Classical Mechanics

Haydn Cheng

June 8, 2025

Contents

1 Newtonian mechanics

1.1 Point Particle

The translation Newton's second law the net force on a system $\mathbf{F}_{\text{net}}$ with its mass m and its center of mass velcoity $\mathbf{v}_{\text{c.m.}}$ by

$$\mathbf{F}_{\text{net}} = \frac{d\mathbf{p}_{\text{c.m.}}}{dt} = \frac{d(m\mathbf{v}_{\text{c.m.}})}{dt}. \tag{1.1}$$

This equation works completely fine when the mass m varies with time. However, the center of mass velocity is hard to find in such cases. A more useful version is the modified Newton's second law

$$\mathbf{F}_{\text{net}} = m\mathbf{a} + \mathbf{v}_{\text{rel}}\frac{dm}{dt}, \tag{1.2}$$

where $\mathbf{a}$ is the acceleration of m and $\mathbf{v}_{\text{rel}}$ is the velocity of the mass relative to the added mass.

This equation is simply an application of the impulse-momentum theorem of a system experiencing a net force $\mathbf{F}_{\text{net}}$ in a time period dt consisting of a mass m moving at velocity $\mathbf{v}$ initially and an added mass dm moving at velocity $\mathbf{u}$ initially.

$$\mathbf{F}_{\text{net}}dt = d\mathbf{p} = (m+dm)(\mathbf{v}+d\mathbf{v}) - dm\mathbf{u} - m\mathbf{v} \implies \mathbf{F}_{\text{net}} = m\mathbf{a} + \frac{dm}{dt}(\mathbf{v}-\mathbf{u}). \tag{1.3}$$

We see that the net force $\mathbf{F}_{\text{net}}$ is responsible for both the usual acceleration $\mathbf{a}$ of the mass m and the acceleration of the added mass dm from its original velocity $\mathbf{u}$ to the mass's velocity $\mathbf{v}$ in time dt

For cases where the mass is decreasing, we have, if the removed mass is moving at velocity $\mathbf{u}$ afterwards,

$$\mathbf{F}_{\text{net}}dt = d\mathbf{p} = (m+dm)(\mathbf{v}+d\mathbf{v}) + (-dm)\mathbf{u} - m\mathbf{v} \implies \mathbf{F}_{\text{net}} = m\mathbf{a} + \frac{dm}{dt}(\mathbf{v}-\mathbf{u}), \tag{1.4}$$

which is the same as the equation for adding mass, just that $\mathbf{v}_{\text{rel}}$ is now the velocity of the mass relative to the removed mass. Here the negative sign in $-dm$ is due to the fact that m is defined as the mass of the massive object, so dm is inherently negative if mass is being removed.

reduced mass = work in cm frame?

Example: Falling Chain off a Table (1).

Question: An uniform incompressible, inextendable and stretched chain of length L and mass m is stretched out on a frictionless horizontal table with part of its length h hanging down through a hole in the table. Find the time it take for the chain to fall off.

Solution: The key words here are incompressible, inextendable and stretched which dictates the chain move together at a constant velocity. If we consider the chain as a whole, there are three forces acting on it: gravity, normal reaction by the table and normal reaction at the corner. The normal reaction must arises to provide horizontal acceleration for the center of mass of the whole system. If we consider the infinietesimal element of chain at the corner, we can conclude that $N = \sqrt{2}T$. Since there are three unknown force, we know that considering the chain along is not sufficient to determine its equations of motion. Instead, we need to divide the system into the part on the table and the part that has fallen off the table (and the infinitesimal chain at the corner which is used to redirect the direction of tension at that point).
Let x be the fraction of the chain that has fallen off the table. The modified Newton's second law read

$$T = (1-x)m\ddot{x}L \quad \text{and} \quad xmg - T = xm\ddot{x}L \implies xmg = mL\ddot{x}. \tag{1.5}$$

Note that it does not need the additional term due to change in mass in the two systems since $\mathbf{v}_{\text{rel}} = 0$.
Solving, we get $t = \sqrt{L/g}\cosh^{-1}(L/h)$.
In fact, using the concept of generalized coordinate (in this case, along the chain) we can obtain the equation of motion in one line.

Example: Falling Chain(2).

Question: Solve the previous problem but the chain is slack initially and is coiled around a hole in a table.

Solution: The slack part of the chain can have no tension – this is the definition of being slack. However, when we consider the part of the chain on the table, we see that it experience a force since its mass is decreasing and the relative velocity between the removed mass and the mass itself is non-zero, how is that possible?
The velocities of different parts of the chain are now moving at different velocities, which is a promising sign to use the modified Newton's second law. Consider part of the chain that has fallen off the table, the modified Newton's second law reads

$$xmg = xma + \frac{dm}{dt}v = xma + \frac{mv^2}{L}. \tag{1.6}$$

This is a non-linaer first order differential equation, which can be solved by guessing $x = At^n$. This is indeed the solution and solving for A and n gives $t = \sqrt{3L/g}$.

Example: Falling Chain(3).

Question: An uniform chain of length L and mass m stetched vertically just above the surface of a weighing scale and then release from rest. Find the reading of the scale as the chain falls onto the scale.

Solution: Since the chain is slack at the scale, the tension is zero at the bottom of the falling chain. The falling chain essentially undergoes free fall and the increase in normal reaction is simply provided only to the infinitesimal element of the chain that comes to stop when it hit the scale.
Let x be the length of the falling chain, measured from the scale up to the positive vertical direction. The speed of the falling chain can be found via energy conservation, which gives

$$\frac{dx}{dt} = -\sqrt{2g(l-x)}. \tag{1.7}$$

There are a few systems we can consider:

1. The infinitesimal element that comes to stop after hitting the scale,
2. the falling chain
3. the chain on the scale that is at rest, and
4. the whole chain.

Firstly, we consider the infinitesimal element that comes to stop after hitting the scale. Having expereinces from previous examples, we see that this is the most natural choice as this is what the normal reaction is actually acting on and the normal reaction is dedicated entirely and only to stop this element of the chain from penetrating the scale. The Newton's second law reads

$$N - gdm = dma \implies N = \frac{mdx}{L}\frac{dx/dt}{dt} = \frac{m}{L}\left(\frac{dx}{dt}\right)^2 = 2\left(\frac{l-x}{l}\right)mg. \tag{1.8}$$

Adding N with $((l-x)/l)mg$, which is the contribution of the stationary chain to the reading of the scale gives $N_{\text{tot}} = 3\left((l-x)/l\right)mg$.
Secondly, we can consider the modified Newton's second law of the falling chain

$$N - \frac{x}{l}mg = \frac{x}{l}m(-g) + \left(-\frac{dx}{dt}\right)\frac{d}{dt}\left(\frac{l-x}{l}m\right) \implies N = 2\left(\frac{l-x}{l}\right)mg. \tag{1.9}$$

Adding N with $((l-x)/l)mg$, which is the contribution of the stationary chain to the reading of the scale gives $N_{\text{tot}} = 3\left((l-x)/l\right)mg$.
Thirdly, we can consider the part of the cahin on the scale that is at rest, but this is basically equivalent to consider the infinitesimal element in the first method, since the stationary chain does not add anything special.

At last, we can consider the whole chain. The center of mass of the chain is located at

$$x_{\text{c.m.}} = \frac{1}{m}\left(\frac{l-x}{l}m(0) + \frac{x}{l}m\frac{x}{2}\right) = \frac{x^2}{2l}. \tag{1.10}$$

The center of mass velocity is therefore

$$\frac{dx_{\text{c.m.}}}{dt} = \frac{x}{l}\left(\frac{dx}{dt}\right) = -\frac{x}{l}\sqrt{2g(l-x)}. \tag{1.11}$$

The center of mass acceleration is therefore

$$\frac{d^2x_{\text{c.m.}}}{dt^2} = \frac{1}{l}\left(\frac{d^2x}{dt^2}\right)^2 + \frac{x}{l}\frac{d^2x}{dt^2} = \left(\frac{l-x}{l}\right)2g + \frac{x}{l}g. \tag{1.12}$$

The unmodified Newton's second law therefore gives

$$N_{\text{tot}} = m\frac{d^2x_{\text{c.m.}}}{dt^2} = 3\left(\frac{l-x}{l}\right)mg. \tag{1.13}$$

Example: Cart with Time-Varing Mass (1).

Question: Consider a cart with mass m moving at a constant velocity v under a force F. In time period dt an infinitesimal mass dm is fell on the cart vertically. What is the equation of motion of the cart?

Solution: Consider the cart as a system, we have from the modified Newton's second law

$$F = ma + \frac{dm}{dt}v = \frac{dm}{dt}v. \tag{1.14}$$

On the other hand, we can consider the added infinitesimal mass as a system, then we have

$$F = dma = dm\frac{v}{dt} = \frac{dm}{dt}v. \tag{1.15}$$

As we have found, the force F is not contributing to the usual acceleration $\mathbf{a}$ of the mass m but is entirely entitiled to the acceleration of the infinitesimal mass dm from rest to velocity $\mathbf{v}$.

Example: Cart with Time-Varing Mass (2).

Question: Consider a cart with mass m moving at an acceleartion a under a force F. In time period dt an infinitesimal mass dm is leaked out of the cart vertically. What is the equation of motion of the cart?

Solution: Consider the cart as a system, we ahve from the modified Newton's second law

$$F = ma + \frac{dm}{dt}v_{\mathrm{rel}} = ma. \tag{1.16}$$

As we can see, since there is no need to accelerate or decelerate the removed mass, it can be gone without using any part of the force F, so F is dedicated fully to accelearte the mass m.

Example: Pulling Carpet.

Question: A long, thin, pliable carpet of mass m and lenght l is laid on the floor. One end of the carpet is bent back and then pulled backwards with constant velocity v, just above the part of the carpet which is still at rest on the floor. What is the minimum force F needed to pull the moving part,?

Solution: Again, since different parts of the carpet is moving at different speed, it would be hard to find the
The modified Newton's second law of the moving part reads

$$F = ma + \frac{dm}{dt}v = v\frac{dm}{dt} = \frac{mv}{L}\frac{dx}{dt} = \frac{mv}{L}\frac{\frac{L}{2}}{\frac{L}{v}} = \frac{mv^2}{L}, \tag{1.17}$$

where we have taken the instant when the top and the bottom of the carpet overlap with each other to calculate dm/dt.
In retrospect, we can simply consider how much mass is accelerated from rest to v by force F in time dt since this is the only thing the force is responsible for when the mass is at constant velocity.

Example: Rocket Equation.

Question: Find the speed of a rocket with initial velocity v_0, initial mass m_0 and final mass m, ejecting fuel at velocity v_{rel}. Gravity can be neglected.

Solution: The modified Newton's second law of the rocket reads

$$F = 0 = m\frac{dv}{dt} + \frac{dm}{dt}v_{\mathrm{rel}} \implies v = v_0 + v_{\mathrm{rel}} \ln\left(\frac{m_0}{m}\right). \tag{1.18}$$

Example: Air Cannon.

Question: Consider a cylindrical tube with cross sectional area A with sealed end at one end and a piston of mass m at the other end. The piston is held stationary and the tube contain no air initially. Find the position x of the piston along the tube if the density and pressure of the air are ρ and P_0.

Solution: Since air is incompressible, all air move at the same speed $\dot{x}$ as the piston. The unmodified Newton's second law of the air inside the tube reads

$$P_0 A = \frac{d}{dt}((m+xA\rho)\dot{x}) \implies P_0 At = (m+xA\rho)\dot{x} \implies x = \frac{m}{\rho A}\left(\sqrt{1+\frac{P_0\rho A^2t^2}{m^2}}-1\right). \tag{1.19}$$

Example: Interstellar Travel.

Question: Consider a probe with mass m_0 and cross sectional surface area S to be travelling in interstellar space with initial velocity u and enters a uniform stationary dust cloud of density ρ. Assuming that all the dust in the path of the satellite sticks to the forward facing surface, obtain expressions for the spped of the satellite into the cloud at a later time t.

Solution: Using the modified Newton's second law, we have

$$F = 0 = mdv + vdm = \frac{v}{u}m_0dv + v(\rho vSdt) \implies v(t) = u/\sqrt{1+2\rho Sut/m_0}. \tag{1.20}$$

To account for the rotational motion of an object, torque $\boldsymbol{\tau}$ and angular momentum $\mathbf{L}$[1] are introduced, defined by

$$\boldsymbol{\tau} = \mathbf{r}\times\mathbf{F} \quad \text{and} \quad \mathbf{L} = \mathbf{r}\times\mathbf{p}. \tag{1.21}$$

We can see that both $\boldsymbol{\tau}$ and $\mathbf{L}$ depend on the origin defined as $\mathbf{r}$ is the position vector.

Taking the derivatives of the angular momentum, we yield "Newton's second law for rotation"

$$\frac{d\mathbf{L}}{dt} = \frac{d}{dt}(\mathbf{r}\times\mathbf{p}) = \mathbf{r}\times\frac{d\mathbf{p}}{dt} + \frac{d\mathbf{r}}{dt}\times\mathbf{p} = \mathbf{r}\times\mathbf{F} + \mathbf{v}\times(m\mathbf{v}) = \mathbf{r}\times\mathbf{F} = \boldsymbol{\tau} \tag{1.22}$$

Another quantity that is introduced to simplify the matter (which ultimately comes from the symmetry of time) is kinetic energy $\mathbf{T}$ and potential energy U defined by

[1] One may question the necessity to introduce the concept of torque and angular momentum. Indeed, with Newton's second law, one can virtually solve all mechanics problems without resorting to other physical laws. However, when analyzing rigid bodies with spatial extent (in contrast to a point particle), torque becomes useful because the internal forces in these bodies are generally very complicated. In fact, $\boldsymbol{\tau} = d\mathbf{L}/dt$ is merely an extension of Newton's second law as explained and derived here. With a different viewpoint, Noether's theorem dictates that since the universe is rotationally symmetric, so $\mathbf{L}$ must be conserved, in some sense $\mathbf{L}$ is just some useful conserved quantity that is a consequence of a certain symmetry just like how the Laplace-Runge-Lenz vector is the repercussion of some hidden symmetry in higher dimensions.

$$T = \frac{1}{2}mv^2 \text{ and } \mathbf{F} = -\boldsymbol{\nabla} V \tag{1.23}$$

This definition is motivated when considering the work done on a constant mass m by a net force $\mathbf{F}_{net}$

$$W = \int_1^2 \mathbf{F_{net}} \cdot d\mathbf{r} = \int_1^2 \frac{dm\mathbf{v}}{dt} \cdot d\mathbf{r} = m \int_1^2 d\dot{\mathbf{r}} \cdot \frac{d\mathbf{r}}{dt} = \frac{m}{2} \int_1^2 d(v^2) = \frac{1}{2}mv_2^2 - \frac{1}{2}mv_1^2 \tag{1.24}$$

so that we can say the work done by the net force is equal to the change in the kinetic energy (also known as the work-energy theorem)

$$W = W_{con} + W_{non-con} = \Delta T.^2 \tag{1.25}$$

If we define $\mathbf{F}_{net} = \mathbf{F}_{con} + \mathbf{F}_{non-con}$ which is comprised of both conservative forces such as gravity where $\mathbf{F} \propto \hat{\mathbf{r}}/r^2$ as well as non-conservative forces such as friction.

The defining properties of conservative forces are:

1. $\oint \mathbf{F} \cdot d\mathbf{r} = 0$, or equivalently, from the Stokes' theorem,
2. $\boldsymbol{\nabla} \times \mathbf{F} = 0$, which both imply
3. The work done by conservative forces is independent of the path taken, as if the work done by the conservative force from point 1 to 2 is a constant value and by switching the sign of $d\mathbf{r}$ in $W = \int_1^2 \mathbf{F} \cdot d\mathbf{r}$, we see that the work done from point 2 to 1 adds a negative sign to that constant value and thus the work done of a loop is zero which is equivalent to $\oint \mathbf{F} \cdot d\mathbf{r} = 0$.

Thus from the vector identity $\boldsymbol{\nabla} \times (\boldsymbol{\nabla} V) = 0$ and the second item above ($\boldsymbol{\nabla} \times \mathbf{F} = 0$), we can define the potential energy as mentioned and the work-energy theorem (eq. (1.25)) becomes the conservation of energy

$$W_{non-con} = T + V. \tag{1.26}$$

1.2 Work Energy Theorem

The work energy theorem states that the total work done to the system W is equals to the change in kinetic energy of the system ΔK.E.

$$W = \Delta \text{K.E.}. \tag{1.27}$$

The work done of the system can be split into the work done by conservative force or the work done by non-conservative force

[2] Some authors use ΔW to denote work done, however, as work done should not be interpreted as changes, which would be meaningless, the Δ symbol is omitted. Formally, $đW$ is used to denote the inexact differential, but the complexity of the symbol forbids me to consistently type it in latex.

$$W = W_{\text{con}} + W_{\text{non-con}}, \tag{1.28}$$

we can therefore define the potential energy P.E. as

$$\Delta\text{P.E.} = -W_{\text{con}}, \tag{1.29}$$

so that the work done by the non conservative forces $W_{\text{non-con}}$ is equals to change in total mechnical energy (kinetic energy plus potential energy),

$$W_{\text{non-con}} = \Delta\text{K.E.} + \Delta\text{P.E.}. \tag{1.30}$$

1.3 System of particles

Having laid out the rudimentary principles, we now investigate the motion of a system of particles.

The translational equation of motion of the ith particle is

$$\mathbf{F}_i = \sum_j \mathbf{F}_{j\to i} + \mathbf{F}_{i,ext} = \frac{\mathrm{d}^2(m_i\mathbf{r}_i)}{\mathrm{d}dt^2}. \tag{1.31}$$

Summing over all particles,

$$\sum_i\sum_j \mathbf{F}_{j\to i} + \sum_i F_{i,\text{ext.}} = F_{\text{ext.,net}} = \sum_i \frac{\mathrm{d}^2(m_i\mathbf{r}_i)}{\mathrm{d}dt^2} = \frac{\mathrm{d}^2}{\mathrm{d}dt^2}\sum_i m_i\mathbf{r}_i = (\sum_i m_i)\ddot{\mathbf{R}}, \tag{1.32}$$

where we have used Newton's third law, stating that $\mathbf{F}_{i\to j} = -\mathbf{F}_{j\to i}$ and

$$\mathbf{R} = \frac{\sum_i m_i\mathbf{r}_i}{\sum_i m_i} \tag{1.33}$$

is defined as the position vector of the center of mass of the system.

This tells us that the total linear momentum of the system is the same as if the entire mass were concentrated at the center of mass and moving with it

Now for the rotational equation of motion of the ith particle, we have

$$\mathbf{r}_i \times \mathbf{F}_i = \mathbf{r}_i \times \mathbf{F}_{i,ext} + \mathbf{r}_i \times \sum_j \mathbf{F}_{j\to i} = \frac{dL_i}{dt}. \tag{1.34}$$

Summing over all particles,

$$\sum_i (\mathbf{r}_i \times \mathbf{F}_{i,\text{ext.}}) + \sum_i (\mathbf{r}_i \times \sum_j \mathbf{F}_{j\to i}) = \boldsymbol{\tau}_{\text{ext.}} + \sum_i ((\mathbf{r}_i - \mathbf{r}_j) \times \mathbf{F}_{j\to i}) = \boldsymbol{\tau}_{\text{ext.}} = \sum_i \frac{dL_i}{dt} = \dot{\mathbf{L}}_{\text{tot.}}. \tag{1.35}$$

where we again used Newton's third law and assumed that the internal forces are central, *i.e.,* the force between two particles act on the line connecting them.

To express $\mathbf{L}_{tot}$ in a more convenient form, we define $\mathbf{r}'_i = \mathbf{r}_i - \mathbf{R}$[3] as shown in fig. 1.1, thus $\mathbf{p}_i = m_i\dot{\mathbf{r}}'_i + m_i\dot{\mathbf{R}}$ and the total angular momentum becomes

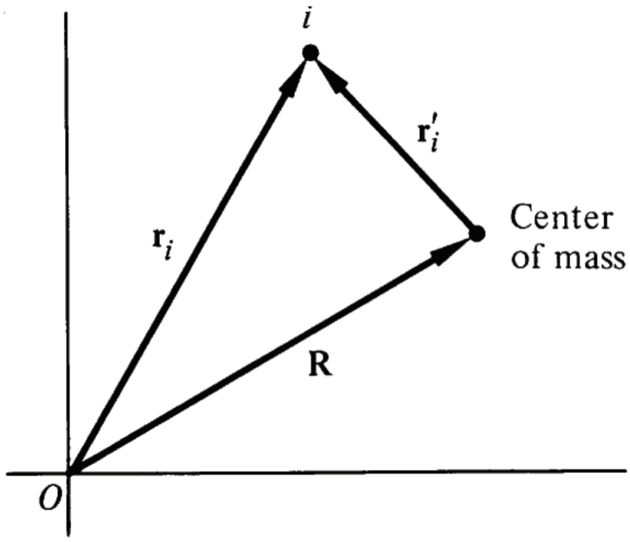


Figure 1.1

$$\mathbf{L}_{tot} = \sum_i \mathbf{r}_i \times \mathbf{p}_i = \sum_i ((\mathbf{r}'_i + \mathbf{R}) \times (m_i\dot{\mathbf{r}}'_i + m_i\dot{\mathbf{R}})) = (\sum_i m_i)\mathbf{R} \times \dot{\mathbf{R}} + \sum_i m_i(\mathbf{r}'_i \times \dot{\mathbf{r}}'_i) \tag{1.36}$$

where the cross terms $\sum_i(\mathbf{R} \times m_i\dot{\mathbf{r}}'_i + \mathbf{r}'_i \times m_i\dot{\mathbf{R}})$ are omitted since $\sum_i m_i\mathbf{r}'_i = 0$ from the definition of the center of mass.

So we see that the total angular momentum of a system of particles (due to $\dot{\mathbf{r}}$) can be split into two parts. The first term is due to the orbital motion of the center of mass about the origin due to transnational motion (due to $\dot{\mathbf{R}}$) and the second is due to the spinning motion of the particles around their center of mass (due to $\dot{\mathbf{r}}'$).

The same reasoning applies to the kinetic energy for a system of particles, where one term is attributed to the collective movement, while another arises from the rotational motion about the center of mass

$$T = \sum_i \frac{1}{2}m_i v_i^2 = \frac{1}{2}\sum_i m_i(\dot{\mathbf{r}}'_i + \dot{\mathbf{R}})^2 = \frac{1}{2}\sum_i m_i\dot{\mathbf{r}}'^2_i + \frac{1}{2}(\sum_i m_i)\dot{\mathbf{R}}^2 \tag{1.37}$$

[3]We will adopt this convention for the rest of this set of notes

where we neglect the cross term $\sum_i m_i(\dot{\mathbf{r}'_i} \cdot \dot{\mathbf{R}})$ for the same reason explained above.

virial theorem var

1.4 Rigid Body Mechanics

1.4.1 Prerequisites

If one were to choose a theorem that represents the crux of rigid body motion, one would have to pick Chasles' Theorem, which states that it is always possible to describe an arbitrary displacement of a rigid body by a translation of its center of mass plus a rotation around its center of mass (it can rotate about an arbitrary point but the center of mass is the most convenient choice).[4][5] The formal proof requires complex matrix algebra but a simple way to demonstrate the theorem is given in section A.1. Since the translational and the rotational motion of a rigid body are separable, so we almost always assume that the translational motion has already been accounted for. In fact, we will assume that the center of mass is at rest for the rest of this section.

If one were to pick a second theorem, then it would be Euler's Theorem, which states that any displacement of a rigid body such that a point on the rigid body remains fixed is equivalent to a single rotation about some axis that runs through the fixed point. Since the center of mass is always fixed as established above, it tells us that rotation about the center of mass means that all points on the rigid body undergo circular motion with respect to the closest point on an axis that runs through the center of mass where the direction of the axis defines the rotational motion and is in the same direction as the angular velocity which will be explored more in the next subsection. The proof of Euler's Theorem will not be given here due to its complexity.

Before diving into the physics of rigid body motion, some conventions of notations used in this set of notes should be explicitly stated first, as different texts would use different notations.

1. The spaced-fixed coordinate system, which is stationary in the lab frame has axes ($\hat{\mathbf{x}}, \hat{\mathbf{y}}$ and $\hat{\mathbf{z}}$) which obeys the right-hand rule. Quantities observed from the lab frame (or the space frame) are the same as quantities measured from the spaced-fixed coordinate system.

2. The body-fixed coordinate system has axes ($\hat{\mathbf{1}}, \hat{\mathbf{2}}$ and $\hat{\mathbf{3}}$) which also follows the right-hand rule and always coincide with the principle axes of the body. Quantities observed from the body frame are the same as quantities measured from the body-fixed coordinate system.

3. The instantaneous inertial frame with axes labeled $\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2$ and $\hat{\mathbf{e}}_3$ is an inertial frame which its axes coincide with the body's principal axes only at time instant t. This frame is not rotating with respect to the lab frame so it is equally superior.

4. The Euler angles that are used to transform between these two coordinate systems

[4] Another way to construct any displacement is first to do a rotation and then translate parallel to the axis of rotation, we reverse the order of translation and rotation while adding a constraint on the translating direction.

[5] Yet another interesting and useful fact is that if the motion of the body is planer (*i.e.,* the angular velocity is perpendicular to the linear velocity), then there always exists an instantaneous axis of rotation (which need not be inside the rigid body) that is parallel to the angular velocity such that any infinitesimal displacement can be constructed by rotating around this axis (This is the third way in which a displacement of a rigid body can be constructed). The proof of this fact is given here. The instantaneous axis of rotation can be constructed geometrically mentioned in idea 33 of this handout by Jaan Kalda.

are rotated in the z-x-z sequence.

1.4.2 Angular velocity vector

Before handling the rather complicated mathematical treatments, it is useful to define what do we mean by angular velocity.

