notes on Classical Dynamics". **Course given for final year undergraduates at the University of Cambridge (2004)**.

[16] **C. Torre**. "Lecture Notes on Intermediate Classical Mechanics". **Course given at Utah State University, Fall 2012**.

[17] **B. Svistunov**. " Notes on Two-Body Problem. Central Potential. 1 D Motion". **Part of a course of "Mechanics I", given at UMassAmherst, Fall 2008**.

[18] **K. T. McDonald**. "Lecture notes on Mechanics". **Course Phys 205, given at Princeton University, (1988)**.

[19] **A. Cherkaev, and E. Cherkaev**. "Calculus of Variations and Applications". **Lecture notes (2003)**.

[20] **R. Kidd and S. Fogg**. "A Simple Formula for the Large-angle Pendulum Period". Published in the Journal of **Phys. Teach. 40 (2), 81-83 (2002)**.

[21] **L. Millet**. "The Large-angle Pendulum Period". Published in the Journal of **Phys. Teach. 41 (3), 162-163 (2003)**.

[22] **G. Hite**. "Approximations for the Period of a Simple Pendulum". Published in the Journal of **Phys. Teach. 43 (5), 290-292 (2005)**.

[23] **A. Belendez, and A. Hernandez, A. Marquez, T. Belendez, and C. Neipp**. "Analytical Approximations for the Period of a Nonlinear Pendulum". Published in the Journal of **European Journal of Physics, Vol. 27, 539-551 (2006)**.

[24] **F. Lima and P. Arun**. "An Accurate Formula for the Period of a Simple Pendulum Oscillating Beyond the Small-angle Regime". Published in the Journal of **Am. J. Phys. 74, 892-895 (2006)**.

[25] **A. Belendez, J. Rodes, T. Belendez, and A. Hernandez**. "Reply to Comment on "Approximation for the Large-angle Simple Pendulum Period". Published in the Journal of **European Journal of Physics, Vol. 30, 25 - 28 (2009)**.

[26] **S. Nasri**. " How the Sun Shines". Posted on http://faculty.uaeu.ac.ae/snasri/Notes/My-How-Sun-Radiates-v3-2016.pdf.