Angular velocity, similar to linear velocity, is a quantity describing a body's (more rigorously, the body-fixed frame's) motion that is independent of the choice of a coordinate system or origin. One may imagine there is an "angular-speedometer" that can measure the angular velocity of a rigid body undergoing any general motion. However, it is frame-dependent, meaning that the angular velocity observed in the lab frame is different from that observed from another.

Suppose we have 3 orthogonal frames: the lab frame, which is not rotating and fixed in space.[6] And two other frames whose origins remained fixed (as our interests do not lie on the translational motions and rotational motions can be analyzed separately from translational motions) and can rotate freely about their origins. Each of these two frames possess their own angular velocity vector as observed from the fixed lab frame, which passes through their origins and the direction defines their rotational motion as guaranteed by Euler's theorem, where every points co-rotating with the frame trace out a circle with the center at the closet point to the rotation axis.

As linear velocity is defined as the time derivative of the displacement vector, one may be tempted to define an "angular displacement vector", describing how an object undergo rotation and the angular velocity can be simply defined as the time derivative of the "angular displacement vector". However, this is not possible for the fact that finite rotations do not commute in 3-dimensional space (for 2D case, rotations do commute as there are only 2 degrees of freedom which can be assigned to positive and negative signs) as one can play with literally any object to try it out, so

$$\Delta\boldsymbol{\theta} \stackrel{?}{=} \Delta\theta_x\hat{\mathbf{x}} + \Delta\theta_y\hat{\mathbf{y}} \neq \Delta\theta_y\hat{\mathbf{y}} + \Delta\theta_x\hat{\mathbf{x}}. \tag{1.38}$$

However, we *can* define an "infinitesimal angular displacement vector" as angular infinitesimal displacements do commute (less obvious but one still gets a feeling by playing with an object but limiting the angles rotated to be very small), so

$$\delta\boldsymbol{\theta} = \delta\theta_x\hat{\mathbf{x}} + \delta\theta_y\hat{\mathbf{y}} = \delta\theta_y\hat{\mathbf{y}} + \delta\theta_x\hat{\mathbf{x}}. \tag{1.39}$$

To prove the above result, we consider fig. 1.2. Without loss of generality, we define the $z-$ axis of the lab frame (which is arbitrarily defined) to coincide with the angular velocity vector of the rotating frame, and the $\mathbf{r}$ vector to be the position vector of any point co-rotating with the rotating frame. The direction of rotation $\delta\boldsymbol{\theta}$ can be x or y axis in the above equation.

[6] From the similarity between angular velocity and linear velocity, one may think there is no universally superior frame of reference when analyzing rotational motion due to relativity. However, rotation is absolute as one may determine whether it is rotating from local measurement, e.g. whether the equipotential surface of a bucket of water is parabolic or horizontal. Although there is still debate on this topic, e.g. here, we take this fact for granted as we are still in the realm of Newtonian physics.

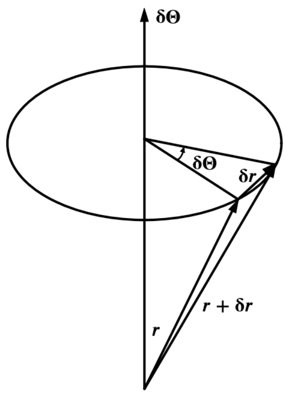


Figure 1.2

From it, it is clear that

$$\delta\mathbf{r} = \delta\boldsymbol{\theta} \times \mathbf{r}. \tag{1.40}$$

Considering two successive rotation through $\delta\boldsymbol{\theta}_1$ and $\delta\boldsymbol{\theta}_2$, we have

$$\delta\mathbf{r}_{12} = \delta\boldsymbol{\theta}_1 \times \mathbf{r} + \delta\boldsymbol{\theta}_2 \times (\mathbf{r} + \delta\mathbf{r}) = (\delta\boldsymbol{\theta}_1 + \delta\boldsymbol{\theta}_2) \times \mathbf{r} = \delta\mathbf{r}_{21} \tag{1.41}$$

if we neglect the higher-order term. An alternate proof providing more intuition but more tedious is given in section A.2.

Dividing eq. (1.40) by δt, we have

$$\mathbf{v} = \frac{\delta\mathbf{r}}{\delta t} = \frac{\delta\boldsymbol{\theta}}{\delta t} \times \mathbf{r}. \tag{1.42}$$

In a more general case where the origin is moving at a velocity $\mathbf{v}_O$, then the velocity of point P in the rigid body will be

$$\mathbf{v}_P = \mathbf{v}_O + \boldsymbol{\omega} \times \mathbf{r}_{O\to P} \tag{1.43}$$

which is a very useful equation since it relates the velocity of any two points in the rigid body with the common angular velocity (note that O need not be the center of mass of the rigid body, as if true, $\mathbf{v}_P = \mathbf{v}_{c.m.} + \boldsymbol{\omega} \times \mathbf{r}_{c.m.\to\mathbf{P}}$ and $\mathbf{v}_\mathbf{O} = \mathbf{v}_{c.m.} + \boldsymbol{\omega} \times \mathbf{r}_{c.m.\to\mathbf{O}}$ yields the general result. In fact, eq. (1.43) can be regarded as the mathematical definition for the angular velocity vector.

Another very useful property of the angular velocity vector is that the law of angular velocity addition to find the relative angular velocity between different frames is exactly analogous to the law of linear velocity addition, where

$$\boldsymbol{\Omega}_{1\text{rel.}2} = \boldsymbol{\Omega}_{1\text{rel.}3} - \boldsymbol{\Omega}_{2\text{rel.}3}. \tag{1.44}$$

To prove this, we first define clearly what do we mean by relative velocity in the linear case. Suppose we have a point P_1 co-rotating with S_1 and P_2 fixed in S_2. In the lab frame S_3, the displacement vectors of P_1 and P_2 are defined as the changes in their linear positions as measured in the lab frame. The linear velocity vectors are defined as the displacement vectors divided by a regular time interval, and the relative velocity of the 2 points (or 2 frames) is the difference in their linear velocity vectors. In the angular velocity case, we can simply follow the same argument as "angular displacement vector" is well defined as long as the time interval concerned tends to zero. It is helpful to visualize the "angular displacement vector" in the 2D case, where the time interval concerned is not limited to infinitesimally small, then it becomes clear that angular velocity vectors do add like linear velocity vectors by considering the most simple case: S_1 rotating with the angular velocity $\omega_1\hat{\mathbf{z}}$ and S_2 with $\omega_2\hat{\mathbf{z}}$, then after a time interval Δt, the angular displacement vectors are $\boldsymbol{\theta}_1 = \omega_1 \Delta t\hat{\mathbf{z}}$ and $\boldsymbol{\theta}_2 = \omega_2 \Delta t\hat{\mathbf{z}}$ and the relative displacement vector is $\boldsymbol{\theta}_{1rel.2} = (\omega_1 - \omega_2)\Delta t\hat{\mathbf{z}}$ thus the relative angular velocity vector is $\boldsymbol{\omega}_{1rel.2} = (\omega_1 - \omega_2)\hat{\mathbf{z}}$. The same thing applies in our 3D world, just that $\Delta \to \delta$ and it is harder to visualize the entire process.

Example: Rolling Cylinder.

Question: A cylinder of raidus r is rolling down a inclinded plane with angle α to the horizontal. The axis of the cylinder moes at constant velocity v. Let (x, y) be a point on the cylinder at distance a from the cylinder's axis. Determine the condition on a that the point sometimes moves in an exactly upwards manner.

Solution: We set the origin at the center of mass of the cylinder when $t = 0$, then we have

$$\mathbf{v}_a = \mathbf{v}_{\text{c.m.}} + \boldsymbol{\omega} \times \mathbf{r_a}. \tag{1.45}$$

In components form,

$$\begin{cases} \dot{x} &= v\cos\alpha + \frac{v}{r}(y + vt\sin\alpha), \\ \dot{y} &= -v\sin\alpha - \frac{v}{r}(x - vt\sin\alpha). \end{cases} \tag{1.46}$$

When $\dot{x} = 0$, we get

$$y = -r\cos\alpha - vt\sin\alpha \implies \dot{y} = -v\sin\alpha \implies x = vt\cos\alpha. \tag{1.47}$$

Imposing the condition $x^2 + y^2 \leq a^2$, *i.e.,* the point where $\dot{x} = 0$ is inside the cylinder, we get

$$a \geq r\cos\alpha. \tag{1.48}$$

1.5 Tensor of Inertia

1.5.1 Angular Momentum and Energy

Now we return to eq. (1.36) and try to evaluate the abstract summation form of the spin angular momentum due to rotation about the center of mass $\mathbf{L}_{rot} = \sum_i m_i(\mathbf{r}'_i \times \dot{\mathbf{r}}'_i)$ when the rigid body is rotating about its center of mass at an angular velocity $\boldsymbol{\omega}$.

Now from fig. 1.2 we can conclude the general relationship that if a vector $\mathbf{r}$ is rotating about a fixed origin with angular velocity $\boldsymbol{\omega}$, then we have the relation

$$\frac{d\mathbf{r}}{dt} = \boldsymbol{\omega} \times \mathbf{r}. \tag{1.49}$$

Therefore, $\mathbf{L}_{rot}$ becomes

$$\mathbf{L}_{rot} = \sum_i m_i(\mathbf{r}'_i \times \dot{\mathbf{r}}'_i) = \sum_i m_i(\mathbf{r}'_i \times (\boldsymbol{\omega} \times \mathbf{r}'_i)) = \sum_i m_i(|\mathbf{r}'_i|^2 \boldsymbol{\omega} - \mathbf{r}'_i(\mathbf{r}'_i \cdot \boldsymbol{\omega})). \tag{1.50}$$

From here, we can explicitly write out the x, y and z components of $\mathbf{L}_{rot}$ as

$$\begin{aligned}
L_{rot,x} &= \sum_i m_i((x_i'^2 + y_i'^2 + z_i'^2)\omega_x - x'_i(x'_i\omega_x + y'_i\omega_y + z'_i\omega_z)) \\
&= \sum_i m_i((y_i'^2 + z_i'^2)\omega_x - (x'_i y'_i)\omega_y - (x'_i z'_i)\omega'_z), \\
L_{rot,y} &= \sum_i (m_i(x_i'^2 + z_i'^2)\omega_y - (y'_i z'_i)\omega_z - (x'_i y'_i)\omega_x), \\
\text{and } L_{rot,z} &= \sum_i m_i((x_i'^2 + y_i'^2)\omega_z - (x'_i z'_i)\omega_x - (y'_i z'_i)\omega_y).
\end{aligned} \tag{1.51}$$

In matrix form,

$$\mathbf{L}_{rot} = \begin{pmatrix} L_{rot,x} \\ L_{rot,y} \\ L_{rot,z} \end{pmatrix} = \begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix} \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix} = \tilde{\mathbf{I}}\boldsymbol{\omega}. \tag{1.52}$$

Similarly, for the abstract sum for the kinetic energy in eq. (1.37) due to the rotational motion, it now becomes

$$T_{rot} = \frac{1}{2}\sum_i m_i \dot{\mathbf{r}}_i'^2 = \frac{1}{2}\sum_i m_i(\boldsymbol{\omega}\times\mathbf{r}'_i)\cdot(\boldsymbol{\omega}\times\mathbf{r}'_i) = \frac{1}{2}\boldsymbol{\omega}\cdot\sum_i m_i\mathbf{r}'_i\times(\boldsymbol{\omega}\times\mathbf{r}'_i) = \frac{1}{2}\boldsymbol{\omega}\cdot\mathbf{L} \tag{1.53}$$

where we used the vector identity $(\mathbf{A} \times \mathbf{B}) \cdot \mathbf{C} = \mathbf{A} \cdot (\mathbf{B} \times \mathbf{C})$.

Now one of the great advantages of the use of principal axes is the simplification of eq. (1.53), as it now becomes

var 13.7 diagonal-ization and iner-tia ten-sor

$$T_{rot} = \frac{1}{2} I_{xx}\omega_x^2 + \frac{1}{2} I_{yy}\omega_y^2 + \frac{1}{2} I_{zz}\omega_z^2. \tag{1.54}$$

1.5.2 Parallel Axis Theorem

If the tensor of inertia about the center of mass $\tilde{\mathbf{I}}_{c.m.}$ and the displacement vector pointing from the center of mass to point $P = X\hat{\mathbf{i}} + Y\hat{\mathbf{j}} + Z\hat{\mathbf{k}}$ are known, then the tensor of inertia about point P will be

$$\begin{aligned}
I_{xx,P} &= \sum_i m_i(y_{i,P}'^2 + z_{i,P}'^2) = \sum_i m_i((y'_{i,c.m.} - Y)^2 + (z'_{i,c.m.} - Z)^2) \\
&= \sum_i m_i((y_{i,c.m.}^2 + z_{i,c.m.}^2) + (Y^2 + Z^2) - 2(y_{i,c.m.}Y + Z_{i,c.m.}Z)) \\
&= I_{xx,c.m.} + \sum_i m_i(Y^2 + Z^2) \\
\text{and } I_{xy,P} &= -\sum_i m_i(x_{i,P}y'_{i,P}) = \sum_i m_i((x_{i,c.m.} - X)(y_{i,c.m.} - Y) \\
&= \sum_i m_i((x_{i,c.m.}y_{i,c.m.}) - XY - (Xy_{i,c.m.} + x_{i,c.m.}Y)) \\
&= I_{xy,c.m.} - \sum_i m_i XY \ \ etc.
\end{aligned} \tag{1.55}$$

where the last term in each of the equations vanishes due to the property of the center of mass.

1.5.3 Perpendicular Axis Theorem

The perpendicular axis theorem states that for a planar lamina the moment of inertia about an axis perpendicular to the plane of the lamina is equal to the sum of the moments of inertia about two mutually perpendicular axes in the plane of the lamina, which intersect at the point where the perpendicular axis passes through. This theorem applies only to planar bodies and is valid when the body lies entirely in a single plane, with the exception that the body has cylindrical symmetry about the perpendicular axis, such as a cylinder.

1.5.4 Euler's Equations

With all the prerequisites explained, we are now ready to tackle the seemingly simple differential equation $\boldsymbol{\tau} = d\mathbf{L}/dt$. Consider a time instant t when a rigid body is rotating with $\boldsymbol{\omega}$. Since the body frame is non-inertial, we cannot apply this rotational Newton's law here. What we can do, however, is to consider an inertial frame that only coincides with the body frame at time t.

It is very important to have this picture in mind: at time t, the inertial frame axes $\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3$ are the same as the body axes $\hat{\mathbf{1}}, \hat{\mathbf{2}}, \hat{\mathbf{3}}$. Then, after time dt, the axes of the body-fixed coordinate system rotate by an angle of ωdt along $\boldsymbol{\omega}$ while the inertial frame axes remained stationary. So from the inertial frame, the body axes actually rotate with $\boldsymbol{\omega}$. We then repeat this procedure infinite time.

Writing out the equation of motion in this inertial coordinate system, we have

$$\begin{aligned}\boldsymbol{\tau} = \frac{d\mathbf{L}}{dt} = \frac{d}{dt}(L_1\hat{\mathbf{1}} + L_2\hat{\mathbf{2}} + L_3\hat{\mathbf{3}}) &= \frac{dL_1}{dt}\hat{\mathbf{1}} + L_1\frac{d\hat{\mathbf{1}}}{dt} + \frac{dL_2}{dt}\hat{\mathbf{2}} + L_2\frac{d\hat{\mathbf{2}}}{dt} + \frac{dL_3}{dt}\hat{\mathbf{3}} + L_3\frac{d\hat{\mathbf{3}}}{dt} \\ &= \frac{dL_1}{dt}\hat{\mathbf{1}} + \frac{dL_2}{dt}\hat{\mathbf{2}} + \frac{dL_3}{dt}\hat{\mathbf{3}} + (\boldsymbol{\omega}\times\hat{\mathbf{1}})L_1 + (\boldsymbol{\omega}\times\hat{\mathbf{2}})L_2 + (\boldsymbol{\omega}\times\hat{\mathbf{3}})L_3\end{aligned} \tag{1.56}$$

where $i = 1, 2$ and 3 and we used eq. (1.49) since the body-fixed axes ($\hat{\mathbf{1}}, \hat{\mathbf{2}}$ and $\hat{\mathbf{3}}$) are rotating angular velocity $\boldsymbol{\omega}$ about the inertial instantaneous frame as mentioned.

Splitting the vector equation into three components, we have three non-linear coupled first-order differential equations

$$\begin{aligned}\tau_1 &= I_1\dot{\omega}_1 + \omega_2\omega_3(I_3 - I_2) \\ \tau_2 &= I_2\dot{\omega}_2 + \omega_3\omega_1(I_1 - I_3) \\ \tau_3 &= I_3\dot{\omega}_3 + \omega_1\omega_2(I_2 - I_1).\end{aligned} \tag{1.57}$$

An alternate derivation of Euler's equations with discrete time interval considerations can be found in Chapter 8.7.2 of Kleppner. An alternate proof of Euler's equations by the Euler-Lagrange equation can be found in Chapter 13.18 of Cline.

One has to be reminded that although the set of equations are given in body-fixed coordinates and thus are only valid at time t where the body frame coincides with the inertial frame, since t is arbitrarily chosen, the equations of motion tell us things that are more general than the behaviors of the system at that mere instant. In fact, we can create an infinite number of instantaneous inertial frames such that Euler's equations are always valid. In retrospect, the introduction of an instantaneous inertial frame was merely to derive Euler's equations and nothing more. From now on there are only 2 frames that matter: the body frame and the lab frame.

Also, since the Euler equations only depend on the principal moments of inertia I_1, I_2 and I_3, thus all bodies having the same principal moments of inertia will behave exactly the same even though the bodies may have very different shapes. The simplest geometrical shape of a body having three different principal moments is a homogeneous ellipsoid. Thus, the rigid body motion often is described in terms of the equivalent ellipsoid that has the same principal moments of inertia.

Example: Kleppner (3rd. ed) Example 8.16

Question: Due to $\boldsymbol{\omega}$ not necessarily parallel to $\mathbf{L}$, many peculiar phenomena are observed in rigid body motion. One of which is the Tennis Racket Theorem (also known as the Intermediate axis theorem), which states that the rotations about the 2 principal axes which have the largest and the smallest moment of inertia are stable while the rotation about the intermediate axis is not. Prove it.

Solution:
To explain this phenomenon, we suppose that the body initially spins with $\boldsymbol{\omega} = \omega_1\hat{\mathbf{e}}_1$ and receives small perturbations on ω_2 and ω_3. Then according to the Euler's equations, we have $\omega_1 =$ constant and

$$\frac{\mathrm{d}^2\omega_2}{\mathrm{d}dt^2} + (\frac{(I_1 - I_2)(I_1 - I_3)}{I_2 I_3}\omega_1^2)\omega_2 = 0 \tag{1.58}$$

as one can easily verify. So we see that ω_2 undergo simple harmonic motion if I_1 is the largest or the smallest moment of inertia, but increase exponentially with time and the motion is unstable.[a]

[a] For a more intuitive explanation, refer to the explanation given by the famous mathematician Terrance Tao as well as this video by the famous YouTuber Veritasium.

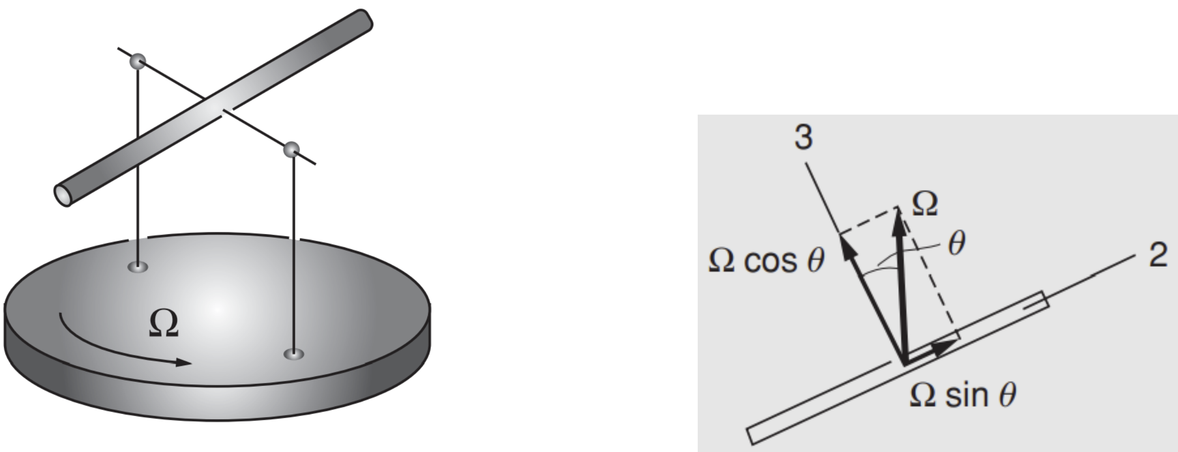
Example: Kleppner (3rd. ed) Example 8.17

Question: A uniform rod is mounted on a horizontal frictionless axle through its center. The axle is carried on a turntable rotating at a constant angular velocity $\boldsymbol{\Omega}$ as depicted in fig. 1.3 . Find $\theta(t)$ shown in the figure.

Solution: Referring to the figures, we have $\omega_1 = \dot{\theta}, \omega_2 = \Omega\sin\theta$ and $\omega_3 = \Omega\cos\theta$. Substituting them into the Euler's equations and leveraging the small angle approximation $\sin\theta \approx \theta$ gives

$$\ddot{\theta} + (\frac{I_3 - I_2}{I_1})\Omega^2\theta = 0. \tag{1.59}$$

So we conclude that θ undergo simple harmonic motion with angular frequency $\gamma = \sqrt{\frac{I_3 - I_2}{I_1}}\Omega$.



(a) (b)

Figure 1.3

1.6 Torque-free Precession

One of the most classic applications of Euler's equations is a torque-free procession. Consider a symmetric top with I_1 being the moment of inertia about the symmetric axis and $I_2 = I_3 = I_\perp$. Then the equations give $\omega_1 = \text{constant} = \omega_s$ and

$$\frac{\mathrm{d}^2\omega_2}{\mathrm{d}dt^2} + (\frac{I_1 - I_\perp}{I_\perp})^2\omega_s^2\omega_2 = 0. \tag{1.60}$$

So ω_2 undergo simple harmonic motion with angular frequency $\gamma = \left|\frac{I_1 - I_\perp}{I_\perp}\right|\omega_s$

$$\omega_2 = \omega_\perp \cos\gamma t \tag{1.61}$$

where $\omega_\perp$ depends on the initial condition .

Further calculation would give that

$$\omega_3 = \pm\ \omega_\perp \sin\gamma t \tag{1.62}$$

where the positive sign corresponds to the case where $I_1 > I_\perp$ indicates the body is short and fat so the spin is clockwise, and vice versa.

To get qualitatively what really happens, refer to fig. 1.4 .

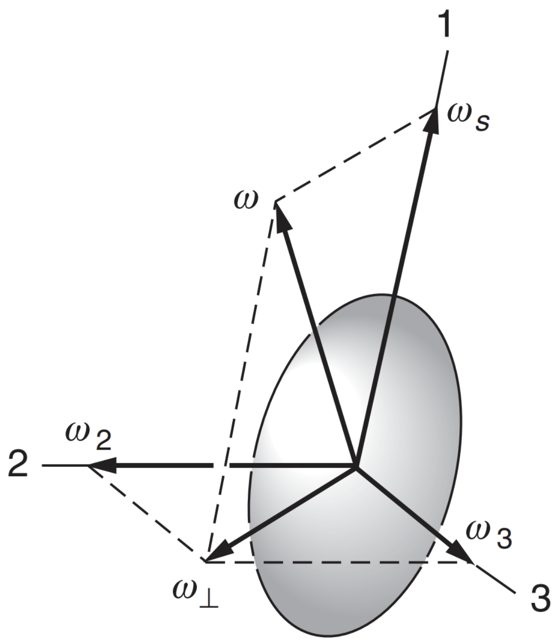


Figure 1.4

$\omega_1 = \omega_s = \text{constant}$ simply means that in the body frame, the component of $\boldsymbol{\omega}$ on $\hat{\mathbf{1}}$ has a fixed magnitude ω_s. [7]

[7]This also means that at every time instant t, $\hat{\mathbf{2}}$ and $\hat{\mathbf{3}}$ revolve about $\hat{\mathbf{e}}_1$ (technically not $\hat{\mathbf{1}}$ since $\hat{\mathbf{1}}$ is not fixed in the instantaneous inertial frame so it is meaningless to talk about rotation around $\hat{\mathbf{1}}$ in this

The solution for ω_2 and ω_3 means that they are actually components of $\boldsymbol{\omega}_\perp$ on $\hat{\mathbf{2}}$ and $\hat{\mathbf{3}}$ respectively when $\boldsymbol{\omega}_\perp$ is rotating about $\hat{\mathbf{1}}$ at the angular speed γ when observed in the body frame.

Combining these two insights, we can say that $\boldsymbol{\omega}_\perp$ rotate about $\hat{\mathbf{1}}$ at the angular speed $\gamma + \omega_s$ when observed from the lab frame by simple angular velocity addition.

Furthermore, since $I_2 = I_3 = I_\perp$ and $\mathbf{L}_2 = I_2\boldsymbol{\omega}_2$ and $\mathbf{L}_3 = I_3\boldsymbol{\omega}_3$ therefore $\mathbf{L}_\perp = \mathbf{L}_2 + \mathbf{L}_3 = I_\perp(\boldsymbol{\omega}_2 + \boldsymbol{\omega}_3) = I_\perp\boldsymbol{\omega}_\perp$ which means that $\hat{\mathbf{1}}, \boldsymbol{\omega}_1 = \omega_s\hat{\mathbf{1}}, \mathbf{L}_1 = I_1\boldsymbol{\omega}_1, \boldsymbol{\omega}_\perp, \mathbf{L}_\perp = I_\perp\boldsymbol{\omega}_\perp, \boldsymbol{\omega} = \boldsymbol{\omega}_1 + \boldsymbol{\omega}_\perp, \mathbf{L} = \mathbf{L}_1 + \mathbf{L}_\perp$ are all in the same plane, and since $\hat{\mathbf{1}}$ is fixed in the body frame, the only degree of freedom is that all the vectors mentioned above rotate about $\hat{\mathbf{1}}$ with the same angular speed.[8] But we already found out that one of the vectors, namely $\boldsymbol{\omega}_\perp$ has an angular speed of γ, so all the vectors mentioned have the same angular velocity $\gamma\hat{\mathbf{1}}$.

We have already solved the problem in the body frame, next we transform it back into the lab frame, which is what we care about the most.

In the space-fixed inertial frame, since there are no external torques in torque-free precession, $\mathbf{L}$ is now fixed in place.

From the analysis done in the body frame, we must bear this fact in mind: all the vectors concerned in this problem are in the same plane. To visualize, it is helpful to imagine that all the vectors are on a piece of paper with $\hat{\mathbf{1}}$ and $\boldsymbol{\omega}_\perp$ being the two adjacent edges of the paper and $\mathbf{L}$ being the diagonal (it is always possible since the size of the paper is arbitrary). In the body frame, $\hat{\mathbf{1}}$ is held still so the piece of paper rotates about one vertical edge with angular speed γ similar to how a door rotates about a door hinge.

However, refer to fig. 1.5 where now we wish to fix $\mathbf{L}$ in place in space frame meaning that the 2 corners (the tip and the tail of $\mathbf{L}$) are now stationary and the piece of paper rotates about $\mathbf{L}$. This picture explains intuitively why although $\boldsymbol{\omega}$ and $\mathbf{L}$ has the same angular velocity in the body frame but when switched to the lab frame, where $\mathbf{L}$ is fixed, $\boldsymbol{\omega}$ is not fixed but is now co-rotating with $\hat{\mathbf{1}}$ about $\mathbf{L}$ with the same angular speed Ω_p. Mathematically, $\mathbf{L}$ is also rotating with the angular speed Ω_p, just that the axis of rotation is $\mathbf{L}$ itself, so it is equivalent to having no rotation at all.

To find this new common angular speed Ω_p, we can utilize the angular velocity addition formula eq. (1.44), where frame 1 is a frame where $\boldsymbol{\omega}$ is at rest, frame 2 is lab frame and frame 3 is the body frame. So

$$\boldsymbol{\Omega}_{\boldsymbol{\omega}\text{ rel. lab}} = \boldsymbol{\Omega}_{\boldsymbol{\omega}\text{ rel. body}} - \boldsymbol{\Omega}_{\text{ lab rel. body}}. \tag{1.63}$$

or

frame and also $\hat{\mathbf{1}}, \hat{\mathbf{2}}$ and $\hat{\mathbf{3}}$ are relatively fixed so no axis is rotating about another axis but since $\hat{\mathbf{1}}$ and $\hat{\mathbf{e}}_1$ coincide at that moment, this saying is generally accepted) at constant angular speed ω_s when observed from the instantaneous inertial frame.

[8]The fact that the angles between all the vectors are fixed is trivial in the body frame considering the mathematical form of each vector listed above. To prove this fact in space frame, notice that $\mathbf{L}$ of the body is fixed in *torque-free* precession, and we have shown that $\omega_1 = \omega_s =$ constant and $\omega_\perp$ is constant as well, so α shown in fig. 1.5 must be constant. To be extra cautious, we can say since $T_{rot} = \frac{1}{2}\boldsymbol{\omega}\cdot\mathbf{L} = \frac{1}{2}\omega L\cos\alpha$ (eq. (1.53)) must be constant since there is no external work done, so α must be constant.

$$\Omega_p \hat{\mathbf{z}} = \gamma \hat{\mathbf{1}} - (-\boldsymbol{\omega}). \tag{1.64}$$

Resolving this vector equation along $\hat{\mathbf{1}}$ gives

$$\begin{aligned} \Omega_p \cos\alpha &= \gamma + \omega_s \\ \Omega_p &= \frac{I_1 \omega_s}{I_\perp \cos\alpha}. \end{aligned} \tag{1.65}$$

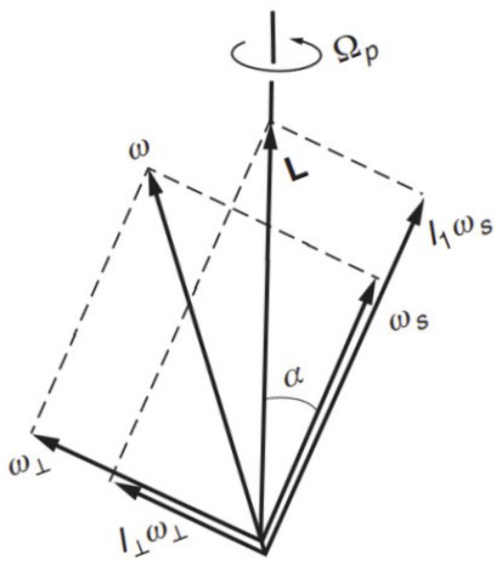


Figure 1.5

The intuitive explanation as to why $\Omega_p \cos\alpha = \gamma + \omega_s$ is as follows:

Firstly, as mentioned, $\omega_1 = \omega_s =$ constant implies that $\hat{\mathbf{2}}$ and $\hat{\mathbf{3}}$ revolve around $\hat{\mathbf{1}}$ (technically, $\hat{\mathbf{e}}_1$) at ω_s. However, even then, we have calculated that $\boldsymbol{\omega}_\perp$ (and also $\boldsymbol{\omega}$ and other relevant vectors) still have angular speed γ in the body frame where $\hat{\mathbf{2}}$ and $\hat{\mathbf{3}}$ are at rest. This means that those sets of vectors rotate at the angular speed $\gamma + \omega_s$ about $\hat{\mathbf{1}}$ in the lab frame.

Secondly, we resort to the "2D paper model" developed above. We now know that for the "door hinge" mode (rotate about $\hat{\mathbf{1}}$), the angular speed observed from the lab frame is $\gamma + \omega_s$. We want to know what the angular speed observed from the lab frame is when rotating about $\mathbf{L}$. To answer this, we have to remember the vector property of angular velocity $\boldsymbol{\Omega}_p$. We utilize this fact and resolve $\boldsymbol{\Omega}_p$ along $\hat{\mathbf{1}}$ (and its perpendicular direction). The former angular speed which equals to $\Omega_p \cos\alpha$ should be identical to the angular speed of the set of vectors when $\hat{\mathbf{1}}$ is fixed which we calculated to be $\gamma + \omega_s$.

1.6.1 Euler Angles

The description of rigid body rotation is greatly facilitated by transforming from the space-fixed (lab) coordinates $(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$ to the body-fixed coordinates $(\hat{\mathbf{1}}, \hat{\mathbf{2}}, \hat{\mathbf{3}})$ since the inertia tensor measured with this coordinate is always diagonal. They can be related, as introduced in the "Maths", by

$$(\hat{\mathbf{1}}, \hat{\mathbf{2}}, \hat{\mathbf{3}}) = \boldsymbol{\lambda}(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}). \tag{1.66}$$

As mentioned in "Maths", only 3 independent angles are needed for any rotational transformation. By convention, the Euler angles ϕ, θ, ψ are used. Refer to fig. 1.6.

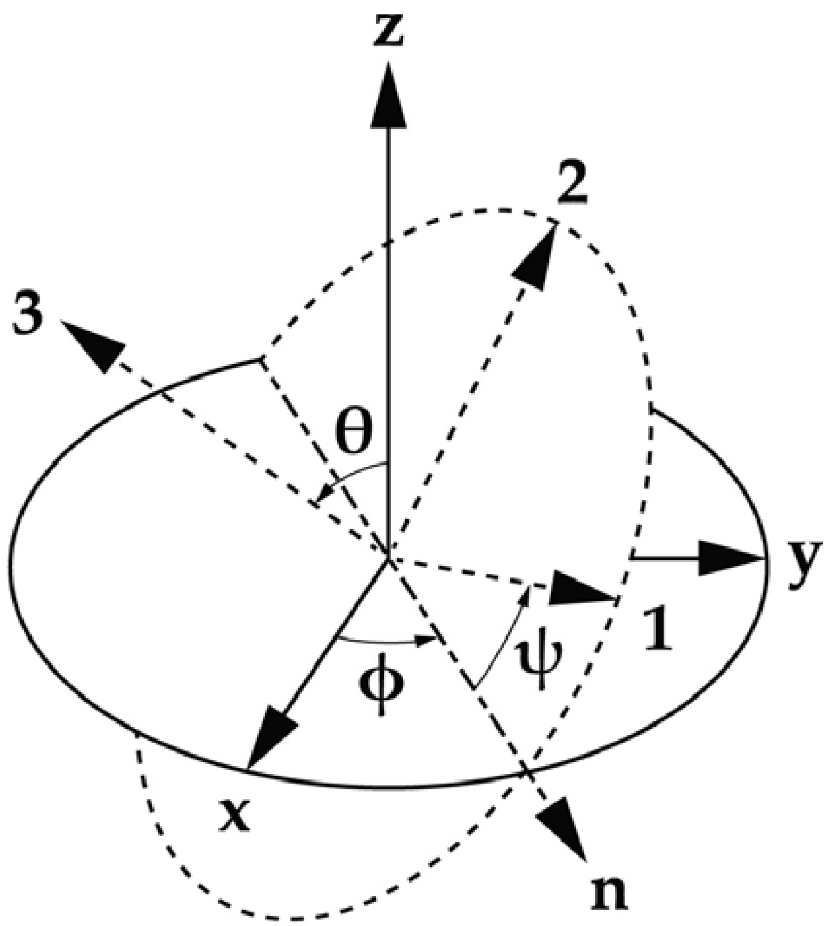


Figure 1.6

The unit vector defined by $\hat{\mathbf{n}} = \hat{\mathbf{z}} \times \hat{\mathbf{3}}$ is called the line of nodes.

Firstly, $\hat{\mathbf{x}}$ is made to coincide with the line of node $\hat{\mathbf{n}}$, then while keeping $\hat{\mathbf{x}}$ unchanged, $\hat{\mathbf{z}}$ is made to coincide with $\hat{\mathbf{3}}$ (which is possible since the line of node is defined to be $\hat{\mathbf{n}} = \hat{\mathbf{z}} \times \hat{\mathbf{3}}$). Lastly, while keeping $\hat{\mathbf{z}}$ uncharged, $\hat{\mathbf{x}}$ is made to be coincide with 1 axis. As $\hat{\mathbf{x}}$ and $\hat{\mathbf{z}}$ are in place, due to the orthogonality of the systems, $\hat{\mathbf{y}}$ is bound to coincide with $\hat{\mathbf{2}}$.

The rotational matrices of each rotation are

$$\boldsymbol{\lambda}_\phi = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \boldsymbol{\lambda}_\theta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix}, \quad \text{and} \quad \boldsymbol{\lambda}_\psi = \begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{1.67}$$

Therefore the total rotational matrix is

$$\boldsymbol{\lambda} = \boldsymbol{\lambda}_\phi\boldsymbol{\lambda}_\theta\boldsymbol{\lambda}_\psi = \begin{pmatrix} \cos\phi\cos\psi - \sin\phi\cos\theta\sin\psi & \sin\phi\cos\psi + \cos\phi\cos\theta\sin\psi & \sin\theta\sin\psi \\ -\cos\phi\sin\psi - \sin\phi\cos\theta\cos\psi & -\sin\phi\sin\psi + \cos\phi\cos\theta\cos\psi & \sin\theta\cos\psi \\ \sin\phi\sin\theta & -\cos\phi\sin\theta & \cos\theta \end{pmatrix}. \tag{1.68}$$

The angular velocity will be

$$\boldsymbol{\omega} = \dot{\phi}\hat{\mathbf{z}} + \dot{\theta}\hat{\mathbf{n}} + \psi\hat{\mathbf{3}}. \tag{1.69}$$

Expressing $\hat{\mathbf{z}}$ and $\hat{\mathbf{n}}$ in terms of the body-fixed coordinates, we have

$$\begin{aligned} \hat{\mathbf{z}} &= \sin\theta\sin\psi\hat{\mathbf{1}} + \sin\theta\cos\psi\hat{\mathbf{2}} + \cos\theta\hat{\mathbf{3}} \\ \hat{\mathbf{n}} &= \cos\psi\hat{\mathbf{1}} - \sin\psi\hat{\mathbf{2}}. \end{aligned} \tag{1.70}$$

So

$$\boldsymbol{\omega} = \dot{\phi}\sin\theta\sin\psi + \dot{\theta}\cos\psi)\hat{\mathbf{1}} + (\dot{\phi}\sin\theta\cos\psi - \dot{\theta}\sin\psi)\hat{\mathbf{2}} + (\dot{\psi} + \dot{\phi}\cos\theta)\hat{\mathbf{3}}. \tag{1.71}$$

By playing a similar game, the angular velocity can be expressed in terms of the space-fixed coordinates, with

$$\boldsymbol{\omega} = (\dot{\theta}\cos\phi + \dot{\psi}\sin\theta\sin\phi)\hat{\mathbf{x}} + (\dot{\theta}\sin\phi - \dot{\psi}\sin\theta\cos\phi)\hat{\mathbf{y}} + (\dot{\phi} + \dot{\psi}\cos\theta)\hat{\mathbf{z}}. \tag{1.72}$$

The validity of the results can be verified by confirming that the dot product of $\boldsymbol{\omega}$ with itself

$$\boldsymbol{\omega}\cdot\boldsymbol{\omega} = \omega_1^2 + \omega_2^2 + \omega_3^2 = \omega_x^2 + \omega_y^2 + \omega_z^2 = \dot{\phi}^2 + \dot{\theta}^2 + \dot{\psi}^2 + 2\dot{\phi}\dot{\psi}\cos\theta \tag{1.73}$$

is an invariant under coordinates transformation as any scalar properties like mass, Lagrangian, or Hamiltonian should.

The advantage of working in the body-fixed coordinates is that the inertia tensor is diagonal, which greatly simplifies the work needed in expressing the kinetic energy as

$$T_{rot} = \frac{1}{2}(I_1(\dot{\phi}\sin\theta\sin\psi + \dot{\theta}\cos\psi)^2 + I_2(\dot{\phi}\sin\theta\cos\psi - \dot{\theta}\sin\psi)^2) + I_3(\dot{\phi}\cos\theta + \dot{\psi})^2). \tag{1.74}$$

2 Waves

2.1 Normal Modes

2.1.1 Equal Masses

We start with the simple case with the equation

$$m\ddot{\mathbf{x}} = -K\mathbf{x}, \tag{2.1}$$

where K is symmetric (if the system conserves energy as we will show below), thus having orthogonal eigenvectors $\mathbf{v}_i$ with eigenvalues λ_i.

We have already discussed how to solve this kind of vector differential equation in the Ordinary Differential Equations notes. In short, we substitute $\mathbf{x} = P\mathbf{q}$ into the equation, where $P = (\mathbf{v}_1, \dots, \mathbf{v}_n)$, we get

$$m\ddot{\mathbf{q}} = -K'\mathbf{q}, \tag{2.2}$$

where $K = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$. One would go on to obtain n decoupled equations with variables q_i, known as the normal coordinates, for which the solutions are

$$q_i = Ae^{i\sqrt{\frac{\lambda_i}{m}}t} + Be^{-i\sqrt{\frac{\lambda_i}{m}}t}. \tag{2.3}$$

Therefore the solution for $\mathbf{x}$ is

$$\mathbf{x} = \mathbf{v}_1\left(A_1e^{i\sqrt{\frac{\lambda_1}{m}}t} + B_1e^{-i\sqrt{\frac{\lambda_1}{m}}t}\right) + \cdots + \mathbf{v}_n\left(A_ne^{i\sqrt{\frac{\lambda_n}{m}}t} + B_ne^{-i\sqrt{\frac{\lambda_n}{m}}t}\right). \tag{2.4}$$

Alternatively, we can guess $\mathbf{x} = \mathbf{v}e^{i\omega t}$ to get[1]

$$m\omega^2\mathbf{v} = K\mathbf{v}, \tag{2.5}$$

[1]This method can be interpreted as separation of variables, which we will use to solve the wave equation, or just by observing the general solution obtained earlier are linear combinations of $\mathbf{v}e^{i\omega t}$.

which shows that $\mathbf{v}_i$ are the eigenvectors with eigenvalues $\lambda_i = m\omega_i^2$, which gives the relative amplitudes of each masses in a certain mode of oscillation. One would go on to find the n eigenvalues $\lambda_i = m\omega_i^2$ and the corresponding n eigenvectors and thus getting the same result as above.

Example: Normal Modes.

Question: Consider a double pendulum consisting of two masses suspended by identical massless rods of length ℓ. The upper and lower masses are m and αm respectively, where α is a constant. With x_1 and x_2 denoting the horizontal displacements of the upper and lower masses respectively, solve for the frequencies of the normal modes of the system in terms of α. If $\alpha = 1$ calculate the ratio of the amplitudes of the two displacements.
Now assume that $\alpha \ll 1$ so that quadratic terms in alpha can be neglected. Determine the value that α should take for the lower freqeuncy mode to have one quarter of the frequency of a single pendulum of length ℓ and mass m. Describe qualitatively what happens to the higher-frequency normal mode as $\alpha \to 0$.

Solution: The equations of motions of the system can be written as

$$\begin{pmatrix} \ddot{x}_1 \\ \ddot{x}_2 \end{pmatrix} = -\frac{g}{\ell}\begin{pmatrix} 1+2\alpha & -\alpha \\ -\alpha & \alpha \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}. \tag{2.6}$$

Substituting $(x_1, x_2) = (X_1, X_2)e^{i\lambda\sqrt{g/\ell}t}$, we find

$$\begin{pmatrix} -1-2\alpha+\lambda^2 & \alpha \\ \alpha & -\alpha+\lambda^2 \end{pmatrix}\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \mathbf{0}. \tag{2.7}$$

Requiring the detemrinant to be zero we get

$$\omega_{1,2}^2 = \left(\frac{1+3\alpha}{2} \pm \frac{\sqrt{5\alpha^2+2\alpha+1}}{2}\right)\frac{g}{\ell}. \tag{2.8}$$

For $\alpha = 1$, $\lambda^2 = 2 \pm \sqrt{2}$, and we find the eigenvectors

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1-\sqrt{2} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1+\sqrt{2} \end{pmatrix} \tag{2.9}$$

for plus and minus respectively.
For the low frequency to be a quarter of that of a simple pendulum we require that

$$\begin{gathered} \frac{1}{4}\sqrt{\frac{g}{\ell}} = \left(\left(\frac{1+3\alpha}{2} \pm \frac{\sqrt{5\alpha^2+2\alpha+1}}{2}\right)\frac{g}{\ell}\right)^{1/2} \\ 7+24\alpha = 8\sqrt{1+2\alpha+5\alpha^2} \approx 8(1+\alpha) \implies \alpha = \frac{1}{16}. \end{gathered} \tag{2.10}$$

Note that we should not square both sides and then omit the $(24\alpha)^2$ and $5\alpha^2$ terms, since then we would mix some second order correction into account. To properly take the second order correction into account we can sipmly expand the square root

up to second order. The moral is that when dealing with small angles never square the equation.
Physically when the lower mass becomes negligible ($\alpha \to 0$), the high-frequency normal mode sipmly tends to the natural freqeuncy of the upper pendulum along $\sqrt{g/\ell}$. The lower mass still follows with an amplitude ration of $X_2/X_1 = 2$, but its inertia is negligible so it does not affect the frequency.

2.1.2 Unequal Masses

If the masses in the system are unequal, then the equation becomes

$$M\ddot{\mathbf{x}} = -K\mathbf{x}. \tag{2.11}$$

The generalized eigenvalues and eigenvectors can be found by solving

$$\det(K - \lambda_i M) = 0, \tag{2.12}$$

where the eigenvalues are real, and orthogonal, in the sense that if $\lambda_i \neq \lambda_j$, then $\mathbf{v}_i^T M \mathbf{v}_j = 0$. The only difference is that the solution is now

$$\mathbf{x} = \mathbf{v}_1 \left(A_1 e^{i\sqrt{\lambda_1}t} + B_1 e^{-i\sqrt{\lambda_1}t} \right) + \cdots + \mathbf{v}_n \left(A_n e^{i\sqrt{\lambda_n}t} + B_n e^{-i\sqrt{\lambda_n}t} \right). \tag{2.13}$$

As the general solution is still linear combination of $\mathbf{v}e^{i\omega t}$, guessing it directly to find ω still works.

Enforcing $\mathbf{x}$ to be real gives $A_i = B_i^*$. Assuming to take the real part we drop off half of the terms and we have

$$\mathbf{x} = \sum_{i=1}^{n} A_i e^{i\sqrt{\lambda_i}t} \mathbf{v}_i. \tag{2.14}$$

A more detailed explanation as to why this simplified form contains the same information is explained in section 2.3.3.

However, despite its simplicity this form is still not so useful due to it invovling complex numbers. To understand the solution and to apply and boundary and initial conditions more easily we take the real part to get the real solution

$$\mathbf{x} = A_1 \mathbf{v}_1 \cos(\omega_1 t + \varphi_1) + \cdots + A_n \mathbf{v}_n \cos(\omega_n t + \varphi_n), \tag{2.15}$$

which can be derived from both eqs. (2.14) and (2.15) directly, again showing that the "solution must be real" reasoning is the same as taking the real part of one of the solutions.

Different terms corresponds to different normal modes, a pattern of unforced oscillation in which all parts of the system oscillate a single frequency with a fixed relative phase and amplitude. If the initial condition matches an eigenvector $\mathbf{v}_i$ corresponding to the eigenfrequency ω_i, *i.e.*, $\mathbf{x}(0) = C\mathbf{v}_i$ and $\dot{\mathbf{x}}(0) = \mathbf{0}$,[2][3] then the masses will oscillate together at the same, single frequency ω_i.

If one of the eigenvalue is $\omega^2 = 0$, then the corresonding solution should be $x_1 = x_2 = \cdots = x_n = X_{\text{c.m.}} + V_{\text{c.m.}} t$.

2.1.3 Beats

Beating occurs when the normal mode frequencies are close to each other. For a two equal masses system, each connected to a wall with a spring with spring constant k and connected to each other with a spring with spring constant κ, the normal mode frequencies and their corresponding eigenvectors are

$$\omega_s = \sqrt{\frac{k}{m}}, \quad \omega_f = \sqrt{\frac{k+2\kappa}{m}} \text{ and } \mathbf{v}_s = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \mathbf{v}_f = \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \tag{2.16}$$

The general solution for $\mathbf{x}(t) = (x_1(t), x_2(t))$ is therefore

$$\mathbf{x} = A_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} \cos \omega_s t + A_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} \sin \omega_s t + B_1 \begin{pmatrix} 1 \\ -1 \end{pmatrix} \cos \omega_f t + B_2 \begin{pmatrix} 1 \\ -1 \end{pmatrix} \sin \omega_f t. \tag{2.17}$$

If $\kappa \gg k$, then we have the strong coupling case. Since the middle spring is very stiff, the masses behave nearly as one rigid body in the lower-frequency (symmetric) mode with frequency $\omega_s = \sqrt{k/m}$, while in the high-frequency (antisymmetric) mode they oscillate against the stiff coupling spring at $\omega_f = \sqrt{2\kappa/m}$.

If $\kappa \ll k$, then we have the weak coupling case where $\omega_f \approx \omega_s = \sqrt{k/m}$. Since the middle spring is very soft, so the two masses are almost uncoupled and each mass oscillates essentially on its own end-spring but energy is transferred between the two masses via the weak coupling spring, a phenomenon known as beating.

he initial condition $\mathbf{x}(0) = (0, A)$ and $\dot{\mathbf{x}}(0) = (0, 0)$ we have

$$\begin{cases} x_1(t) = A(\cos \omega_s t - \cos \omega_f t)/2 = A \sin \Omega t \sin \epsilon t, \\ x_2(t) = A(\cos \omega_s t + \cos \omega_f t)/2 = A \cos \Omega t \cos \epsilon t, \end{cases}, \quad \Omega \equiv \frac{\omega_s + \omega_f}{2} \text{ and } \epsilon \equiv \frac{\omega_f - \omega_s}{2}. \tag{2.18}$$

The solutions are plotted in fig. 2.1. We can see that the rapid Ω oscillation is enveloped by the slow ϵ oscillation. An important point to note is that the beat freqeuncy is the frequency of the "bubbles" in the envelope curve, so this frequency is $\omega_{\text{beat}} = 2\epsilon = \omega_f - \omega_s$, but not ϵ.

[2]Here C is just a scaling constant, since only the relative positions of the masses matter but not the absolute positions.

[3]Of course one can also have the relative positions not equal to the eigenvector but have the velocities compensate for it so that the system still exhibits a single normal mode but it is less intuitive and less pratical.

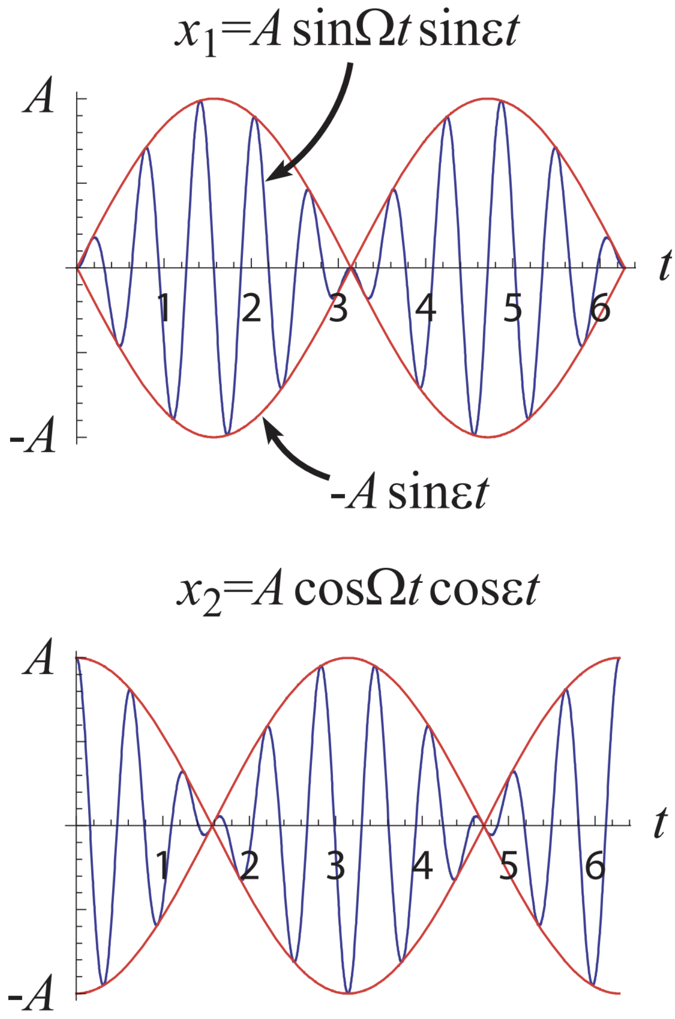


Figure 2.1

Note that there is no restriction on the slope of the $x_1(t)$ and $x_2(t)$ curves at the intersection between the curves $\pm\cos\Omega t$. Note also that as long as the solution has the form $x(t) = A\cos(\omega_1)t + B\cos(\omega_2 t)$ there would be beating even if $A \neq B$, as the cosine terms still go in and out of phase.

2.1.4 Damped and Forced Oscillators

If there is damping, then the equation becomes

$$m\ddot{\mathbf{x}} = -\gamma\dot{\mathbf{x}} - K\mathbf{x}, \tag{2.19}$$

which can be reduced to the normal case via a substitution $\mathbf{x} = \mathbf{y}e^{-\frac{\gamma t}{2}}$. All friction does is reduce the frequency of each normal mode and introduce an overall damping factor. Agian, guessing $\mathbf{x} = \mathbf{v}e^{i\omega t}$ still works.

If there is also a driving force, then the equation becomes

$$m\ddot{\mathbf{x}} = -\gamma\dot{\mathbf{x}} - K\mathbf{x} + \mathbf{F}e^{i\omega t}, \tag{2.20}$$

which adds a particular solution $\mathbf{x}_P = \mathfrak{Re}(\mathbf{C}e^{i\omega t})$, which upon substitution gives

$$\mathbf{C} = \left(K + (i\gamma\omega - \omega^2)\mathbb{I}\right)^{-1}\mathbf{F}. = (P(K' + (i\gamma\omega - \omega^2)\mathbb{I})P^{-1})^{-1}\mathbf{F} = PGP^{-1}\mathbf{F}, \tag{2.21}$$

where $G = \mathrm{diag}((\lambda_i - \omega^2 + i\gamma\omega))^{-1}$. If the driving force frequency is close to one of the normal-node frequencies, say $\omega \approx \omega_1$, then G is dominated by the entry with $i = 1$, so we get

$$\mathbf{C} \approx \mathbf{v}_1 \frac{\mathbf{v}_1^T\mathbf{F}}{i\gamma\omega} \implies \mathbf{x}_P \approx \mathbf{v}_1 \frac{\mathbf{v}_1^T\mathbf{F}}{\gamma\omega}\sin(\omega t), \tag{2.22}$$

and the system is simply in the corresponding normal mode.

2.1.5 Energy Conservation

Multiplying both sides of eq. (2.11) with $\dot{\mathbf{x}}^T$, we get

$$\frac{d}{dt}\left(\frac{\mathbf{x}^T M\mathbf{x}}{2}\right) = -\dot{\mathbf{x}}^T K\mathbf{x} = -\frac{d}{dt}\left(\frac{\mathbf{x}^T K\mathbf{x}}{2}\right), \tag{2.23}$$

where the last equality holds since $K = K^T$ for symmetric K. From the above equation, we conclude that kinetic energy corresponds to the LHS of the above equation and the potential energy the RHS.

In terms of normal coordinates, we have

$$E = \frac{1}{2}(\mathbf{x}^T M\mathbf{x}) + \frac{1}{2}(\mathbf{x}^T K\mathbf{x}) = \frac{1}{2}(\mathbf{q}^T M'\mathbf{q}) + \frac{1}{2}(\mathbf{q}^T K'\mathbf{q}). \tag{2.24}$$

In general, if the eigenvectors are not normalized, we can introduce a matrix $\Gamma = P^T P$ such that instead of M' we have $\Gamma M'$ and instead of K' we have $\Gamma K'$.

In equal masses case,

$$E = \frac{1}{2}\sum_{i=1}^{n}\left(m\dot{q_i}^2 + \lambda_i q_i^2\right). \tag{2.25}$$

Writing the potential energy and kinetic energy in quadratic forms $U = \mathbf{x}^T V\mathbf{x}/2$ and $K = \dot{\mathbf{x}}^T T\dot{\mathbf{x}}/2$ we can use energy conservation to derive the equation of motion

$$\frac{dE}{dt} = \frac{d}{dt}(U + K) = 0 = \dot{\mathbf{x}}^T V\mathbf{x} + \mathbf{x}^T V\dot{\mathbf{x}} + \ddot{\mathbf{x}}^T T\dot{\mathbf{x}} + \dot{\mathbf{x}}^T T\ddot{\mathbf{x}} \tag{2.26}$$

Dividing $\dot{\mathbf{x}}$ and noting that the first and second, and the third and fourth terms are equal, since V and T are symmetric we yield

$$T\ddot{\mathbf{x}} + V\mathbf{x} = 0. \tag{2.27}$$

2.1.6 N Masses

For N masses oscillating in transverse or longitudinal direction, the equation reads

$$m\frac{d^2}{dt^2}\begin{pmatrix}\vdots\\x_{n-1}\\x_n\\x_{n+1}\\\vdots\end{pmatrix}=\begin{pmatrix} & & & \vdots & & & \\ \cdots & k & -2k & k & & & \\ & & k & -2k & k & & \\ & & & k & -2k & k & \cdots \\ & & & & \vdots & & \end{pmatrix}\begin{pmatrix}\vdots\\x_{n-1}\\x_n\\x_{n+1}\\\vdots\end{pmatrix}. \tag{2.28}$$

As ususal, we guess $\mathbf{x}=\mathbf{v}e^{i\omega t}$, but instead of taking the determinant, we look at the n^{th} equation

$$-\omega^2 v_n=\omega_0^2(v_{n-1}-2v_n+v_{n+1}), \tag{2.29}$$

with $v_0=v_{n+1}=0$ to cover $n=1$ or n cases.

It turns out that $v_n=A\sin(n\theta)+B\cos(n\theta)$ is a general solution to the equation, as can proved by induction, and to accomodate for $v_0=v_{n+1}=0$, we require $B=0$ and $\theta=m\pi/(N+1),$ where $m=1,\dots,N.$ Substituting into the above equation, we have

$$\omega_m=2\omega_0\sin\left(\frac{m\pi}{2(N+1)}\right)\quad\text{and}\quad v_n=A\sin\left(nm\pi/(N+1)\right). \tag{2.30}$$

Here n is the label of each mass, m is the label of each normal mode and N is the total number of masses, which also equals to the total number of normal modes.

The complete general solution is therefore

$$\begin{aligned}\mathbf{x}=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}&=A_1\begin{pmatrix}\sin\left(\pi/(N+1)\right)\\\sin\left(2\pi/(N+1)\right)\\\vdots\\\sin\left(N\pi/(N+1)\right)\end{pmatrix}\cos(\omega_1 t+\varphi_1)+A_2\begin{pmatrix}\sin\left(2\pi/(N+1)\right)\\\sin\left(4\pi/(N+1)\right)\\\vdots\\\sin\left(2N\pi/(N+1)\right)\end{pmatrix}\cos(\omega_2 t+\varphi_2)\\&+\dots+A_n\begin{pmatrix}\sin\left(N\pi/(N+1)\right)\\\sin\left(2N\pi/(N+1)\right)\\\vdots\\\sin\left(N^2\pi/(N+1)\right)\end{pmatrix}\cos(\omega_n t+\varphi_n).\end{aligned} \tag{2.31}$$

Again, we have omitted the $-\omega_i$ solution by forcing the solution to be real, concluding that half of the terms can be neglected as long as we take the real part at the end.

When $N\to\infty$, we have

$$\omega_m=\frac{m\pi}{L}\sqrt{\frac{T}{\rho}}=m\omega_1,\quad m\ll N. \tag{2.32}$$

The cases where $m = 1, 2, 3, 4, 5$ for $N = 6$ and the case where $m = N = 18$ are shown in fig. 2.2.

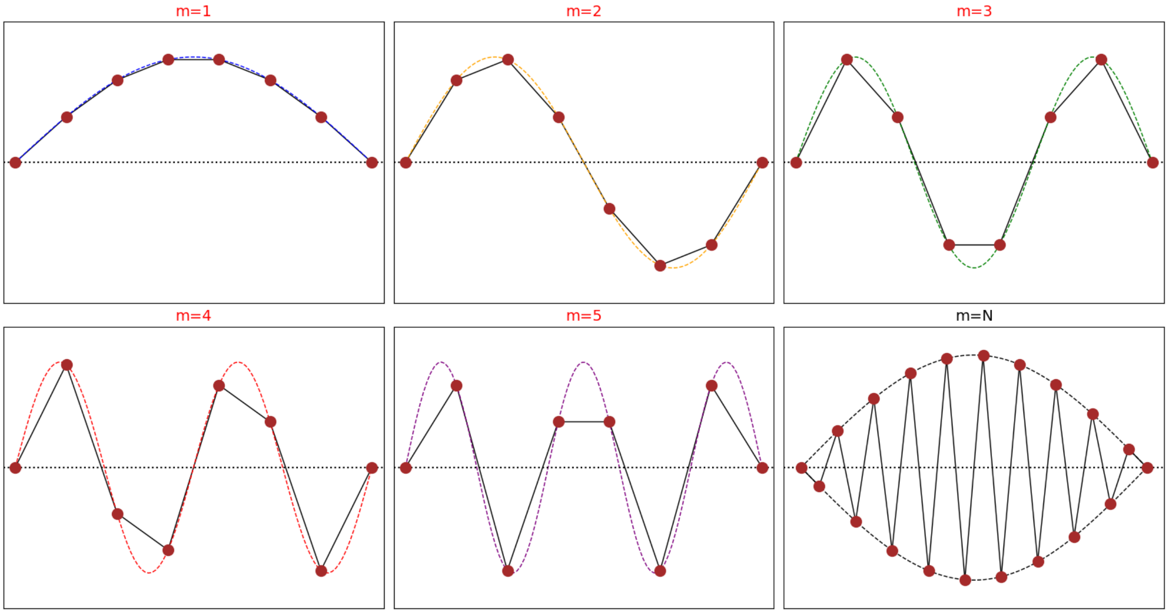


Figure 2.2

Essentially, a system with N particles would have N normal modes. A normal mode is a state that the system is in oscillating at the normal mode frequency. A certain normal mode can be excited by tuning the initial condition in a specific way, normally by releasing each masses at rest each at a specific amplitude. A general initial condition can be written as a linaer combination of the initial conditions of the N normal modes, thus exciting the system into different normal mode by a different extent.

Therefore, the relative amplitudes of the masses at a certain normal mode can be found by considering the continuous cases, and selecting the mass elements to be on the continuous wave at regular interval, as shown in fig. 2.2.

On the other hand, the frequencies of each normal mode can be found by dividing a quarter circle with raidus $2\omega_0$ into $N + 1$ equal intervals, and finding the values of the resulting points. The case for $N = 3$ is illustrated in fig. 2.3.

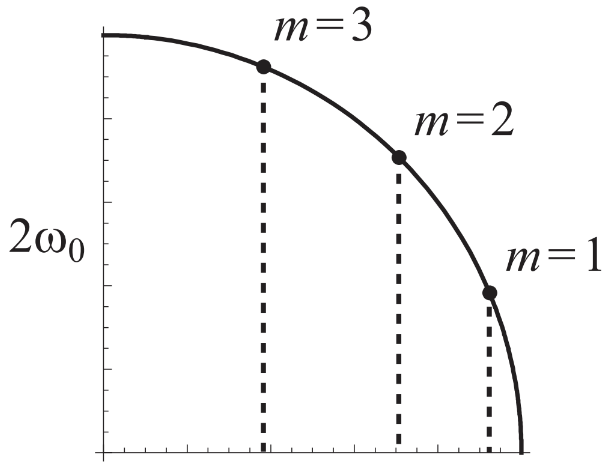


Figure 2.3

In the above discussion, we have restricted ourselves to $1 \leq m \leq N$, but in theory m can take any values. In discrete case this is not a problem, since the frequencies and the relative amplitudes are the same regardless of $m = 3$ or 17. So this ambuiguity is purely mathematical, which can simply resolved by restricted ourselves to the values of m in the range $1 \leq m \leq N$, and affect nothing physical.

However, in the continuous case, this means that there is no way to tell what mode the string is really in if we only look at six equally spaced points. This ambiguity is known as aliasing, or the Nyquist effect.

2.2 Examples of Wave Equations

2.2.1 Longitudinal Oscillations of a String

When $N \to \infty$, then the equations become the wave equation

$$\rho \frac{\partial^2 \xi}{\partial t^2} = E \frac{\partial^2 \xi}{\partial x^2}, \tag{2.33}$$

where $\rho = m/\Delta x, E = k\Delta x$, and we have change the notation for displacement from x to ξ, so that x denotes the equilibrium position.

Alternatively, one can consider the force acting on an infinitesimal mass element to get

$$\rho A \delta x \frac{\partial^2 \xi}{\partial t^2} = \delta F, \quad F = EA \frac{\partial \xi}{\partial x}, \tag{2.34}$$

where the second equation is from the definition of the Young's modulus E to get the same result.

2.2.2 Transverse Osillations of a String

Consider a string with tension T and density μ. Let x the coordinate along the string and $\psi(x)$ be the transverse displacement.

Assuming the slope of the string is small throughout, and consider the horizontal forces acting on a mass element, we can conclude that the tension of the string is constant throughout. If we consider the vertical forces, then we get

$$\mu \frac{\partial^2 \psi}{\partial t^2} dx = T \frac{\partial^2 \psi}{\partial x^2}. \tag{2.35}$$

2.2.3 Acoustic Waves

Refer to fig. 2.4, from mass conservation we have

$$(\rho_0 + \rho_1) A(x + dx + \xi(x + dx, t) - (x + \xi(x, t))) = \rho_0 A dx \implies \frac{\rho_1}{\rho_0} = \frac{1}{1 + \frac{\partial \xi}{\partial x}} - 1 \approx -\frac{\partial \xi}{\partial x}. \tag{2.36}$$

where the relevant force is now

$$F = A(p_0 + p_1), \quad p_1 = \left(\frac{\rho_0 + \rho_1}{\rho_0} \right)^\gamma p_0 \approx \frac{\gamma \rho_1}{\rho_0} \rho_1 \approx -\gamma p_0 \frac{\partial \xi}{\partial x}. \tag{2.37}$$

Newton's second law thus gives

$$\rho_0 \frac{\partial^2 \xi}{\partial t^2} = \gamma p_0 \frac{\partial^2 \xi}{\partial x^2}. \tag{2.38}$$

More generally, the compressibility of a gas κ is defined exactly the same as the Young's modulus E as

$$F = -\kappa A \frac{\partial \xi}{\partial x} \implies \kappa = -V \frac{\partial p}{\partial V}. \tag{2.39}$$

so the wave equation can also be written as

$$\rho \frac{\partial^2 \xi}{\partial t^2} = \kappa \frac{\partial^2 \xi}{\partial x^2}. \tag{2.40}$$

For example, if the exapansion is isothermal instead of adiabetic, then we have

$$\kappa = -V \left(-\frac{RT}{V^2} \right) = p \implies v = \sqrt{\frac{p_0}{\rho_0}}. \tag{2.41}$$

The velocity of the adiabetic compression case can also derived generally since

$$\kappa = -V \left(-\frac{p V^\gamma \gamma}{V^{\gamma+1}} \right) = \gamma p \implies v = \sqrt{\frac{\gamma p_0}{\rho_0}}. \tag{2.42}$$

Since $p_1 \propto \partial \xi / \partial x$, so the excess pressure p_1 also satisfies the same wave equation.

The characteristic impedance is

$$Z = \frac{p_1}{\partial\xi/\partial t} = \frac{\kappa\partial\xi/\partial x}{\partial x/\partial t} = \frac{\kappa k}{\omega} = \sqrt{\rho\kappa}. \tag{2.43}$$

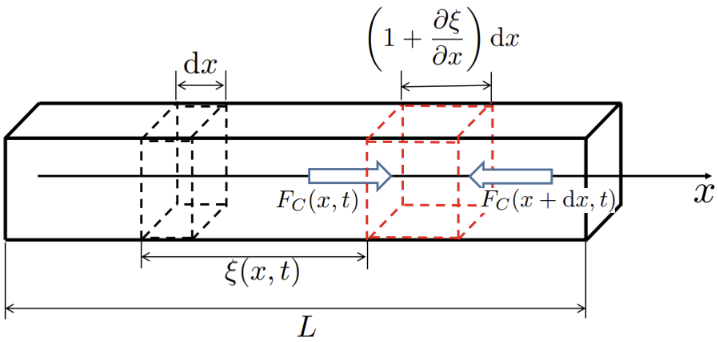


Figure 2.4

2.3 Solution to the Wave Equation

2.3.1 d'Alembert's Solution

General Solution

We start with the wave equation

$$\frac{\partial^2 y}{\partial t} = c^2 \frac{\partial^2 y}{\partial x^2}, \tag{2.44}$$

subject to the initial conditions

$$y(x, t = 0) = y_0(x) \quad \text{and} \quad \frac{\partial y}{\partial t}(x, t = 0) = \dot{y}_0(x). \tag{2.45}$$

Introducing the variables $u = x - ct$ and $v = x + ct$, we can rewrite the time derivative by

$$\partial_t = u_t\partial_u + v_t\partial_v = c(\partial_v - \partial_u) \implies \partial_{tt} = c^2(\partial_{uu} - 2\partial_{uv} + \partial_{vv}). \tag{2.46}$$

Similarly, the spatial derivative can be rewritten as

$$\partial_{xx} = \partial_{uu} + 2\partial_{uv} + 2\partial_{vv}. \tag{2.47}$$

Subtituting into the wave equation we get

$$\frac{\partial^2 y}{\partial u\partial v} = 0. \tag{2.48}$$

The general solution of the wave equation is therefore given by

$$y(x,t) = f(x-ct) + g(x+ct), \tag{2.49}$$

with velocity

$$\frac{\partial y(x,t)}{\partial t} = \dot{y}(x,t) = \frac{\partial f(x-ct)}{\partial t} + \frac{\partial g(x-ct)}{\partial t} = c(g'(x-ct) - f'(x+ct)). \tag{2.50}$$

It is important to note that f and g are functions of single variable, and f' and g' are just normal derivatives.

The initial conditions becomes

$$f(x) + g(x) = y_0(x) \quad \text{and} \quad c\left(g'(x) - f'(x)\right) = \dot{y_0}(x). \tag{2.51}$$

Integrating the second equation above with respect to x, we get

$$f(x) - g(x) = -\frac{1}{c}\int^x \dot{y_0}(s)ds + C. \tag{2.52}$$

We can then solve for $f(x)$ and $g(x)$ to get and obtain the solution of y as

$$y(x,t) = \frac{1}{2}\left(y_0(x-ct) + y_0(x+ct) + \frac{1}{c}\int_{x-ct}^{x+ct} \dot{y_0}(s)ds\right). \tag{2.53}$$

Infinite String

Consider the initial condition

$$y_0(x) = \begin{cases} a\left(1 + x/L\right), & -L \le x < 0, \\ a\left(1 - x/L\right), & 0 \le x < L, \\ 0, & \text{otherwise.} \end{cases} \quad \text{and} \quad \dot{y_0}(x) = 0. \tag{2.54}$$

We find that the solution is

$$y(x,t) = \frac{1}{2}(y_0(x-ct) + y_0(x+ct)), \tag{2.55}$$

as indicated in fig. 2.5. The red line is the actual displacement while the blue and purple lines are the right- and left-traveling waves respectively.

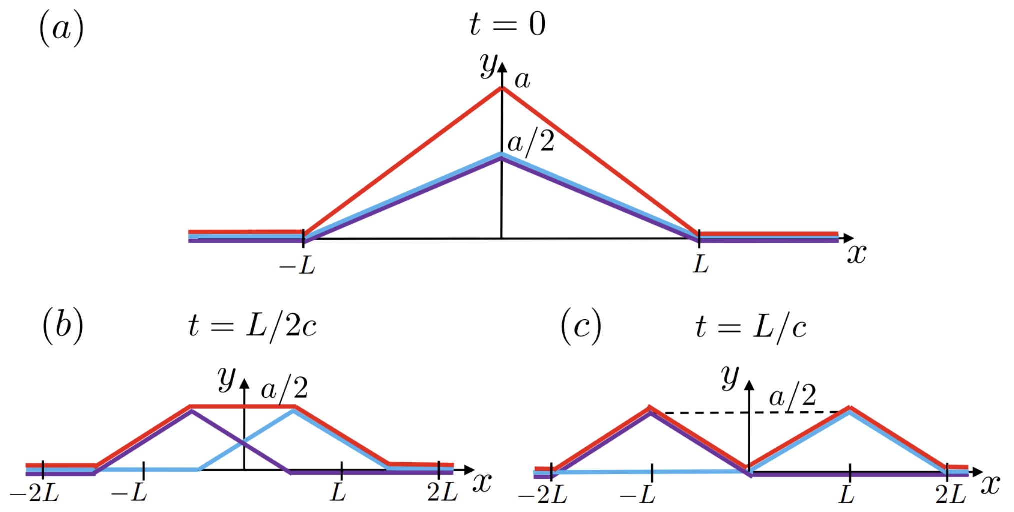


Figure 2.5

If the initial condition is given by

$$y_0(x) = 0 \quad \text{and} \quad \dot{y}_0(x) = \begin{cases} V, & -L \leq x \leq L, \\ 0 & \text{otherwise}, \end{cases} \tag{2.56}$$

then the solution is

$$y(x,t) = \frac{1}{2c}\int_{x-ct}^{x+ct} \dot{y}_0(s)ds = \begin{cases} -VL/2c, & x < L, \\ Vx/2c, & -L \leq x < L, \\ VL/2c, & x \geq L, \end{cases} \tag{2.57}$$

as indicated in fig. 2.6.

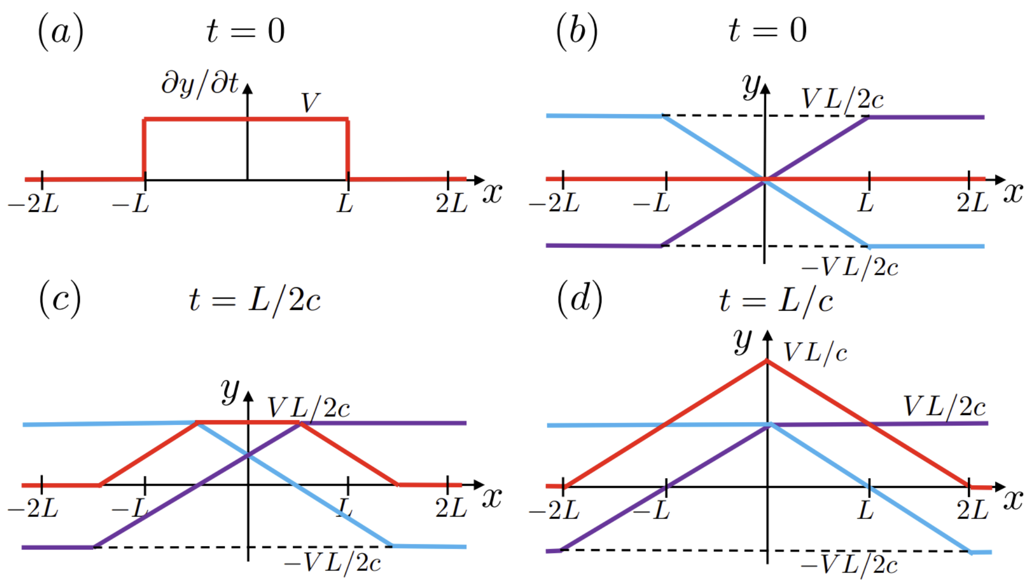


Figure 2.6

If we want a wave of the shape $y_0(x)$ traveling towards positive x, we would need to have

$$y(x,t) = f(x-ct) = y_0(x-ct) \implies \dot{y_0}(x) = \frac{\partial y}{\partial t}(x,0) = -cy_0'(x). \tag{2.58}$$

Semi-infinite String

In all the cases above, we have ignored the ends of the stretched string by assuming that it is infinite. The hidden boundary condition that we have assumed was that $y \to 0$ as $x \to \pm\infty$.

If the string is not infinite, then we have to impose suitable boundary conditions, such as $y = 0$ at fixed points, or $\partial y/\partial x = 0$ at free points, due to the necessary of zero force acting on an infinitesimal mass element.

For example, for the intial condition

$$y_0(x) = \begin{cases} ax/L, & 0 \le x < L, \\ a\left(2 - x/L\right), & L \le x < 2L, \\ 0, & x \ge 2L, \end{cases} \tag{2.59}$$

and

$$\dot{y_0}(x) = cy_0'(x) = \begin{cases} ac/L, & 0 \le x < L, \\ ac/L, & L \le x < 2L, \\ 0, & x \ge 2L, \end{cases} \tag{2.60}$$

which is sketched in fig. 2.7, we would obtain for $f(u)$ and $g(v)$

$$f(u) = \begin{cases} ??, & u < 0, \\ 0, & u \ge 0, \end{cases} \quad \text{and} \quad g(v) = \begin{cases} ??, & v < 0, \\ y_0(v), & v \ge 0. \end{cases} \tag{2.61}$$

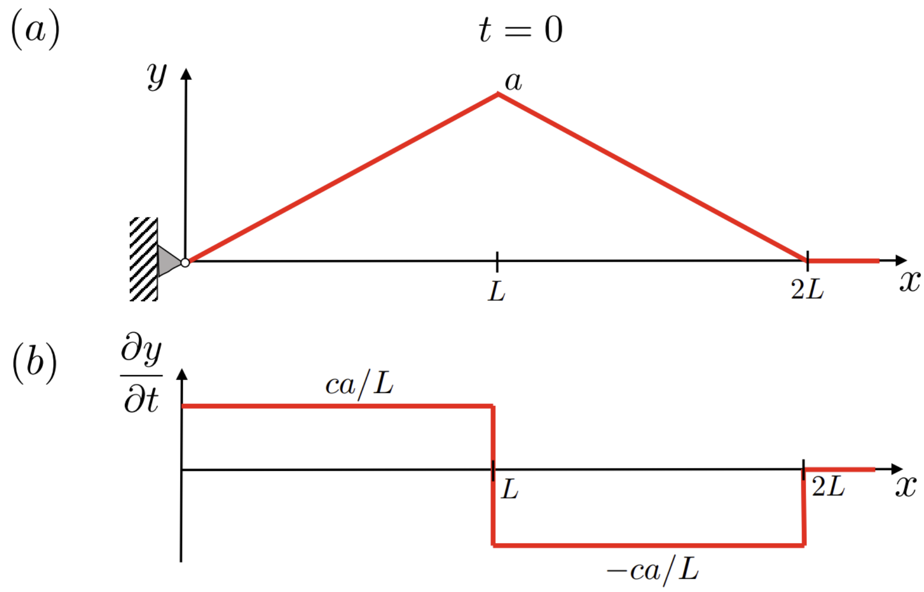


Figure 2.7

It is not until we impose the boundary condition $y = 0$ at $x = 0$ do we get $f(u) = -g(-u)$ and the solution is therefore determined, as shown in fig. 2.8.

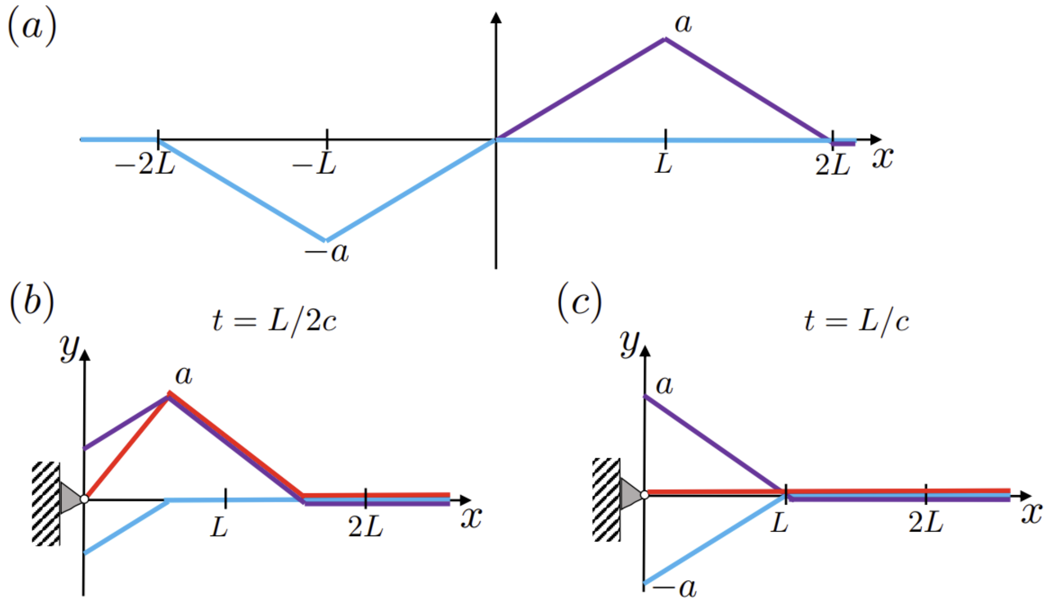


Figure 2.8

If instead we have a free end, so that $\partial y/\partial x = 0$ at $x = 0$ we get $f(u) = g(-u)$ and the solution is shown in fig. 2.9.

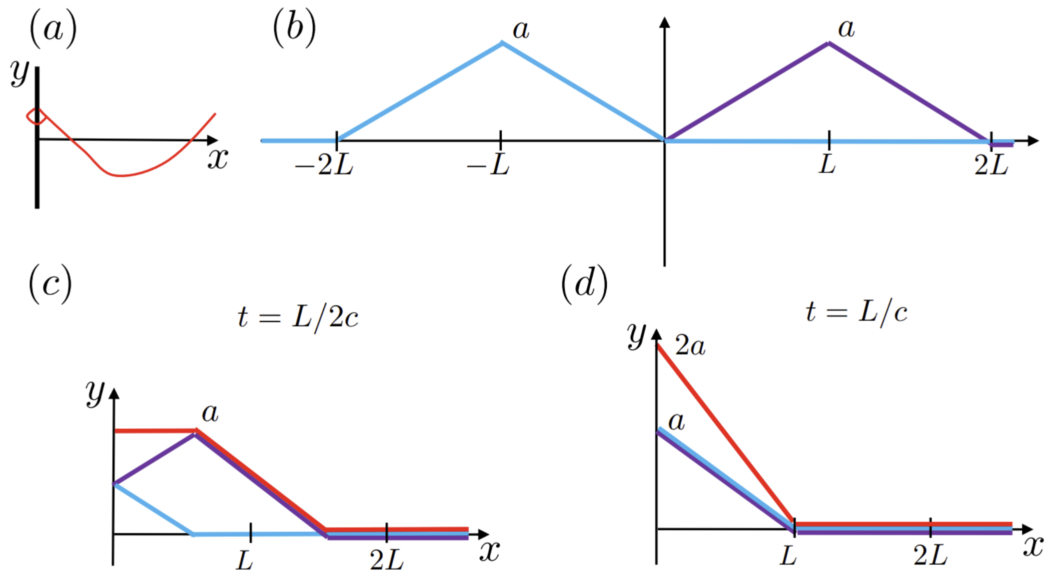


Figure 2.9

Finite String

We will now consider a finite string fixed at both ends at $x = 0$ and L, with the initial conditions

$$y_0(x) = \begin{cases} ax/L, & 0 \leq x < L, \\ a(2 - x/L), & L \leq x < 2L, \end{cases} \quad \text{and} \quad \dot{y}_0(x) = cy_0'(x) = \begin{cases} ac/L, & 0 \leq x < L, \\ -ac/L, & L \leq x < 2L. \end{cases} \tag{2.62}$$

Using the d'Alembert's solution, we get

$$f(u) = \begin{cases} ??, & u < 0, \\ 0, & 0 \leq u \leq 2L, \end{cases} \quad \text{and} \quad g(v) = \begin{cases} y_0(v), & 0 \leq v < 2L, \\ ??, & v \geq 2L. \end{cases} \tag{2.63}$$

Due to the boundary condition at $x = 0$ and L, we can show that $f(u) = -g(-u)$ and $g(v) = -f(4L - v)$ respectively, so we have

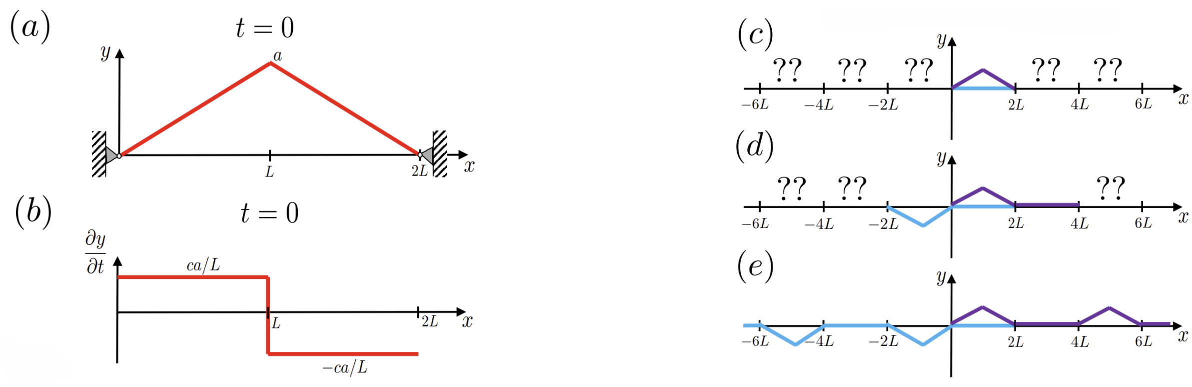
$$f(u) = \begin{cases} ??, & u < 2L, \\ -y_0(-u), & -2L \leq u < 0, \\ 0, & 0 \leq u \leq 2L, \end{cases} \quad \text{and} \quad g(v) = \begin{cases} y_0(v), & 0 \leq v < 2L, \\ 0, & 2L \leq v < 4L, \\ ??, & v \geq 4L. \end{cases} \tag{2.64}$$

The complete solution can thus be constructed step by step as shown in fig. 2.10.

2.3.2 Separation of Variables

General Approach

Just as how we guess $\mathbf{x} = \mathbf{v}e^{i\omega t}$ as the solution to the equation



(a) (b)

Figure 2.10

$$m\ddot{\mathbf{x}} = -K\mathbf{x}, \tag{2.65}$$

we guess $y(x,t) = X(x)T(t)$[4] to the equation

$$\frac{\partial^2 y}{\partial t^2} = c^2 \frac{\partial^2 y}{\partial x^2}, \tag{2.66}$$

where $c^2\partial^2 y/\partial x^2$ now plays the role of $-K$, thus having a spectrum of infinite eigenvalues ω^2 instead of just finite number of them. Substitution gives

$$\frac{1}{c^2 T}\frac{d^2T}{dt^2} = \frac{1}{X}\frac{d^2X}{dx^2}. \tag{2.67}$$

Since the LHS and RHS are functions of different variables, they can only be equal if they equals to the same constant $-k^2$, where we have require the constant to be negative because we expect oscillating behaviour but not exponentials, which in wome case is what we expect such as in evanescent waves or in matter waves. Substituting, we have

$$\frac{d^2T}{dt^2} = -(kc)^2 T \quad \text{and} \quad \frac{d^2X}{dx^2} = -k^2 X, \tag{2.68}$$

which has the general solution

$$y(x,t) = T(t)X(x) = \left(Ae^{ikct} + Be^{-ikct}\right)\left(Ce^{ikx} + De^{-ikx}\right). \tag{2.69}$$

The $y(x,t)$ we get above is called a stationary wave, or standing wave, or a nomral mode, since it only contains one frequency of oscillation $\omega = ck$.

The complete solution is the linear combination of stationary waves with different wavenumbers k and frequencies $\omega = ck$, *i.e.*,

[4] To be less formal one can also just substitute $y(x,t) = X(x)e^{i\omega t}$ and obtain the general solution $y(x,t) = e^{ikct}(Ce^{ikx} + De^{-ikx})$, but then k need to range from $-\infty \to +\infty$ (excluding $k = 0$), recovering the other half of the solutions given in eq. (2.69).

$$y(x,t) = \sum_{m=1}^{\infty} \left(Ae^{ikct} + Be^{-ikct}\right)\left(Ce^{ikx} + De^{-ikx}\right), \tag{2.70}$$

where the linearity factor K_m is absorbed into A, B, C, D.

As we see there are infinite modes due to infinite amount of particles present, and for each mode there are four undetermined constants, refering to the second order nature of the wave equation and the fact that there are two variables x and t.

Solving the wave equation now becomes a matter of finding the coefficients A, B, C, D to satisfy the boundary conditions and initial conditions.

2.3.3 Complex and Real Waves

To represent a forward propagating wave, the most traditional and general expressions are

$$y(x,t) = A\cos(kx - \omega t + \delta) = A\cos(\omega t - kx - \delta), \quad A \in \mathbb{R}. \tag{2.71}$$

However since trigonometric functions are harder to deal with we usually take the shortcut

$$y(x,t) = \mathfrak{Re}(\tilde{A}e^{i(kx-\omega t)}) = \mathfrak{Re}(\tilde{A}^* e^{i(\omega t - kx)}), \quad \tilde{A} \equiv Ae^{i\delta} \in \mathbb{C}. \tag{2.72}$$

It is customary to omit the $\mathfrak{Re}$ sign and assume that taking the real part of a complex wave is understood, so

$$y(x,t) = Ae^{i(kx-\omega t)} = A^* e^{i(\omega t - kx)}. \tag{2.73}$$

In physics, we use the $kx - \omega t$ notation to denote a forward propagating wave, while engineneers prefer the $\omega t - kx$ notation. Some texts like to write a general forward propagating wave as

$$y(x,t) = Ae^{i(kx-\omega t)} + A^* e^{(\omega t - kx)} \tag{2.74}$$

to make the wave real but this is entirely uneccessary and, in fact, just a more complicated procedure to take the real part of the first term.

Similarly, $-kx - \omega t$ is used to denote a backward propagating wave in physics.

At this point, one might think that the general solution in eq. (2.69) using the separation of variables method can be further simplified, since terms like $Ae^{i(\Lambda ct)} + Be^{-i(\Lambda ct)}$ can be represented by just the first term given that $A = B^*$ since the solution must be real. However, this is not correct since the whole premise of this simplication is that we would take the real part at the end and recover the same real wave as if we did not neglect the second term.

The difference is that in eq. (2.69), we are taking the real part of the product of two complex numbers, which is generally not equal to the product of their real parts, but we also have to consider the product of their imaginary parts, which complicate things even more. Therefore in such cases, it is best to restrict ourselves to the current form of eq. (2.69).

Linear Combination of Stationary Waves

Stationary wave is a wave in which certain points, called nodes remain fixed in space while other points, called antinodes, oscillate with maximum amplitude, so the amplitude of oscillation waries with position. A standing waves stores energy locally and does not transmit any net power. It is the superposition of two waves of the same frequency and amplitude travel in opposite directions in the same medium.

The most general form of a stationary wave with wavenumber k (thus frequency $\omega = kc$) is given by eq. (2.69)

$$\begin{aligned} y(x,t) &= \left(Ae^{ikct} + Be^{-ikct}\right)\left(Ce^{ikx} + De^{-ikx}\right) \\ &= ACe^{ik(x+ct)} + BDe^{-ik(x+ct)} + BCe^{ik(x-ct)} + ADe^{-ik(x-ct)}, \end{aligned} \tag{2.75}$$

which we have shown to be consisting of a forward and a backward travelling wave.

If we enforce $y(x,t)$ is real for all x and t, we have $A = B^* \equiv T_0 e^{i\epsilon}$ and $C = D^* \equiv X_0 e^{i\delta}$, we see that the general solution in eq. (2.69) from the separation of variables is in fact a stationary wave

$$y(x,t) = T_0(e^{i(\Lambda ct+\epsilon)} + e^{-i(\Lambda ct+\epsilon)})X_0(e^{i(\Lambda x+\delta)} + e^{-i(\Lambda x+\delta)}) = X_0T_0\cos(\Lambda ct+\epsilon)\cos(\Lambda x+\delta). \tag{2.76}$$

Note that we do not limit the generality of the solution by enforcing $y(x,t)$ to be real, since we, in fact, live in a real world. If one insist that $y(x,t)$ being complex is more general, it is also fine, however the real initial and boundary conditions make the imaginary part zero, forcing it to be real.

In short, this procedure is not a must, but would help us more intuitively understand the solution, since we are real creatures.

The complete general solution is the linear combination of these stationary waves with different wavenumbers k (thus $\omega = ck$).

The boundary conditions, usually specifying two fixed nodes or antinodes, fix the values of δ and Λ, while the initial conditions, usually the initial displacement and velocity, fix the values of X_0T_0 and ϵ.

Example: Stretched String Fixed at Two Points.

Question: Consider a stretched string fixed at $x = 0$ and L, with initial condition $y_0(x)$ and $\dot{y}_0(x)$ as usual. Find the general solution to the wave equation.

Solution: We have the boundary conditions $X(0) = X(L) = 0$, so

$$C = -D \quad \text{and} \quad kL = m\pi, \qquad m = 1, 2, \ldots \tag{2.77}$$

where negative m is not considered as we can combine the the coefficients of negative m and positive m, while $k = 0$ is not considered since then the general solution given in eq. (2.69) no longer works, but is given by $X(x) = C + Dx$, which gives $C = D = 0$, and is trivial.
If the boundary conditions permit C or D to be non-zero then we would have to take this solution into account as well.
The general solution of the wave equation is therefore

$$y(x,t) = \sum_{m=1}^{\infty} \left(A\cos\left(\frac{m\pi ct}{L}\right) + B\sin\left(\frac{m\pi ct}{L}\right)\right) \sin\left(\frac{m\pi x}{L}\right), \tag{2.78}$$

which is the sum of all possible stationary waves each with its own m (thus ω).
To satisfy the initial conditions $y(x,0) = y_0(x)$ and $\dot{y}(x,0) = \dot{y}_0(x)$, we Fourier decompose $y_0(x)$ and $\dot{y}_0(x)$ into linear combinations of sine functions

$$y_0(x) = \sum_{m=1}^{\infty} Y_m \sin\left(\frac{m\pi x}{L}\right) \quad \text{and} \quad \dot{y}_0(x) = \sum_{m=1}^{\infty} Y'_m \sin\left(\frac{m\pi x}{L}\right), \tag{2.79}$$

and compare them with A and B to get

$$A_m = Y_m \quad \text{and} \quad \frac{m\pi c}{L} B_m = Y'_m. \tag{2.80}$$

Example: A General Boundary Condition.

Question: Consider a stretched string with tension T attached to a vertical spring at $x = 0$ with spring constant K. Find the amplitude ratio of incident and reflected waves.

Solution: The boudary condition is

$$-T\frac{\partial y}{\partial x}(0,t) - Ky(0,t) = 0.^{a} \tag{2.81}$$

Substitute

$$y = Ae^{i(kx+\omega t)} + Be^{i(kx-\omega t)}, \tag{2.82}$$

we have

$$\frac{A}{B} = e^{i\varphi}, \quad \varphi = \pi + 2\tan^{-1}\left(\frac{kT}{K}\right). \tag{2.83}$$

---

[a] The negative signs can be explained by using the vector equation $\mathbf{F}_{\text{net}} = \mathbf{F}_{\text{tension}} = \mathbf{F}_{\text{spring}}$, where $\mathbf{F}_{\text{tension}} = -T\partial y/\partial x$ is due to the fact with the slope of the string is positive then the force is negative, while $\mathbf{F}_{\text{spring}} = -Ky$ is due to the fact with the displacement of the string is positive then the force is negative.

Example: Stretched Strings with Different Density or Tension.

Question: Consider a stretched strings composed of two strings of different linear density μ_1 and μ_2, which are tied together at $x = 0$. Find the general solution to the wave equations.
If instead of a continuous string with different density we have two strings with different tension, tied together at $x = 0$ by a massless ring encircling a frictionless pole (so that the change in tension is balanced by the normal reaction), then how would the solutions change?

Solution: The two strings each satisfies their own wave equation, with different wave speed $c_1 = \sqrt{T/\mu_1}$ and $c_2 = \sqrt{T/\mu_2}$, where the tension T remains the same due to balance of the horizontal force on the infinitesimal mass element at $x = 0$.
The boundary conditions are

$$y_1(0,t) = y_2(0,t) \quad \text{and} \quad \frac{\partial y_1}{\partial x}(0,t) = \frac{\partial y_2}{\partial t}(0,t). \tag{2.84}$$

We substitute

$$y_1 = Ae^{i(k_1x+\omega_1t)} + Be^{i(k_1x-\omega_1t)} \quad \text{and} \quad y_2 = Ce^{i(k_2x-\omega_2t)}, \tag{2.85}$$

Imposing the boundary conditions we get

$$\omega_1 = \omega_2, \quad A + B = C \quad \text{and} \quad k_1(A - B) = k_2C. \tag{2.86}$$

Solving for reflection and transmission coefficients $r \equiv B/A$ and $t \equiv C/A$, we have

$$r = \frac{k_1 - k_2}{k_1 + k_2} \quad \text{and} \quad t = \frac{2k_1}{k_1 + k_2}. \tag{2.87}$$

If there are two strings with different tension, then the boundary conditions are modified to be

$$y_1(0,t) = y_2(0,t) \quad \text{and} \quad T_1\frac{\partial y_1}{\partial x}(0,t) = T_2\frac{\partial y_2}{\partial t}(0,t). \tag{2.88}$$

and the reflection and transmission coefficients becomes

$$r = \frac{k_1T_1 - k_2T_2}{k_1T_1 + k_2T_2} \quad \text{and} \quad t = \frac{2k_1T_1}{k_1T_1 + k_2T_2}. \tag{2.89}$$

We usually define the impedance $Z_i \equiv T_i/v_i \propto k_iT_i$, which is the damping constant between the force and velocity, since

$$F_y = T_2\frac{\partial y_2}{\partial x}(0,t) = -\frac{T_2}{v_2}\frac{\partial y_2}{\partial t}(0,t) = -Z_2v_y. \tag{2.90}$$

Example: Stretched Strings connected by a Mass.

Question: Find the general solution to the wave equation if at $x = 0$ the two strings are connnected by a mass M.

Solution: The boundary conditions now becomes

$$y_1(0,t) = y_2(0,t) \;\text{ and }\; T\frac{\partial y_2}{\partial x}(0,t) - T\frac{\partial y_1}{\partial x}(0,t) = M\frac{\partial y_1}{\partial t}(0,t) = M\frac{\partial y_2}{\partial t}(0,t). \quad (2.91)$$

Solving for the reflection and transmission coefficients r and t, we have

$$r = \frac{(k_1 - k_2)T - i\omega^2 M}{(k_1 + k_2)T + i\omega^2 M} = |r|e^{i\theta} \;\text{ and }\; t = \frac{2k_1 T}{(k_1 + k_2)T + i\omega^2 M} = |t|e^{i\phi}. \quad (2.92)$$

The energy of the system is conserved since

$$k_1|r|^2 + k_2|t|^2 = k_1. \quad (2.93)$$

Example: Impedances in Transmission lines.

Question: Refer to fig. 2.11, which shows a system made of inductors and capacitors with L and C being the inductance and capacitance per unit length respectively.

Solution: Consider the piece of the top conductor of length δx, charge δQ accumulates within this piece of conductor due to the difference in currents, so we have $\delta Q = I(x) - I(x + dx)$.
From the definition of capacitance we have

$$\delta V = \frac{1}{C}\delta Q = -\frac{1}{C}\frac{\partial I}{\partial x}\delta t. \quad (2.94)$$

From the definition of inductance we also have

$$\delta V = -L\delta x\frac{\partial I}{\partial t}. \quad (2.95)$$

Combining the two equations yield the wave equation

$$\frac{\partial^2 V}{\partial x^2} = LC\frac{\partial^2 V}{\partial t^2} \;\text{ and }\; \frac{\partial^2 I}{\partial x^2} = LC\frac{\partial^2 I}{\partial t^2}. \quad (2.96)$$

Thus we see that the voltage difference between the lines or the current in the lines corresponds to the displacement of the string in the traditional case.
The impedance of the system is generally given by the push variable (in this case voltage) divided by the flow variable (in this case current), so

$$Z_0 = \frac{V_0}{I_0} = \frac{\omega L}{k} = \sqrt{\frac{L}{C}}, \quad (2.97)$$

where the first equality is justified due to the same equation V and I satisfy.

If the transmission line is now terminated at $x = 0$ by an impedance of Z_T, then the boundary condition is that

$$V(0,t) = Z_T I(0,t). \tag{2.98}$$

Subsituting

$$\begin{aligned} V(x,t) &= Ae^{i(\omega t - kx)} + Be^{i(\omega t + kx)}, \\ Z_0 I(x,t) &= Ae^{i(\omega t - kx)} - Be^{i(\omega t + kx)}, \end{aligned} \tag{2.99}$$

we get the reflection coefficient

$$r = \frac{Z_T - Z_0}{Z_T + Z_0}. \tag{2.100}$$

Therefore the maximum power is transferred to the terminating load if $Z_0 = Z_T$.

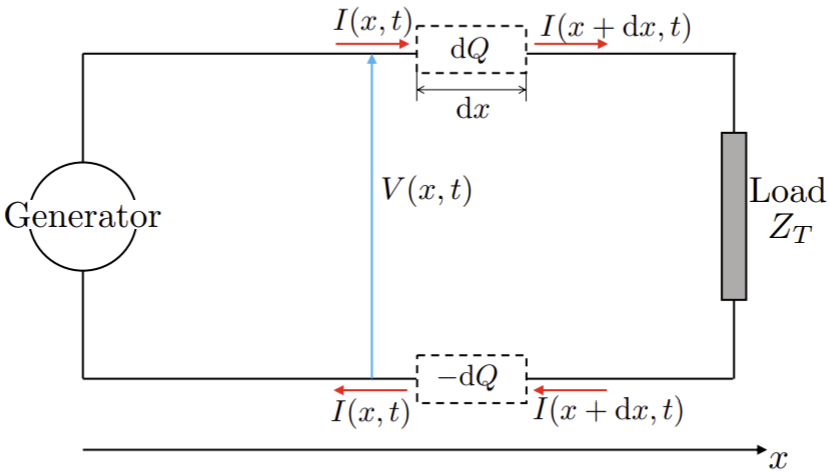


Figure 2.11

Solutions obtaining from separation of variables are no less (and no more) than the d'Alembert's solution for the fact that we can Fourier decompose any function (f and g in this case) to linear combination of exponentials. More specifically, two terms in eq. (2.69) corresponds to the fourier decomposition of $f(x-ct)$ and the remaining two terms $g(x-ct)$.

2.4 Energy in Transverse Oscillation of a String

The length of a stretched string element is given by

$$dl = \sqrt{dx^2 + d\psi^2} \approx dx + \frac{1}{2}\left(\frac{d\psi}{dx}\right)^2 dx, \tag{2.101}$$

so the energy density of the string is

$$\epsilon = \frac{1}{2}\mu\left(\frac{\partial\psi}{\partial t}\right)^2 + \frac{1}{2}T\left(\frac{\partial\psi}{\partial x}\right)^2. \tag{2.102}$$

For simple sinusoidal wave, the average enery density is given by

$$\epsilon_{\text{avg}} = \frac{1}{4}\mu A^2\omega^2 + \frac{1}{4}TA^2k^2 = \frac{1}{2}\mu A^2\omega^2 = \frac{1}{2}TA^2k^2. \tag{2.103}$$

We proceed to study the time evolution of the energy density (*i.e.*, the power density). Firstly, we have

$$\frac{\partial \epsilon}{\partial t} = \mu\frac{\partial y}{\partial t}\frac{\partial^2 y}{\partial t^2} + T\frac{\partial y}{\partial x}\frac{\partial^2 y}{\partial t \partial x} = T\left(\frac{\partial y}{\partial t}\frac{\partial^2 y}{\partial x^2} + \frac{\partial y}{\partial x}\frac{\partial^2 y}{\partial t \partial x}\right) = \frac{\partial}{\partial x}\left(T\frac{\partial y}{\partial x}\frac{\partial y}{\partial t}\right). \tag{2.104}$$

Integrating from $x = 0$ to L, we get the power

$$\frac{dE}{dt} = \left(T\frac{\partial y}{\partial x}\frac{\partial y}{\partial t}\right)\Bigg|_{x=0}^{x=L} = (F_y v_y)\Bigg|_{x=0}^{x=L} \tag{2.105}$$

Thus at each point $F_y v_y$ is the power exerted by the right side to the left side.

Using the d'Alembert's solution, we can rewrite the energy density as

$$\epsilon = \frac{1}{2}\left(\mu c\,(g' - f')^2 + T(f' + g')^2\right) = T(f'^2 + g'^2) = \epsilon_f + \epsilon_g, \tag{2.106}$$

and its rate of change as

$$\frac{\partial \epsilon}{\partial t} = Tc\left(g'^2 - f'^2\right) = c(\epsilon_g - \epsilon_f). \tag{2.107}$$

Integrating from $x = 0$ to L, we have

$$\frac{dE}{dt} = c\,(\epsilon_g - \epsilon_f)\Big|_{x=0}^{x=L}. \tag{2.108}$$

One can refer to the sketches in fig. 2.12 to visualize the terms.

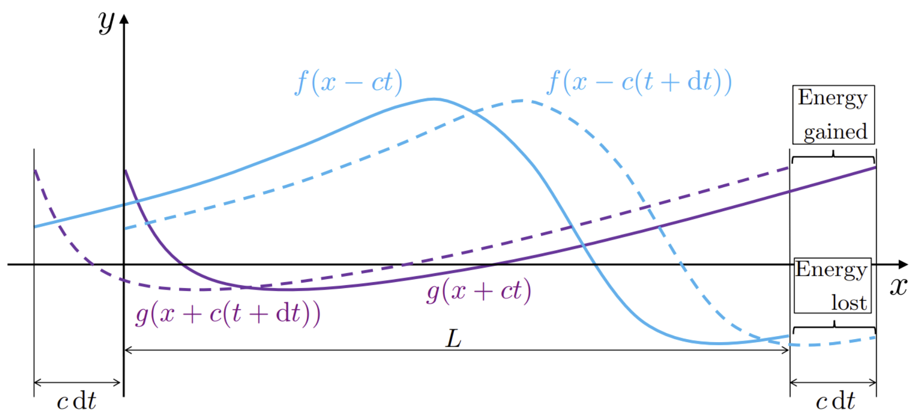


Figure 2.12

Note that the wave equation can be rewritten as

$$\frac{\partial \epsilon_f}{\partial t} + c\frac{\partial \epsilon_f}{\partial x} = 0, \tag{2.109}$$

which means that the energy is conserved for individual waves.

One can also show that if the boundary conditions are simple (*i.e.,* only consist of restrictions on y or $\partial y/\partial x$ but not both), then the total energy is the sum of energy of each mode.

The general solution when the boundary conditions are simple is

$$y(x,t) = A_0 + A_1 t + \sum_{k=1}^{\infty} C_k \cos\left(\sqrt{-\Lambda_k} ct - \varphi_k\right) Q_k(x), \tag{2.110}$$

where the first two terms corresponds to the case where $k = 0$.

The energy of a string is

$$\begin{aligned}
E &= \frac{1}{2}\left(\mu \int_0^L \left(\frac{\partial y}{\partial t}\right)^2 dx + T\int_0^L \left(\frac{\partial y}{\partial x}\right)^2 dx\right) = \frac{1}{2}\int_0^L \left(\mu\left(\frac{\partial y}{\partial t}\right)^2 + Ty\left(\frac{\partial^2 y}{\partial x^2}\right)\right) \\
&= \frac{\mu}{2}\int_0^L \left(A_1 - c\sum_{k=1}^{\infty}\sqrt{-\Lambda_k} C_k \sin\left(\sqrt{-\Lambda_k} ct - \varphi_k\right) Q_k(x)\right)^2 dx \\
&- \frac{T}{2}\int_0^L \left(\sum_{k=1}^{\infty} C_k \cos\left(\sqrt{-\Lambda_k} ct - \varphi_k\right) Q_k(x)\right)\left(\sum_{l=1}^{\infty} \Lambda_l C_l \cos\left(\sqrt{-\Lambda_l} ct - \varphi_l\right) Q_l(x)\right) dx.
\end{aligned} \tag{2.111}$$

where we have performed integration by parts at the second equality of the first line and the boundary terms vanish due to either y or $\partial y/\partial x = 0$ at the boundaries.

If we expand the brackets, we find integrals of the form $A_1 C_k \int_0^L Q_k(x)dx$ and $C_k C_l \int_0^L Q_k(x)Q_l(x)$, but we will show that all the integrals are zero except for integrals in the from $C_k^2 \int_0^L Q_k(x)^2 dx$.

We start by integrating the eigenequation of Q_k after mutiplying both sides by Q_l to get

$$\int_0^L Q_l \frac{d^2 Q_k}{dx^2} dx = -\int_0^L \frac{dQ_l}{dx}\frac{dQ_k}{dx} = \Lambda_k \int_0^L Q_l Q_k dx, \tag{2.112}$$

where we have performed integration by parts at the first equality and the boundary terms vanish for the same reason as above. We can likewise integrate the eigenequation of Q_l after multiplying both sides by Q_k to get the same equation except Λ_k is replaced by Λ_l. By comparing the two equation, we get

$$\int_0^L Q_l Q_k dx = 0, \quad k \neq l. \tag{2.113}$$

This equation also proves that $A_1 \int_0^L Q_k(x)dx$ vanishes since A_1 is non-zero only when Λ_0 is an eigenvalue and hence $Q_0(x)$ is a mode, so either $A_1 = 0$ or $\int_0^L Q_0 Q_k dx = \int_0^L Q_k dx = 0$. After eliminating all the integrals that equals to zero, we have

$$E = \frac{1}{2}\left(\mu L A_1^2 - T\sum_{k=1}^{\infty} \Lambda_k C_k^2 \int_0^L Q_k^2(x)dx\right), \tag{2.114}$$

which is simply the sum of the energy of each mode. If $\int_0^L Q_k^2(x)dx = L/2$, for sinusodial waves, then we can further simply the energy as

$$E = \frac{\mu L A_1^2}{2} - \frac{TL}{4}\sum_{k=1}^{\infty} \Lambda_k C_k^2. \tag{2.115}$$

Example: Energy in Stretched Strings with Different Density.

Question: Prove that the energy is conserved in the example in section 2.3.2 about strings with different density.

Solution: The power transferred from the string 1 to the point at $x = 0$ is

$$P_1 = T\frac{\partial y_1}{\partial x}\frac{\partial y_1}{\partial t} = -T(-k_1 A + k_1 r A)(\omega A + \omega r A) = T k_1 \omega (A^2 - r^2 A^2) = \frac{4\omega^2 A^2 \rho_1 \sqrt{\rho_2} T}{(\sqrt{\rho_1} + \sqrt{\rho_2})^2}, \tag{2.116}$$

while the transmitted power, *i.e.*, power tranferred from the point at $x = 0$ to string 2 is

$$P_2 = T\frac{\partial y_2}{\partial x}\frac{\partial y_2}{\partial t} = T(k_2 t A)(\omega t A) = \frac{4\omega^2 A^2 \rho_1 \sqrt{\rho_2} T}{(\sqrt{\rho_1} + \sqrt{\rho_2})^2}. \tag{2.117}$$

Thus energy is conserved.

2.5 Dispersive Waves

A dispersive medium is one which waves of different wavenumber k (or frequency ω) travel at different phase velocity. A wave-packet composed of sinusodial waves of different wavenumber k thus spread out over time and space, and would not maintain a constant shape. Its dispersion relation $\omega(k)$ gives the angular frequency ω as a function of k. Note that dispersive wvaes does not obey eq. (2.44).

Examples of dispersion includes water waves, where deep-water waves obey $\omega^2 = gk$ and shallow-water waves obey $\omega^2 = gh$; light in glass where the refractive index $n(\lambda)$ depends on wavelength, a prism disperses whitelight into its component colours for wavelength analysis uses this pricniple; and seismic waves where P- and S- waves' speeds depend on frequency.

The general approach to solving dispersive wave equation is still to use the method of separation of variables and to guess $y = Ae^{i(kx-\omega t)}$ to find out the relation between ω and k.

2.5.1 Phase Velocity

The phase velocity

$$v_p(k) = \frac{\omega(k)}{k} \tag{2.118}$$

is the velocity of a single sinusoidal travelling wave with wavenumber k. Different components of the wave with different wavenumbers thus move at different speeds and therefore the wave do not maintain its initial shape while they move.

For the N masses case, the wavelength of the m^{th} normal mode is $\lambda_m = 2L/m \implies m = k_m L/\pi$ and the angular frequency given by eq. (2.30)

$$\omega_m = 2\omega_0 \sin\left(\frac{m\pi}{2(N+1)}\right) = 2\omega_0 \sin\left(\frac{kl}{2}\right), \tag{2.119}$$

where $l = L/(N+1)$ is the distance between each masses. And of course the frequency ω_m reduces to $m\omega_1$ when $kl \ll 1$.

No information is carried in a pure sinusodial wave (except the fact that there is a tone at this frequency, but nothing ever chnages in the signal), so the phase speed is not limited by the speed of light. Imagine a lamp that is always on exactly 100 Watts. You walk into a room and see it – sure, you know there is a lamp, but you don't learn anything else. If the person contrlling that lamp wanted to communicate "A" or "B" to you, leaving it on constantly gives you no way to discriminate. Only when it blinks or chnages brightness can you interprest "short/long flashes" or "bright/dim" as Morse code or some other scheme.

2.5.2 Group Velocity

The group velocity

$$v_g(k) = \frac{d\omega}{dk} \tag{2.120}$$

is the velocity at which a wave-packet envelope which consists of sinusodial waves with different wavenumbers k travels.

A sipmle example for transmitting information is if we combine two waves of similar freqeucies and wavenumbers. Then by superposition the waveform would be

$$y = A\sin\left((k+\delta k)x - (\omega+\delta\omega)t\right) + A\sin\left((k-\delta k)x - (\omega-\delta\omega)t\right) = 2A\cos(\delta kx - \delta\omega t)\sin(kx-\omega t), \tag{2.121}$$

which consists of a rapidly-oscillating wave with wavenumber k and frequency ω (beacuse the wavelength and period are small), travelling at the phase speed $v_p = \omega/k$, and a slowly-oscillating envelope wave with wavenumber δk and frequency $\delta\omega$. The maximum of the envelope wave satisfies

$$\delta k x - \delta\omega t = \text{constant} \implies \frac{dx}{dt} = \frac{\delta\omega}{\delta k}. \tag{2.122}$$

This is precisely the group velocity v_g of those two closely spaced components.

Note that this is not a perfect example to illustrate group velocity since the envelope wave is still a pure sinusoidal wave, so itself also carries no information. But the point stands: in a modulated wave, the underlying rapidly-oscillating wave carries no information and travels at the phase spped $v_p = \omega/k$ while the modulating envelope travels at the group speed $v_g = d\omega/dk$.

Below we will demonstrate the formula for group speed $v_g = d\omega/dk$ more rigorously. Consider the case where the peak of the wave-packet is situated at $x = 0$ at $t = 0$ and the phases of each individual sinusoidal waves with wavenumbers k_i and frequencies ω_i are given by

$$\theta_i = \omega_i t - k_i x + \phi_i. \tag{2.123}$$

For the sinusoidal waves to interfere constructively to give the peak of the wave-packet we want the phase θ to be independent of k, so

$$\frac{d}{dk}(\omega_i t - k_i x + \phi_i) = 0 \implies \frac{x}{t} = \frac{d\omega}{dk}. \tag{2.124}$$

Since the bump consists of wave components with many different values of k, there is an ambiguity about which value of k is the one where we should evaluate $v_g = d\omega/dk$. The general rule is that it is evaluated at the value of k that dominates the bump. That is, it is evaluated at the peak of the Fourier transform of the bump.

You can only assign a single "group velocity" when the pulse remains coherent, *i.e.,* when its spectral components don't wander apart. If the material dispersion is strong enough (or the pulse bandwidth wide enough) that different frequencies outrun each other, the pulse breaks up, and the notion of one group velocity ceases to be meaningful.

2.5.3 Gravity Waves with Surface Tension

Let the equilibrium and the displaced coordinates be (x, y, z) and $(x + \xi(x, y, t), y + \eta(x, y, t), z)$ respectively. The height of the water surface is then given by $h(x, t) = \eta(x, y = 0, t)$. The pressure is $p(x, y, t) = p_a - \rho g y + p_1(x, y, t)$, where p_1 is the excess pressure.

Mass conservation gives

$$dxdydz = \left(\left(1 + \frac{\partial\xi}{\partial x}\right)\left(1 + \frac{\partial\eta}{\partial x}\right) - \frac{\partial\xi}{\partial y}\frac{\partial\eta}{\partial x}\right) dxdydz \implies \frac{\partial\xi}{\partial x} + \frac{\partial\eta}{\partial y} = 0 \tag{2.125}$$

In general, the equation of motion of an infinitesimal volume of water is given by

$$\rho\frac{d^2\xi}{dt^2} = -\frac{\partial p_1}{\partial x} \quad \text{and} \quad \rho\frac{\partial^2\eta}{\partial t^2} = \frac{\partial p_1}{\partial y} - \rho g. \tag{2.126}$$

Differentiating the left equation with respect to x and the right equation with respect to y, one have

$$\rho\frac{\partial^2}{\partial t^2} = \left(\frac{\partial\xi}{\partial x} + \frac{\partial\eta}{\partial y}\right) = -\frac{\partial^2 p_1}{\partial x^2} - \frac{\partial^2 p_1}{\partial y^2} \implies \frac{\partial^2 p_1}{\partial x^2} + \frac{\partial^2 p_1}{\partial y^2} = 0. \tag{2.127}$$

Assuming $h(x,t) = \eta(x, y=0, t) \propto e^{i(kx-\omega t)}$ for typical waves, we have from eq. (2.129) that $p_1(x,y,t) = P_1(y)e^{i(kx-\omega t)}$, which upon substitution gives

$$\frac{d^2P_1}{dy^2} = k^2P_1 \implies p_1(x,y,t) = p_1(x, y=0, t)e^{|k|y}, \tag{2.128}$$

where we have thrown away the another exponential term due to the expectation of $p_1(x, -\infty, t) = 0$.

Refer to fig. 2.13, by balancing the vertical forces acting on an infinitesimal volume of water on the interface, we have

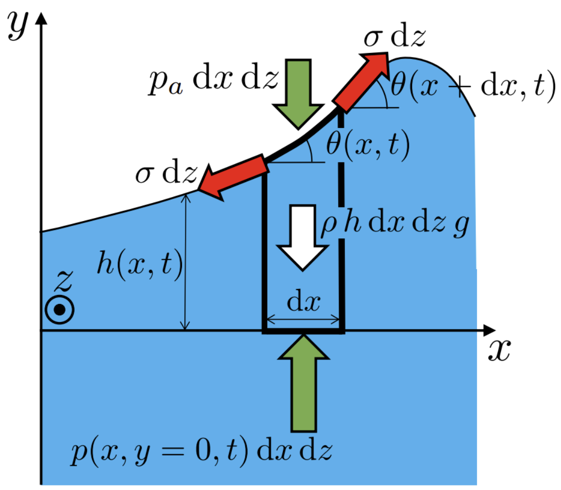


Figure 2.13

$$p(x,y=0,t)dxdz - p_a dxdz - \rho g h(x,t)dxdz + \sigma\sin\theta(x+dx,t)dz - \sigma\sin\theta(x,t)dz = 0$$
$$\implies p(x,y=0,t) = p_a + \rho g h(x,t) + \sigma\frac{\partial^2 h}{\partial x}(x,t), \tag{2.129}$$

where we have neglected the acceleration since it is proportional to $h\,\partial^2 h/\partial t^2$. The horizontal equation of motion is the same as the general case, so

$$a_x(x, y = 0, t) = -\frac{\partial p_1(x, y = 0, t)}{\partial x} = -\left(gh(x, t) - \frac{\sigma}{\rho}\frac{\partial^2 h}{\partial x^2}(x, t)\right) = -ik\left(g + \frac{\sigma k^2}{\rho}\right)h(x, t), \tag{2.130}$$

where we have assumed $h(x, t) = Ae^{i(kx-\omega t)}$ again.

$$a_x(x, y = 0, t) = - \tag{2.131}$$

2.5.4 Wavepackets

Wavepackets are composed of a carrier wave with wavelength k_c and an envelope that can have any shape.

At $t = 0$, the wave packet has the form $y(x, t = 0) = E(x)\cos(k_c x + \varphi)$, where $E(x)$ is the envelope.

2.5.5 Conservation of Waves

For dispersive wave system one can associate a local wave number $\overline{k} = k(x, t)$ and a local angular frequency $\overline{\omega} = \omega(x, t)$ given that $\overline{\omega}$ and $\overline{k}$ are slowly varying functinos of both space and time.

The number of waves per unit time must be equal to the negative number of waves per unit length, so

$$\frac{1}{T} = \frac{1}{\lambda} \implies \frac{\partial \overline{k}}{\partial t} + \frac{\partial \overline{\omega}}{\partial x} = 0. \tag{2.132}$$

Substituting $\omega(k) = \sqrt{a + c^2k^2}$ we also get

$$\frac{\partial \overline{k}}{\partial t} + c_g(\overline{k})\frac{\partial \overline{k}}{\partial x} = 0. \tag{2.133}$$

It then becomes clear that changes in $\overline{k}$ and $\overline{\omega}$ travel at the group velocity, which is also the pseed at which the wave energy propagates.

3

Central Forces

3.1 Reduced Mass

For two particles at $\mathbf{r}_1$ and $\mathbf{r}_2$ we introduce two natural coordintaes (*i.e.,* the center of mass coordinates and the relative position coordinates)

$$\mathbf{R} = \frac{m_1\mathbf{r}_1 + m_2\mathbf{r}_2}{m_1 + m_2} \quad \text{and} \quad \mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2. \tag{3.1}$$

Under central forces (*e.g.,* gravity, collisions, springs) the equation of motions of the two particles are

$$m_1\ddot{\mathbf{r}}_1 = \mathbf{F}(\mathbf{r}_1 - \mathbf{r}_2) \quad \text{and} \quad m_2\ddot{\mathbf{r}}_2 = -\mathbf{F}(\mathbf{r}_1 - \mathbf{r}_2). \tag{3.2}$$

Dividing the first equation by m_1 and the second equation by m_2 and subtracting them we get

$$\mu\ddot{\mathbf{r}} = \mathbf{F}(\mathbf{r}). \tag{3.3}$$

The positions of the particles relative to the center of mass are then

$$\mathbf{r}_1' = \mathbf{r}_1 - \mathbf{R} = \frac{m_2}{m_1 + m_2}\mathbf{r} \quad \text{and} \quad \mathbf{r}_2' = \mathbf{r}_2 - \mathbf{R} = -\frac{m_1}{m_1 + m_2}\mathbf{r}. \tag{3.4}$$

It can be proved easily that the energy of the system can be written as

$$\frac{1}{2}m_1v_1^2 + \frac{1}{2}m_2v_2^2 = \frac{1}{2}M\dot{\mathbf{R}}^2 + \frac{1}{2}\mu\dot{\mathbf{r}}^2, \quad M = m_1 + m_2 \quad \text{and} \quad \mu = \frac{m_1m_2}{m_1 + m_2}. \tag{3.5}$$

As we can see the energy of a system can be splitted into the energy of the center of mass and the energy due to the relative velocity. The former is frame-dependent but the later is not.

This observation urges us to consider the system in a frame in which the center of mass energy vanshies, *i.e.,* $\dot{\mathbf{R}} = 0$, or $\mathbf{R} = \mathbf{C} \equiv 0$, where the energy observed in this frame becomes

$$\frac{1}{2}m_1 v_1^2 + \frac{1}{2}m_2 v_2^2 = \frac{1}{2}\mu\dot{\mathbf{r}}^2. \tag{3.6}$$

In totally inelastic collisions the particles stop moving in the center of mass frame, corresponding to the loss of all kinetic energy.

In the center of mass frame, $\dot{\mathbf{R}} = 0$, which also implies that the total momentum observed in this frame is zero. In other words, the momenta of the two particles must be equal and opposite. This fact is very useful in solving collision problems.

The most classic of all is the problem of 1-dimensional elastic collisions: find the veclocities of the two particles $\mathbf{v}_1'$ and $\mathbf{v}_2'$ after an elastic collision with initial velocities $\mathbf{v}_1$ and $\mathbf{v}_2$. In the center of mass frame, the two particles simply reverse their velocicities after collision, since this is the only way for their velocities to change under the premise that their momenta must be equal and opposite and that the energy is conserved. In symbols, we have

$$(\mathbf{v}_1, \mathbf{v}_2) \overset{\text{c.m.}}{=} (\mathbf{v}_1 - \mathbf{v}_{\text{c.m.}}, \mathbf{v}_2 - \mathbf{v}_{\text{c.m.}}) \overset{\text{collision}}{=} (\mathbf{v}_{\text{c.m.}} - \mathbf{v}_1, \mathbf{v}_{\text{c.m.}} - \mathbf{v}_2) \overset{\text{lab}}{=} (2\mathbf{v}_{\text{c.m.}} - \mathbf{v}_1, 2\mathbf{v}_{\text{c.m.}} - \mathbf{v}_2). \tag{3.7}$$

The coefficient of restitution is defined generally as

$$e = \frac{|\mathbf{v}_{\text{rel}}' \cdot \hat{\mathbf{n}}|}{|\mathbf{v}_{\text{rel}} \cdot \hat{\mathbf{n}}|}, \tag{3.8}$$

where $\hat{\mathbf{n}}$ is the direction of collision, which is a frame-independent quantity since it only depends on the relative velocities of the two objects.

It is very clear in the center of mass frame that if $e = 1$, then $\mathbf{v}_{\text{rel}}' = \mathbf{v}_{\text{rel}}$ and the energy is conserved.

Example: Two-Dimensional Elastic Collision.

Question: A particle of mass m is travelling with initial speed u along the x-axis when it collides elastically and obliquely with a stationary particle of mass M. Find the angle of deflection θ of the mass m in terms of the angle of deflection ϕ of mass M.

Solution: We start with energy and momentum conservation

$$\begin{cases} mu &= mv\cos\theta + MV\cos\phi, \\ mv\sin\theta &= MV\sin\phi, \\ mu^2 &= mv^2 + MV^2. \end{cases} \tag{3.9}$$

We first eliminate V by plugging $V = mv\sin\theta/M\sin\phi$ into the first equation

$$mu = mv\cos\theta + M\left(\frac{mv\sin\theta}{M\sin\phi}\right)\cos\phi \implies v = \frac{u\sin\phi}{\sin(\theta+\phi)}, \tag{3.10}$$

and substitute v and V into the last equation

$$mu^2 = m\left(\frac{\sin^2\phi}{\sin^2(\theta+\phi)}\right)u^2 + M\left(\frac{m\sin\theta}{M\sin\phi}\right)^2\left(\frac{\sin^2\phi}{\sin^2(\theta+\phi)}\right). \tag{3.11}$$

Simplying we get

$$\sin^2(\theta+\phi) = \sin^2\phi + \frac{m}{M}\sin^2\theta. \tag{3.12}$$

Expanding the LHS we have

$$\sin^2\theta\cos^2\phi + \cos^2\theta\sin^2\phi + \frac{1}{2}\sin 2\theta\sin 2\phi = \sin^2\phi + \frac{m}{M}\sin^2\theta \tag{3.13}$$

Writing $\cos^2\theta = 1 - \sin^2\theta$ we have

$$\sin^2\theta\cos 2\phi + \frac{1}{2}\sin 2\theta\sin 2\phi = \frac{m}{M}\sin^2\theta \implies \tan\theta = \frac{M\sin 2\phi}{m - M\cos 2\phi}. \tag{3.14}$$

Example: Two-Dimensional Inealastic Collision.

Question: A particle of mass M_1 is travelling with initial speed u along the x-axis when it collides inelastically and obliquely with a stationary particl of mass M_2. Given that the angle of deflection of mass M_1 is θ, find the speed of mass M_1, *i.e.,* v, following the collision in terms of the coefficient of resitutiton e.
Find the maximum deflection angle θ_{max} given that $M_1 > M_2$.

Solution: Consider the center of mass frame travelling at

$$v_{\text{c.m.}} = \frac{M_1}{M_1+M_2}u \tag{3.15}$$

From the center of mass frame, m and M travels in opposite direction with speed

$$u_{1,\text{c.m.}} = \frac{M_2}{M_1+M_2}u \quad\text{and}\quad u_{2,\text{c.m.}} = \frac{M_1}{M_1+M_2}, \quad M_1u_{1,\text{c.m.}} = M_2u_{2,\text{c.m.}}. \tag{3.16}$$

After the collision, M_1 and M_2 travels in opposite direction with speed $v_{1,\text{c.m.}}$ and $v_{2,\text{c.m.}}$ respectively, with the same constriant $M_1v_{1,\text{c.m.}} = M_2v_{2,\text{c.m.}}$.
Using the definition of the coefficient of restitution we get

$$e = \frac{|\mathbf{v}'_{\text{rel}}|\cdot\hat{\mathbf{n}}}{|\mathbf{v}_{\text{rel}}|\cdot\hat{\mathbf{n}}} = \frac{v_{1,\text{c.m.}} + v_{2,\text{c.m.}}}{u_{1,\text{c.m.}} + u_{2,\text{c.m.}}} = \frac{v_{1,\text{c.m.}}}{u_{1,\text{c.m.}}} = \frac{v_{2,\text{c.m.}}}{v_{2,\text{c.m.}}}. \tag{3.17}$$

Considering the vector addition triangle of M_1 shown in fig. 3.1 we then get

$$v_{1,\text{c.m.}}^2 = v^2 + v_{\text{c.m.}}^2 - 2v_1v_{\text{c.m.}}\cos\theta, \tag{3.18}$$

so we get

$$v = \frac{M_1}{M_1 + M_2} \left(\cos\theta \pm \sqrt{\cos^2\theta - 1 + e^2 \left(\frac{M_2}{M_1} \right)} \right) u. \tag{3.19}$$

For $M_1 = M_2$ and $e = 1$, we have

$$v = \frac{1}{2}(\cos\theta \pm |\cos\theta|). \tag{3.20}$$

For $0 \leq \theta \leq \pi/2$, we choose the positive root for the solution to make sense, so $v = u\cos\theta$. For $\pi/2 \leq \theta \leq \pi$ neither the positive root nor the negative root make any sense, since both gives a negative answer, even through v is a straightly non-negative quantity (it is the norm of the velocity vector of M_1 in the lab frame), so $v = u\cos\theta < 0$. This means that the maximum deflection angle is $\pi/2$ for equal mass elastic collision.
For $M_2 \gg M_1$ and $e = 1$, we have

$$v = \pm 1. \tag{3.21}$$

However, as we have mentioned v must be non-negative, $v = 1$ is the only possible solution, which means the M_1 simply bounces off M_2 with the same speed, which makes perfect sense.
For $e = 0$, the only possible value of θ which makes the terms under the square root non-negative is $\theta = 0$, in which case we get

$$v = \frac{M_1}{M_1 + M_2} u, \tag{3.22}$$

which also make sense since M_1 and M_2 now sticks together and moves at the same speed.
From fig. 3.1 we see that as the angle of deflection in the center of mass frame ϕ varies from 0 to 2π, the angle of deflection in the lab frame increases from 0 to θ_{max} then subsequently decreases to 0. By simple geometry we find

$$\sin\theta_{\text{max}} = \frac{v_{1,\text{c.m.}}}{v_{\text{c.m.}}} = \frac{M_2}{M_1}. \tag{3.23}$$

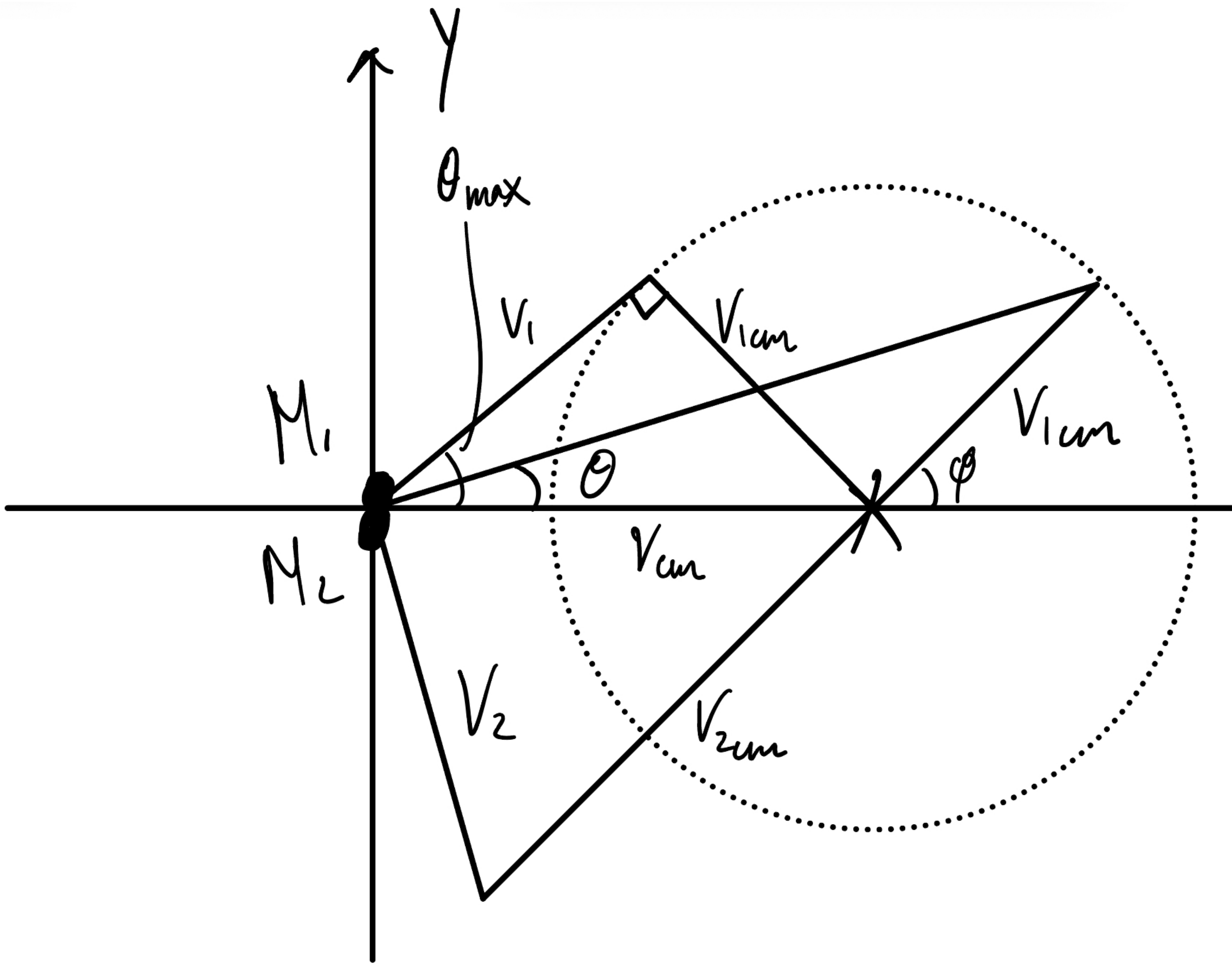


Figure 3.1

Example: Superball.

Question: A spherical superball is perfectly elastic, incompressible and rough. In the figure below a superball of radius a is spinning with an angular velocity Ω_1 as shown. The superball hits a rough, horizontal surface while travelling with components of velocity u_1 normal to the surface and v_1 parallel to the surface. Find the linear and angular velocities u_2, v_2 and Ω_2 with which it rebounds.

Solution: Since the collision is elastic we have

$$e = \left| \frac{\mathbf{v}'_{\text{rel}} \cdot \hat{\mathbf{n}}}{\mathbf{v}_{\text{rel}} \cdot \hat{\mathbf{n}}} \right| = 1 \implies u_2 = -u_1. \tag{3.24}$$

From the linear and angular impulse theorem we have

$$J = -ft = m(v_2 - v_1) \quad \text{and} \quad aJ = -aft = I(\Omega_2 - \Omega_1). \tag{3.25}$$

From the conservation of energy we have

$$m(v_1^2 + u_1^2) + I\Omega_1^2 = m(v_2^2 + u_2^2) + I\Omega_2^2. \tag{3.26}$$

Eliminating J we get

$$\begin{cases} v_1 + a\Omega_1 & = -(v_2 + a\Omega_2), \\ 5v_1 - 2a\Omega_1 & = 5v_2 - 2a\Omega_2, \end{cases} \tag{3.27}$$

where the first equation shows that the tangential velocity of the contact point is also reversed, so the “tangential coefficient of restitution” is also 1.
Solving the system of equation gives

$$v_2 = \frac{3}{7}v_1 - \frac{4}{7}a\Omega_1 \quad \text{and} \quad \Omega_2 = \frac{10}{7}\frac{v_1}{a} + \frac{3}{7}\Omega_1. \tag{3.28}$$

3.2 Effective Potential

Considering the reduced mass concept introduced in the last section we can recast the two-dimensional two-bodies central force problem into a two-dimensional one-body central force problem.

For a central froce $F(r)$, using the conservation of angular momentum

$$J = mr^2\ddot{\theta}, \tag{3.29}$$

the energy of a particle can be written as

$$E = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) + U(r) = \frac{1}{2}m\dot{r}^2 + \frac{J^2}{2mr^2} + U(r), \tag{3.30}$$

where J is the angular momentum about the origin and $U(r)$ is the potential energy.

We have reduced a two-dimensional problem into a one-dimensional one, where we can consider another particle moving in the radial direction experiencing the effective potential

$$U_{\text{eff}} = \frac{J^2}{2mr^2} + U(r), \tag{3.31}$$

where the kinetic energy associated with the angular motion is combined into the potential energy.

For the typical $F(r) \propto -1/r^2$, the effective potential is shown in fig. 3.2.

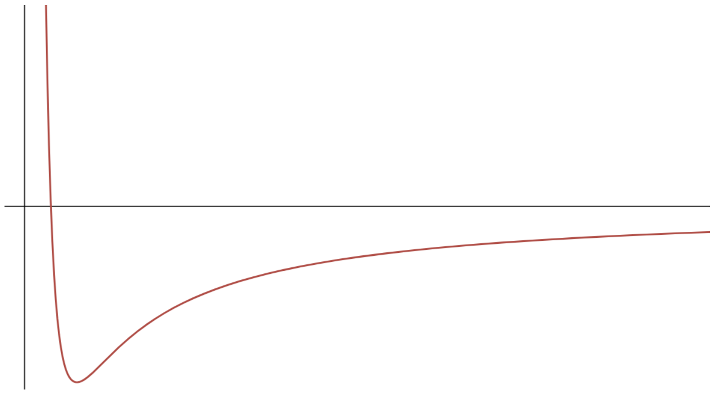


Figure 3.2

An orbit is bounded if the total energy is smaller than the value of its effective potential at infinity, so this case $E > 0$ corresponds to unbounded orbits (hyperbola), and $E < 0$ corresponds to bounded orbits (ellipse or cirlce) and $E = 0$ corresponds to the marginal case (parabola).

For a bounded orbit the energy lies between the minimum point of U_{eff} and the x-axis. Two two points where they coincide are the perigee and apogee, since at those points we have

$$E = \frac{1}{2}m\dot{r}^2 + U_{\text{eff}} = U_{\text{eff}} \implies \dot{r} = 0. \tag{3.32}$$

By this logic we have the minimum point of U_{eff} being the circular orbits, since the radius remain fixed, we have $\dot{r} = 0$ at all times.

Taking the time derivative we get

$$m\ddot{r} = \frac{J^2}{mr^3} - \frac{dU}{dr}, \tag{3.33}$$

which is the same is directly writing down the Newton's second law

$$-\frac{dU}{dr} = m(\ddot{r} - r\dot{\theta}^2). \tag{3.34}$$

The tangential angular velocity plus the gravitational potential consitutes the effective potential, which further reduces the two-dimensional one-body central force problem into a one-dimensional one-body problem.

Zeros of $E - V_{\text{eff}}(r)$ give perigee and apogee. Minima of $V_{\text{eff}}(r)$ correspond to stable circular orbits. The "centrifugal barrier" term $J^2/2mr^2$ explains why the particle cannot reach $r = 0$ unless $J = 0$.

Example: Power Laws (1).

Question: Consider a point mass m moving with angular momentum L

under the action of a potential $V(r) = \beta r^k$. Find the radius r_0 of a circular orbit in terms of β, k, m and L.
The mass is disturbed slightly, such that the radius oscillates about r_0. Find the frequency of these oscillations ω_r and the ratio of this frequency to the orbital frequency ω_θ at r_0.
Use the result to sketch the shape of the perturned orbits for $k = 2, 7$ and $7/4$.

Solution: The effective potential is

$$V_{\text{eff}} = \frac{L^2}{2mr^2} + \beta r^k. \tag{3.35}$$

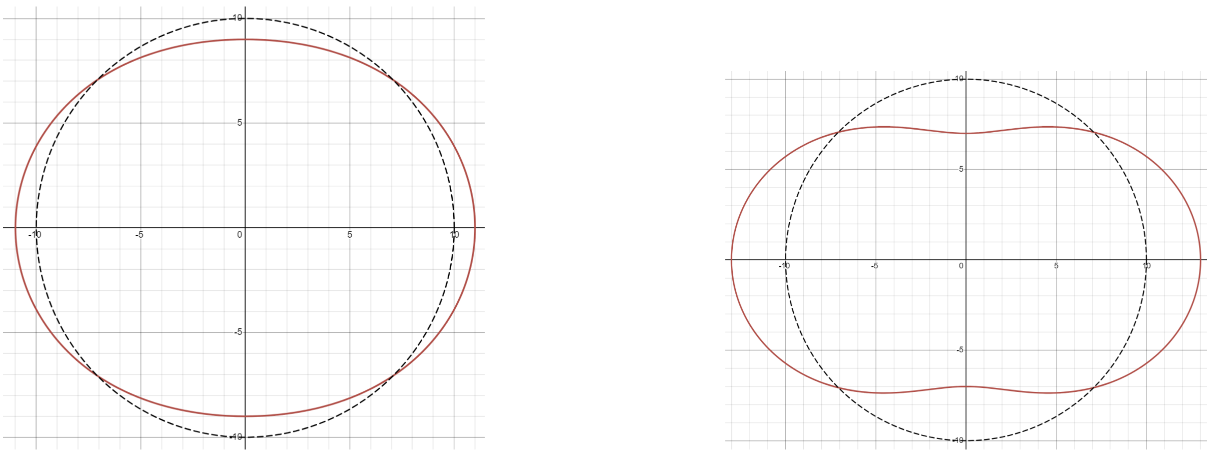
To find the circular orbit we set its derivative to be zero

$$\frac{dV_{\text{eff}}}{dr} = 0 \implies r_0 = {\frac{L^2}{\alpha km}}^{1/(k+2)}. \tag{3.36}$$

The angular frequency of the oscillation is simply

$$\omega = \sqrt{\frac{V''_{\text{eff}}(r_0)}{m}} = \sqrt{k+2}\,\omega_\theta. \tag{3.37}$$

The sketches for $k = 2, 7$ and $7/4$ can be found in fig. 3.3 and fig. 3.4. Figure 3.3a and fig. 3.3b shows the case for $k = 2$, where the amplitude of the raidal disturbation is set to be $0.1r_0$ and $0.3r_0$ respectively, while fig. 3.4a and fig. 3.4b show the case for $k = 7$ and $7/4$ respectively. Note that for small enough amplitude of raidal oscilaltion and for $k = 2$ the orbit is an ellipse, as expected.

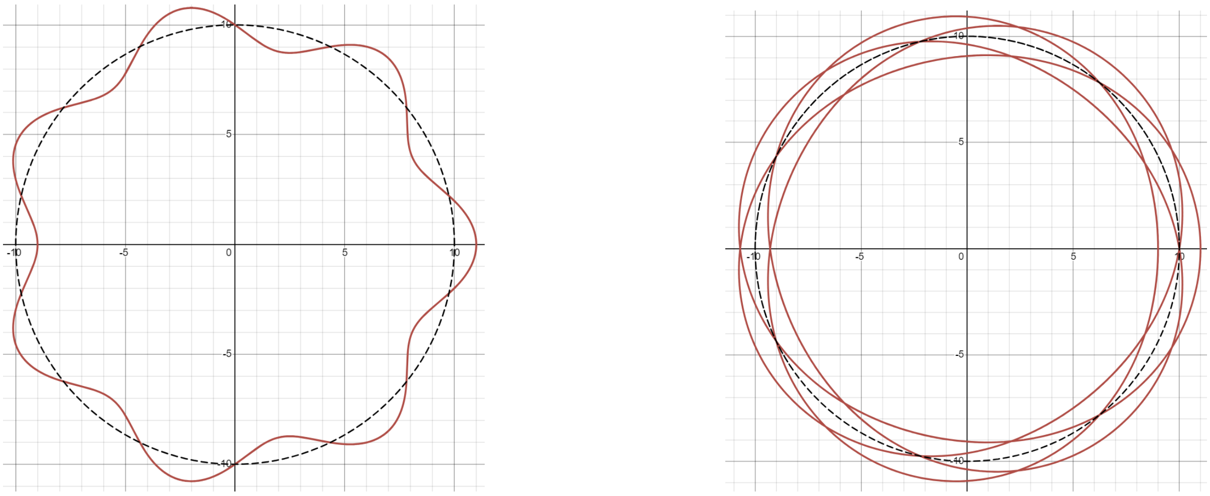


(a) (b)

Figure 3.3

Example: Power Laws (2).

Question: A particle of mass m, angular momentum L and energy $E > 0$ moves in a central potential $V(r) = \beta/r^2$. In the case $L^2/2m - \beta > 0$ solve for the trajectory of the particle in polar coordinates. What happens when $L^2/2m = \beta$?



(a) (b)

Figure 3.4

Solution: From energy and angular momentum conservation we have

$$E = \frac{1}{2}m\dot{r}^2 + \frac{L^2}{2mr^2} + \frac{\beta}{r^2} \quad \text{and} \quad L = mr^2\dot{\theta}. \tag{3.38}$$

As always, use the angular momentum conservation equation to elimiante t dependecy, so we substitute

$$\dot{r} = \frac{L}{mr^2}\frac{dr}{d\theta} \tag{3.39}$$

to get

$$\left(\frac{dr}{d\theta}\right)^2 = \frac{2mE}{L^2}r^4 - a^2r^2, \quad a^2 = 1 + \frac{2m\beta}{L^2}. \tag{3.40}$$

Now we do the famous substitution $u = 1/r$ and get

$$\left(\frac{du}{d\theta}\right)^2 = \frac{2mE}{L^2} - a^2u^2 \implies u(\theta) = \frac{1}{r} = \frac{\sqrt{2mE}}{aL}\sin(a\theta). \tag{3.41}$$

When $L^2/2m = \beta$, the effective potential vansies, and we simply have

$$\dot{r} = \sqrt{\frac{2E}{m}}. \tag{3.42}$$

Integrating we find that the particle crashes onto the origin in finite time, however, since

$$r = \frac{1}{\theta}\sqrt{\frac{L^2}{2mE}}, \tag{3.43}$$

as one can easily show by the same means as the previous case, $\theta \to \infty$ as $r \to 0$, so the particle winds around the origin infinite number of times before crashing.

Wavepacket Construction and Group-Velocity Dispersion

A localized wavepacket can be built by superposing plane waves whose wavenumbers lie in a small band around k_0:

$$\Psi(x,t) = \int_{k_0-\Delta k}^{k_0+\Delta k} A(k)\, e^{\,i\left(kx-\omega(k)\,t\right)}\, \mathrm{d}k,$$

where we take the dispersion relation

$$\omega(k) = \frac{k^2}{m}.$$

Phase and group velocities

$$\begin{aligned} \text{Phase velocity:} \quad v_p &= \frac{\omega(k_0)}{k_0} = \frac{k_0^2/m}{k_0} = \frac{k_0}{m}, \\ \text{Group velocity:} \quad v_g &= \left.\frac{\mathrm{d}\omega}{\mathrm{d}k}\right|_{k_0} = \frac{2k_0}{m}. \end{aligned}$$

Thus the *carrier* oscillations $e^{i(k_0x-\omega_0t)}$ move at v_p, while the slowly–varying envelope (determined by the width Δk) travels at v_g.

Spreading of the packet Expand $\omega(k)$ about k_0:

$$\omega(k) \approx \omega(k_0) + (k-k_0)\,\omega'(k_0) + \tfrac{1}{2}\,(k-k_0)^2\,\omega''(k_0) + \cdots$$

with

$$\omega'(k_0) = v_g, \qquad \omega''(k_0) = \frac{2}{m}.$$

Inserting into the integral and performing the standard Gaussian approximation shows the envelope acquires a factor

$$\exp\!\left[-\,\tfrac{(x-v_gt)^2}{2(\Delta x)^2+i\,(\omega''(k_0)\,t)}\right],$$

where $\Delta x \sim 1/\Delta k$. Because $\omega''(k_0) \neq 0$, the width $\sigma_x(t)$ grows in time, i.e. the wavepacket *spreads out*:

$$\sigma_x(t) \;=\; \sigma_x(0)\,\sqrt{1+\left(\tfrac{\omega''(k_0)\,t}{2\,\sigma_x(0)^2}\right)^2}.$$

This phenomenon is called *group-velocity dispersion.*

4 Lagrangian mechanics

4.1 Euler-Lagrange Equation

Generalized coordianates are a set of parameters $q_1, q_2, \ldots, q_n$ chosen to uniquely specify the configuration of a system such that knowing all q_i fixes every particle's position. The minimum number of independent generalized coordintates needed is the degrees of freedom of the system.

Suppose that the function $q(t)$ minimizes the action S,[1] which is the integral of the Lagrangian $\mathcal{L}(q(t), \dot{q}(t), t)$

$$S = \int_{t_1}^{t_2} \mathcal{L}\left(q(t), \dot{q}(t), t\right) dt \tag{4.1}$$

between two fixed time t_1 and t_2, where the particle transit from the generalized coordinate q_1 to q_2, *i.e.*, $q(t_1) = q_1$ and $q(t_2) = q_2$.

To find $q(t)$, we introduce variation $\delta q(t)$ (and $\dot{q}(t)$) as

$$\delta q(t) = q(\epsilon, t) - q(0, t) = \epsilon\eta(t) \implies \delta\dot{q}(t) = \epsilon\dot{\eta}(t), \tag{4.2}$$

where ϵ is a small paramter and $\eta(t)$ is an arbitrary smooth function that vanishes at the endpoints, *i.e.*,

$$\eta(t_1) = \eta(t_2) = 0. \tag{4.3}$$

The corresponding variation in the Lagrangian $\mathcal{L}$ is

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial q}\delta q + \frac{\partial\mathcal{L}}{\partial\dot{q}}\delta\dot{q} = \epsilon\eta(t)\frac{\partial\mathcal{L}}{\partial q} + \epsilon\dot{\eta}\frac{\partial\mathcal{L}}{\partial t} \tag{4.4}$$

The corresponding variation in the action S is

[1] S is a functional which takes a function as its input and produces a real number as its output.

$$\delta S = \epsilon \int_{t_1}^{t_2} \left(\frac{\partial \mathcal{L}}{\partial q} q(t) + \frac{\partial \mathcal{L}}{\partial \dot{q}} \dot{q}(t) \right) dt = \epsilon \int_{t_1}^{t_2} \left(\frac{\partial \mathcal{L}}{\partial q} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}} \right) \right) \eta(t) dt, \tag{4.5}$$

where in the last equality we used integration by part and the boundary term vanishes due to the eq. (4.3).

Demanding $\delta S = 0$ for minimum S gives the Euler-Lagrange equation

$$\frac{\partial q}{\partial t} = \frac{d}{dt} \frac{\partial \dot{q}}{\partial t}, \tag{4.6}$$

since $\eta(t)$ is an aribitrary function, so the integrand must be zero.

If the Lagragian $\mathcal{L}$ contains more than one generalized coordinate q, we simply apply the Euler-Lagrangian equation to each generalized coordinate q_i separately.

In fact this method is so general that t need not be time but any independent variable and q need not be generalized coordinate but any dependent variable.

4.1.1 Cyclic Coordinates

A cyclic coordinate is one which does not appear explicitly in the Lagragian

$$\frac{\partial L}{\partial q_i} = 0. \tag{4.7}$$

From the Euler-Lagrange equation it follows immediately that

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) = 0 \implies p_i = \frac{\partial L}{\partial \dot{q}_i} = \text{constant}, \tag{4.8}$$

where p_i is called the conjugate momentum.

The Beltrami identity

$$\mathcal{L} - \dot{q} \frac{\partial \mathcal{L}}{\partial \dot{q}} = \text{constant} \tag{4.9}$$

holds for cyclic coordinates.

To prove the identity, we note that

$$\frac{d\mathcal{L}}{dt} = \frac{\partial \mathcal{L}}{\partial t} + \frac{\partial \mathcal{L}}{\partial q} \frac{dq}{dt} + \frac{\partial \mathcal{L}}{\partial \dot{q}} \frac{d\dot{q}}{dt} \quad \text{and} \quad \frac{d}{dt} \left(\dot{q} \frac{\partial \mathcal{L}}{\partial \dot{q}} \right) = \frac{\partial \mathcal{L}}{\partial \dot{q}} \frac{d\dot{q}}{dt} + \dot{q} \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}} \right), \tag{4.10}$$

where $\frac{\partial \mathcal{L}}{\partial t} \neq \frac{d\mathcal{L}}{dt}$ since $\mathcal{L} = \mathcal{L}(q(t), \dot{q}(t), t)$. Combining the two equation gives

$$\frac{d}{dt} \left(\dot{q} \frac{\partial \mathcal{L}}{\partial \dot{q}} \right) = \frac{d\mathcal{L}}{dt} - \frac{\partial \mathcal{L}}{\partial t} + \dot{q} \left(\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}} - \frac{\partial \mathcal{L}}{\partial q} \right) \implies \frac{\partial \mathcal{L}}{\partial t} = \frac{d}{dt} \left(\mathcal{L} - \dot{q} \frac{\partial \mathcal{L}}{\partial \dot{q}} \right). \tag{4.11}$$

Example: Shortest Distance Between Two Points.

Question: Prove that the shortest distance between two points is a straight line.

Solution: We want to minimize the arc length

$$L = \int_{x_1}^{x_2} \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx, \tag{4.12}$$

so we can use the Euler-Lagrange equation, where $\mathcal{L} = \sqrt{1 + y'^2}$, which gives

$$\frac{\partial \mathcal{L}}{\partial y} = 0 = \frac{\partial \mathcal{L}}{\partial y'} \implies y = ax + b. \tag{4.13}$$

Example: Brachistochrone Problem.

Question: Find the shape for the minimum transite time between two points $(0, 0)$ and (x, y) under gravity.

Solution: We want to minimize the transit time

$$t = \int_{x_1}^{x_2} \frac{ds}{v} = \int_{y_1}^{y_2} \frac{\sqrt{1 + x'^2}}{2gy} dy, \tag{4.14}$$

where we have changed the independent variable from x to y. Applying the Euler-Lagrange equation gives

$$\frac{\partial t}{\partial x} = 0 = \frac{d}{dy} \frac{\partial t}{\partial x'}. \tag{4.15}$$

Solving gives

$$x = a(\theta - \sin\theta) \quad \text{and} \quad y = a(1 - \cos\theta), \tag{4.16}$$

where θ is a parameter, and the solution is a cycloid.

Example: Minimal Travel Cost.

Question: The cost of flying an aircraft at height z is $e^{-\kappa z}$ per unit distance of flight-path. Find the shape of the flight path for minimal flying cost from $(-a, 0)$ $(a, 0)$.

Solution: We want to mimize the flying cost

$$C = \int_{-a}^{a} e^{-\kappa z} ds = \int_{-a}^{a} e^{-\kappa z} \sqrt{1 + z'^2} dx. \tag{4.17}$$

Applying the Euler-Lagrange equation gives

$$\frac{\partial \mathcal{L}}{\partial z} = -\kappa e^{-\kappa z}\sqrt{1+z'^2} = \frac{d}{dx}\frac{\partial \mathcal{L}}{\partial z'} = \frac{z'' e^{-\kappa z}}{\sqrt{1+z'^2}} - \frac{\kappa z'^2 e^{-\kappa z}}{\sqrt{1+z'^2}} - \frac{z'' z'^2 e^{-\kappa z}}{(1+z'^2)^{3/2}}. \quad (4.18)$$

Solving gives

$$z(x) = \frac{1}{\kappa}\ln\frac{\cos(\kappa x)}{\cos(\kappa a)}. \quad (4.19)$$

Example: Surface Area of a Cylindrically-Symmetric Soap Bubble.

Question: Consider the surface stretched by a soap bubble film with boundaries being two circular hoops. Find the shape of the soap bubble at equilibrium.

Solution: We want to minimize the surface energy, which is proportional to the surface area

$$S = 2\pi \int \rho\sqrt{dz^2 + d\rho^2} = 2\pi \int_{z_1}^{z_2} \rho\sqrt{1+\left(\frac{d\rho}{dz}\right)^2} dz = 2\pi \int_{\rho_1}^{\rho_2} \rho\sqrt{1+\left(\frac{dz}{d\rho}\right)^2} dz. \quad (4.20)$$

If we choose z to be the independent variable, then the Euler-Lagrange equation reads

$$\frac{\partial \mathcal{L}}{\partial \rho} = \sqrt{1+\rho'^2} = \frac{d}{dz}\frac{\partial \mathcal{L}}{\partial \rho'} = \frac{d}{dz}\left(\frac{\rho\rho'}{\sqrt{1+\rho'^2}}\right), \quad (4.21)$$

which is not an easy equation to solve.
On the other hand, if we choose ρ to be the independent variable, then the Euler-Lagrange equation reads

$$\frac{\partial \mathcal{L}}{\partial z} = 0 = \frac{d}{d\rho}\frac{\partial \mathcal{L}}{\partial z'} = \frac{d}{d\rho}. \quad (4.22)$$

Solving gives

$$\rho = a\cosh\left(\frac{z-b}{a}\right), \quad (4.23)$$

where a and b are constants determined by the boundary conditions.

Example: Fermat's Principle

Question: Refer to fig. 4.1 and derive the Snell's law $n_1 \sin\theta_1 = n_2 \sin\theta_2$ when light propagate from one medium of refractive index n_1 with incidence angle θ_1 to another medium of refractive index n_2 with refracted angle θ_2.

Solution: We want to minimize the transit time

$$t = \int_{y_1}^{y_2} \frac{ds}{v} = \int_{y_1}^{-y_2} \frac{1}{c} n(x,y,z)\sqrt{1+x'^2+z'^2}dy \tag{4.24}$$

The Euler-Lagrange equation for z reads

$$0 + \frac{d}{dy}\left(\frac{1}{c}\left(\frac{n_1 z'}{\sqrt{1+x'^2+z'^2}} + \frac{n_2 z'}{\sqrt{1+x'^2+z'^2}}\right)\right) = 0 \implies z' = 0, \tag{4.25}$$

which is equivalent to saying that the z-coordinate of the light beam remains at $z = 0$ at all times.
The Euler-Lagrange equation for x reads

$$0 + \frac{d}{dy}\left(\frac{1}{c}\left(\frac{n_1 \tan\theta_1}{\sqrt{1+\tan^2\theta_1}} - \frac{n_2 \tan\theta_2}{\sqrt{1+\tan^2\theta_2}}\right)\right) = \frac{d}{dy}\left(\frac{1}{c}\left(n_1 \sin\theta_1 - n_2 \sin\theta_2\right)\right) = 0. \tag{4.26}$$

Therefore $n_1 \sin\theta_1 - n_2 \sin\theta_2$ is a constant, which must be zero since when $n_1 = n_2$ we have $\theta_1 = \theta_2$.
In fact the geometry of this problem is simple enough that directly minimizing the path rather than using the Euler-Lagrage equation is faster. However due to the simplicity of this approach it is not illustrated here.

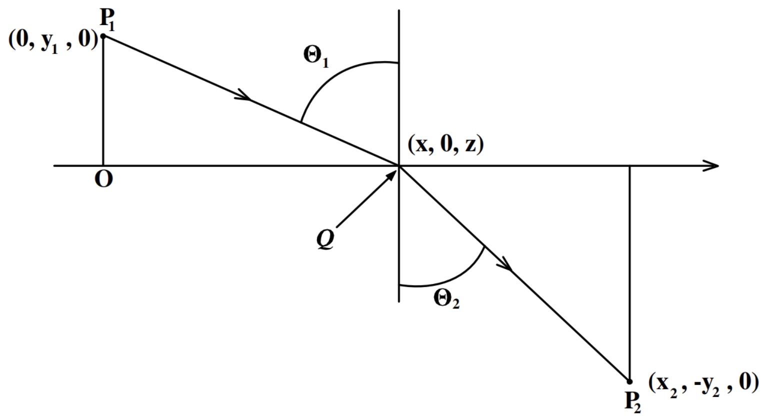


Figure 4.1

Example: Minimum of $(\nabla\phi)^2$ in a Volume.

Question: Find the function $\phi(\mathbf{r})$ that has the minimum value of $(\nabla\phi)^2$ per unit volume.

Solution: We want to minimize

$$J = \frac{1}{V}\int (\nabla\phi)^2 dV = \frac{1}{V}\int \left(\left(\frac{\partial\phi}{\partial x}\right)^2 + \left(\frac{\partial\phi}{\partial y}\right)^2 + \left(\frac{\partial\phi}{\partial z}\right)^2\right). \tag{4.27}$$

The Euler-Lagrange equation for the three coordinates read

$$\frac{\partial \mathcal{L}}{\partial \phi} = 0 = \frac{\partial}{\partial x_i}\left(\frac{\partial \mathcal{L}}{\partial \phi'}\right) \implies \frac{\partial^2 \phi}{\partial x_i^2} = 0 \implies \sum_{i=1}^{3} \frac{\partial^2 \phi}{\partial x_i^2} = 0 \implies \nabla^2 \phi = 0, \quad (4.28)$$

therefore ϕ must satisfy the Laplace's equation in order that $(\boldsymbol{\nabla}\phi)^2$ is minimum.

4.2 Constraints

In the above section we assumed the coordinates q_i are independent. However, there could be constriants in the system which relates the coordinates q_i. The constriant in the form

$$f(q_1, q_2, ..., t) = 0, \quad (4.29)$$

For example, the constriant on a rigid body is that the distance betewen any two points in the body is fixed, *i.e.,* $|\mathbf{r}_i - \mathbf{r}_j| = c_{ij}$.

A system consisting of N free particles has $3N$ degrees of freedom and thus $3N$ independent variables. If there are k holonomic constraints, then we can always find $3N-k$ independent variables (known as the generalized coordinates), which is the minimum number of variables that can still fully describe the state of the system. We can then apply Euler-Lagrange equations to the $3N-k$ generalized coordinates separately to obtain the system's equation of motion.

For example, for a double pendulum, we can use the two equations which state that the lengths of the two rods are constant to eliminate two of the four Cartesian variables. Alternatively, we can simply use the two generalized coordinates θ_1 and θ_2.

In contrary, a constraint that cannot be expressed in the above form is called a nonholonomic constraint. For example, refering to fig. 4.2, consider a disc with radius a rolling on the horizontal x-y plane such that the plane of the disc is always vertical. In addition to the x and y coordinates of the center of mass, we need specify its orientation using the angle between the disc's symmetric axis and the x-axis θ and its rotated angle ϕ.

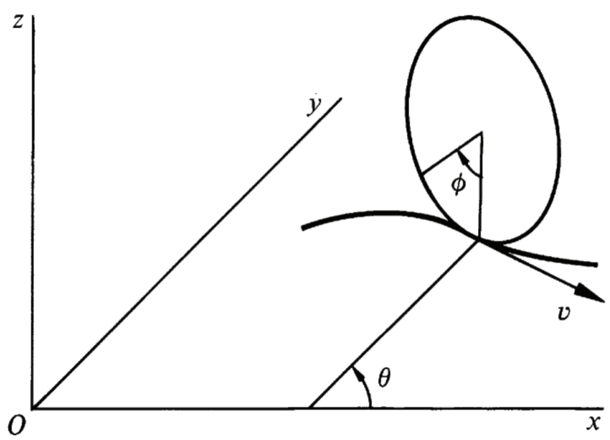


Figure 4.2

The angular velocity of the disc is

$$\boldsymbol{\omega} = \dot{\theta}\hat{\mathbf{z}} + \dot{\phi}(\cos\theta\hat{\mathbf{x}} + \sin\theta\hat{\mathbf{y}}), \tag{4.30}$$

since the rotations in θ and ϕ occur about orthogonal axes (one horizontal and one vertical), so the angular velocity is simply the vector sum of the two contributions. One can show formally using the rotational matrix that $\boldsymbol{\omega} = \dot{R}R^T$ is indeed what we have claimed.

The constraint of the disc's motion is that the relative velocity of the contact point between the disc and the plane must be zero, *i.e.,*

$$\begin{aligned} \mathbf{v}_{\text{c.m.}} + \boldsymbol{\omega} \times \mathbf{r}_{\text{contact}} &= 0 \\ (\dot{x}\hat{\mathbf{x}} + \dot{y}\hat{\mathbf{y}}) + \left(\dot{\theta}\hat{\mathbf{z}} + \dot{\phi}(\cos\theta\hat{\mathbf{x}} + \sin\theta\hat{\mathbf{y}})\right) \times (-a\hat{\mathbf{z}}) &= 0 \\ (\dot{x}\hat{\mathbf{x}} + \dot{y}\hat{\mathbf{y}}) + a\dot{\phi}\cos\theta\hat{\mathbf{y}} - a\dot{\phi}\sin\theta\hat{\mathbf{x}} &= 0 \\ (\mathbf{x} - a\dot{\phi}\sin\theta)\hat{\mathbf{x}} + (\mathbf{y} + a\dot{\phi}\cos\theta)\hat{\mathbf{y}} &= 0 \\ \mathbf{x} = a\dot{\phi}\sin\theta \quad \text{and} \quad \mathbf{y} &= -a\dot{\phi}\cos\theta. \end{aligned} \tag{4.31}$$

However, these equations of constriant are not relating the dependent variables x, y, θ, ϕ but rather $x, y, \dot{\theta}, \dot{\phi}$ and neither of these equations can be integrated without first solving the problem itself, so they are nonholonomic since we cannot eliminate the dependent variables using these equations.

Another example of a nonholonomic constriant is a particle being put on the surface of a sphere of radius a. In this case the constriant appears in the form of an inequality $x^2 + y^2 + z^2 > a$.

The constriant can even come in integral form. For example, in the catenary problem the arc length $\int_{x_1}^{x_2} \sqrt{1 + y'^2}dx = l$ is constrained to be equal to a fixed length l.

4.3 Lagrange Multipliers for Holonomic Constraints.

Beside using generalized coordinates to solve for the equation of motions with prescence of constraint forces as detailed in section 4.2, we can use the method of Lagrange multipliers.

Suppose there are m holonomic algebraic constraints for the n variables q_i $(1 \leq i \leq n)$, *i.e.,* $g_k(\mathbf{q}) = 0$ $(1 \leq k \leq m)$.

Example: Rolling Cylinder.

Question: As shown in fig. 4.3, a uniform solid cylinder of mass M_c and radius R rolls without slipping down the inclided surface of a wedge-shaped block with mass M_b and angle of inclination α.

1. Justify that the system can be described in terms of two generalized coordinates and determine the linear accelerations $\ddot{X}$ and $\ddot{s}$, where X and s are the position of the block and distance along the inclined surface to the cylinder, respectively.

2. Consider the case where the block is fixed in position determine the constraint forces by treating the system as having three generalised coordintaes (r, s, θ). Hence find the rolling-without-slipping condition to hold.

Solution:

1. The two coordinates θ and s are related by $R\theta = s + \text{constant}$, through a constriant force f, so we can just consider either θ or s and omit the constriant force f.
   The Lagragian is

$$\mathcal{L} = \frac{1}{2}M_b\dot{X}^2 + \frac{1}{2}M_c\left(\left(\dot{X} + \dot{s}\cos\alpha\right)^2 + (\dot{s}\sin\alpha)^2\right) + \frac{1}{2}\left(\frac{1}{2}M_cR^2\right)\dot{\theta}^2 - M_cgs\cos\alpha. \tag{4.32}$$

   Applying the Euler-Lagrange equations then give

$$\ddot{X} = \frac{2M_cg\cos^2\alpha}{3(M_b + M_c) - 2M_c\cos^2\alpha} \quad \text{and} \quad \ddot{s} = \frac{-2(M_b + M_c)M_cg\cos^2\alpha}{M_c\cos\alpha(3(M_b + M_c) - 2M_c\cos^2\alpha)}. \tag{4.33}$$

2. The Lagragian is now

$$\mathcal{L} = \frac{1}{2}M_c^2\dot{s}^2 + \frac{1}{2}M_c\dot{r}^2 + \frac{1}{2}\left(\frac{1}{2}M_cR^2\right)\dot{\theta}^2 - M_cg(s\cos\alpha + r\sin\alpha) + Nr + fs + fR\theta. \tag{4.34}$$

   The Euler-Lagrange equations now read

$$\begin{aligned}
\frac{\partial\mathcal{L}}{\partial r} &= N - M_cg\sin\alpha = \frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{r}} = M_c\ddot{r} = 0, \\
\frac{\partial\mathcal{L}}{\partial s} &= f - M_cg\cos\alpha = \frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{s}} = M_c\ddot{s}, \\
\frac{\partial\mathcal{L}}{\partial\theta} &= fR = \frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{\theta}} = \frac{1}{2}M_cR^2\ddot{\theta}.
\end{aligned} \tag{4.35}$$

   From the first equation, and using the $\ddot{s} = -R\ddot{\theta}$,[a] the last two equations combined to give

$$N = M_cg\sin\alpha \quad \text{and} \quad f = \frac{1}{3}M_cg\cos\alpha \implies \mu \geq \frac{|f|}{|N|} = \frac{1}{3}\tan\alpha. \tag{4.36}$$

   Here note that we note that θ and s are related by the constriant force f. Either we include both θ and s in the Lagrangian to find f, or we simply treat either θ or s as the generalized coordinate and omit f completely.

[a] This is because we have assumed clockwise as the positive direction for θ. If we have defined anti-clockwise as positive, then we would have $\ddot{s} = R\ddot{\theta}$, but then the $fR\theta$ term in the Lagrangian should have an extra negative sign.

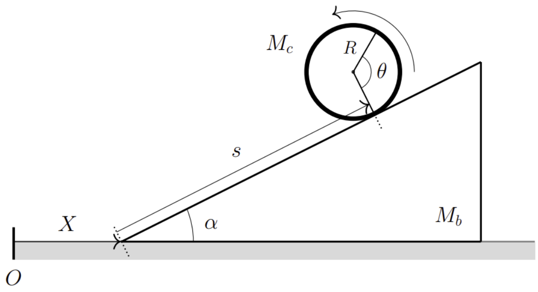


Figure 4.3

4.4 Hamiltonian Mechanics

4.4.1 Hamiltonian

The Hamiltonian H of a system is defined as

$$H = \sum_i \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} - L. \tag{4.37}$$

A system is isolated if its Lagrangian has no explicit time dependence

$$\frac{\partial L}{\partial t} = 0. \tag{4.38}$$

In those cases the Hamiltonian H is conserved

$$\begin{aligned}
\frac{dE}{dt} &= \sum_i \ddot{q}_i \frac{\partial L}{\partial \dot{q}_i} + \sum_i \dot{q}_i \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{dL}{dt} \\
&= \sum_i \ddot{q}_i \frac{\partial L}{\partial \dot{q}_i} + \sum_i \dot{q}_i \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \left(\sum_i \frac{\partial L}{\partial q_i}\dot{q}_i + \sum_i \frac{\partial L}{\partial \dot{q}_i}\ddot{q}_i + \frac{\partial L}{\partial t}\right) \\
&= \sum_i \dot{q}_i \left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial t} = 0.
\end{aligned} \tag{4.39}$$

Appendices

$\mathcal{A}$

Rigid Body Mechanics

A.1 Chasles' Theorem

We begin by considering two masses m_1 and m_2 located at $\mathbf{r}_1$ and $\mathbf{r}_2$ respectively connected by a thin, rigid and massless rod.

The "rigid body condition" is that the distance between the two masses remained unchanged, *i.e.,*

$$
\begin{aligned}
d(|\mathbf{r}_1 - \mathbf{r}_2|) &= 0 \\
|\mathbf{r}_1 - \mathbf{r}_2| &= c \\
|\mathbf{r}_1 - \mathbf{r}_2|^2 = (\mathbf{r}_1 - \mathbf{r}_2) \cdot (\mathbf{r}_1 - \mathbf{r}_2) &= c^2 \\
d((\mathbf{r}_1 - \mathbf{r}_2) \cdot (\mathbf{r}_1 - \mathbf{r}_2)) &= 2(\mathbf{r}_1 - \mathbf{r}_2) \cdot d(\mathbf{r}_1 - \mathbf{r}_2) = 0 \\
d\mathbf{r}_1 &= d\mathbf{r}_2 \text{ or } (d\mathbf{r}_1 - d\mathbf{r}_2) \perp (\mathbf{r}_1 - \mathbf{r}_2)
\end{aligned}
\tag{A.1}
$$

Now since $d\mathbf{r}'_1 = d\mathbf{r}_1 - d\mathbf{R} = (\frac{m_2}{m_1+m_2})(d\mathbf{r}_1 - d\mathbf{r}_2)$ and $d\mathbf{r}'_2 = d\mathbf{r}_2 - d\mathbf{R} = -(\frac{m_1}{m_1+m_2})(d\mathbf{r}_1 - d\mathbf{r}_2)$, so when $d\mathbf{r}_1 = d\mathbf{r}_2$ in the first case, it means that the body undergo pure translation without rotating. And the second case corresponds to a case of translation plus rotation since

1. $d\mathbf{r}'_1 \perp (\mathbf{r}_1 - \mathbf{r}_2)$:

$$
d\mathbf{r}'_1 \cdot (\mathbf{r}_1 - \mathbf{r}_2) = (\frac{m_2}{m_1 + m_2})(d\mathbf{r}_1 - d\mathbf{r}_2) \cdot (\mathbf{r}_1 - \mathbf{r}_2) = 0 \tag{A.2}
$$

2. $d\mathbf{r}'_2 \perp (\mathbf{r}_1 - r_2)$: the proof is the same as above

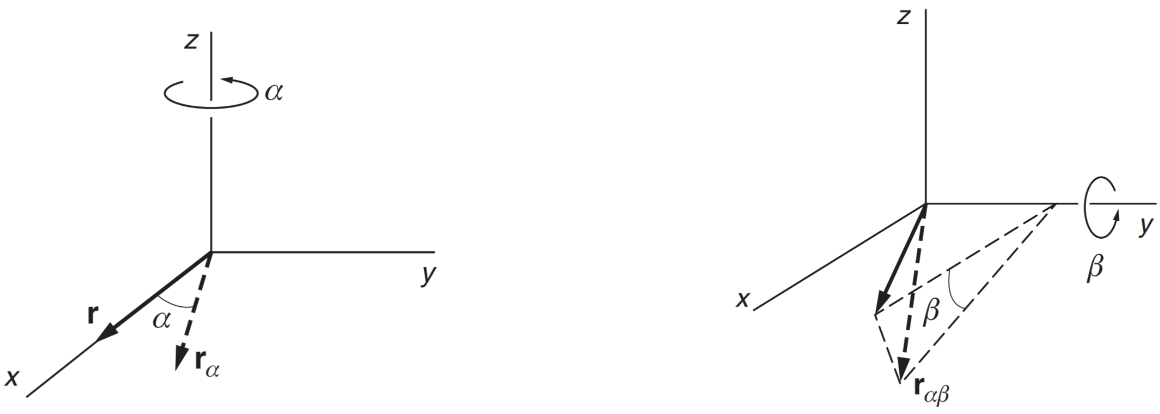
3. $\frac{d\mathbf{r}'_1}{r'_1} = -\frac{d\mathbf{r}'_2}{r'_2}$:

$$
\frac{d\mathbf{r}'_1}{r'_1} = (\frac{m_2}{m_1 + m_2})\frac{(d\mathbf{r}_1 - d\mathbf{r}_2)}{r'_1} = (\frac{m_1}{m_1 + m_2})\frac{(d\mathbf{r}_1 - \mathbf{r}_2)}{r'_2} = -\frac{d\mathbf{r}'_2}{r'_2}. \tag{A.3}
$$

A.2 Noncommutability of finite rotations

fig. A.1 illustrates the essence of the general proof of this fact, where we consider the rotation of the position vector $\mathbf{r} = r\hat{\mathbf{i}}$ through an angle α about the z axis and β about the

y axis but in different order. Rotating about z axis by an angle α, $\hat{\mathbf{i}}$ becomes $\cos\alpha\hat{\mathbf{i}}+\sin\alpha\hat{\mathbf{j}}$ while rotating about y axis by an angle β, $\hat{\mathbf{i}}$ becomes $\cos\beta\hat{\mathbf{i}}-\sin\beta\hat{\mathbf{k}}$, so



(a) (b)

Figure A.1

$$\mathbf{r}_{\alpha\beta} = r\cos\alpha(\cos\beta\hat{\mathbf{i}} - \sin\beta\hat{\mathbf{k}}) + r\sin\alpha\hat{\mathbf{j}} = r\cos\alpha\cos\beta\hat{\mathbf{i}} + r\sin\alpha\hat{\mathbf{j}} - r\cos\alpha\sin\beta\hat{\mathbf{k}}$$
$$\text{and } \mathbf{r}_{\beta\alpha} = r\cos\alpha\cos\beta\hat{\mathbf{i}} + r\cos\beta\sin\alpha\hat{\mathbf{j}} - r\sin\beta\hat{\mathbf{k}}. \tag{A.4}$$

It is evident that while finite size of α and β would result in a difference between $\mathbf{r}_{\alpha\beta}$ and $\mathbf{r}_{\beta\alpha}$, but if we take the limit $\alpha \ll 1$ and $\beta \ll 1$, then $\mathbf{r}_{\alpha\beta} = \mathbf{r}_{\beta\alpha}$ and the angular displacement vector $\Delta\boldsymbol{\theta} = \Delta\alpha\hat{\mathbf{k}} + \Delta\beta\hat{\mathbf{j}}$ is well defined. In particular, the displacement of $\mathbf{r}$ is

$$\Delta\mathbf{r} = \mathbf{r}_{\alpha\beta} - \mathbf{r} = \mathbf{r}_{\beta\alpha} - \mathbf{r} = r\alpha\hat{\mathbf{j}} - r\beta\hat{\mathbf{k}} = \Delta\boldsymbol{\theta} \times \mathbf{r}. \tag{A.5}$$

The linear velocity will then be

$$\mathbf{v} = \lim_{\Delta t\to 0}\frac{\Delta\mathbf{r}}{\Delta t} = \lim_{\Delta t\to 0}\frac{\Delta\boldsymbol{\theta}\times\mathbf{r}}{\Delta t} = \boldsymbol{\omega}\times\mathbf{r}. \tag{A.6}$$

where the angular velocity vector $\boldsymbol{\omega}$ is defined as

$$\boldsymbol{\omega} = \lim_{\Delta t\to 0}\frac{\Delta\theta}{\Delta t} \tag{A.7}$$

In this case, $\boldsymbol{\omega} = \frac{d\beta}{dt}\hat{\mathbf{j}} + \frac{d\alpha}{dt}\hat{\mathbf{k}}$.