

Classical Mechanics

Haydn Cheng

March 21, 2025

Contents

1	Newtonian mechanics	3
1.1	Point Particle	3
1.2	Work Energy Theorem	9
1.3	System of particles	9
1.4	Rigid Body Mechanics	12
1.4.1	Prerequisites	12
1.4.2	Angular velocity vector	13
1.5	Tensor of Inertia	15
1.5.1	Angular Momentum and Energy	15
1.5.2	Parallel Axis Theorem	16
1.5.3	Euler's Equations	17
1.6	Torque-free Precession	18
1.6.1	Euler Angles	22
2	Waves	25
2.1	Normal Modes	25
2.1.1	Basics	25
2.1.2	Unequal Masses	26
2.1.3	Damped and Forced Oscillators	26
2.1.4	Energy Conservation	27
2.1.5	N Masses	27
2.2	Examples of Wave Equations	29
2.2.1	Longitudinal Oscillations of a String	29
2.2.2	Transverse Oscillations of a String	29
2.2.3	Acoustic Waves	30
2.3	Solution to the Wave Equation	31
2.3.1	d'Alembert's Solution	31
2.3.2	Separation of Variables	36
2.4	Energy in Transverse Oscillation of a String	41
2.5	Dispersive Waves	44
2.5.1	Phase Velocity	44
2.5.2	Group Velocity	45

2.5.3	Gravity Waves with Surface Tension	45
2.5.4	Wavepackets	46
3	Lagrangian mechanics	47
3.1	Euler-Lagrange Equation	47
3.2	Constraints	51
3.3	Lagrange Multipliers for Holonomic Constraints.	53
	Appendices	54
A	Rigid Body Mechanics	55
A.1	Chasles' Theorem	55
A.2	Noncommutability of finite rotations	55

Newtonian mechanics

1.1 Point Particle

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

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

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

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

where \mathbf{a} is the acceleration of m and \mathbf{v}_{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}_{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.

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

We see that the net force \mathbf{F}_{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,

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

which is the same as the equation for adding mass, just that \mathbf{v}_{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 infinitesimal 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

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

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

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

This is a non-linear 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 stretched 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

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

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 experiences 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

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

Adding N with $((l-x)/l)mg$, which is the contribution of the stationary chain to the reading of the scale gives $N_{\text{tot}} = 3((l-x)/l)mg$.

Secondly, we can consider the modified Newton's second law of the falling chain

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

Adding N with $((l-x)/l)mg$, which is the contribution of the stationary chain to the reading of the scale gives $N_{\text{tot}} = 3((l-x)/l)mg$.

Thirdly, we can consider the part of the chain 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

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

The center of mass velocity is therefore

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

The center of mass acceleration is therefore

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

The unmodified Newton's second law therefore gives

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

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

$$F = ma + \frac{dm}{dt}v = \frac{dm}{dt}v. \quad (1.14)$$

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

$$F = dma = dm \frac{v}{t} = \frac{dm}{dt}v. \quad (1.15)$$

As we have found, the force F is not contributing to the usual acceleration \mathbf{a} of the mass m but is entirely entitled 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 New-

ton's second law

$$F = ma + \frac{dm}{dt}v_{\text{rel}} = ma. \quad (1.16)$$

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 accelerate the mass m .

Example: Pulling Carpet.

Question: A long, thin, pliable carpet of mass m and length 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

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

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

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

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

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

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

$$\boldsymbol{\tau} = \mathbf{r} \times \mathbf{F} \quad \text{and} \quad \mathbf{L} = \mathbf{r} \times \mathbf{p}. \quad (1.20)$$

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”

$$\frac{d\mathbf{L}}{dt} = \frac{d}{dt}(\mathbf{r} \times \mathbf{p}) = \mathbf{r} \times \frac{d\mathbf{p}}{dt} + \frac{d\mathbf{r}}{dt} \times \mathbf{p} = \mathbf{r} \times \mathbf{F} + \mathbf{v} \times (m\mathbf{v}) = \mathbf{r} \times \mathbf{F} = \boldsymbol{\tau} \quad (1.21)$$

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

$$T = \frac{1}{2}mv^2 \quad \text{and} \quad \mathbf{F} = -\nabla V \quad (1.22)$$

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

$$W = \int_1^2 \mathbf{F}_{net} \cdot d\mathbf{r} = \int_1^2 \frac{dm\mathbf{v}}{dt} \cdot d\mathbf{r} = m \int_1^2 d\dot{\mathbf{r}} \cdot \frac{d\mathbf{r}}{dt} = \frac{m}{2} \int_1^2 d(v^2) = \frac{1}{2}mv_2^2 - \frac{1}{2}mv_1^2 \quad (1.23)$$

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)

$$W = W_{con} + W_{non-con} = \Delta T.^2 \quad (1.24)$$

¹One 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, $\boldsymbol{\tau} = \frac{d\mathbf{L}}{dt}$ 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.

²Some 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, dW is used to denote the inexact differential, but the complexity of the symbol forbids me to consistently type it in latex.

If we define $\mathbf{F}_{net} = \mathbf{F}_{con} + \mathbf{F}_{non-con}$ which is comprised of both conservative forces such as gravity where $\mathbf{F} \propto \frac{\mathbf{r}}{r^2}$ 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 Stoke's theorem,
2. $\nabla \times \mathbf{F} = 0$, 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 $\nabla \times (\nabla V) = 0$ and the second item above ($\nabla \times \mathbf{F} = 0$), we can define the potential energy as mentioned and the work-energy theorem (eq. (1.24)) becomes the conservation of energy

$$W_{non-con} = T + V. \quad (1.25)$$

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 $\Delta K.E.$

$$W = \Delta K.E.. \quad (1.26)$$

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

$$W = W_{con} + W_{non-con}, \quad (1.27)$$

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

$$\Delta P.E. = -W_{con}, \quad (1.28)$$

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

$$W_{non-con} = \Delta K.E. + \Delta P.E.. \quad (1.29)$$

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 i th particle is

$$\mathbf{F}_i = \sum_j \mathbf{F}_{j \rightarrow i} + \mathbf{F}_{i,ext} = \frac{d^2(m_i \mathbf{r}_i)}{dt^2}. \quad (1.30)$$

Summing over all particles,

$$\sum_i \sum_j \mathbf{F}_{j \rightarrow i} + \sum_i \mathbf{F}_{i,ext} = \mathbf{F}_{ext,net} = \sum_i \frac{d^2(m_i \mathbf{r}_i)}{dt^2} = \frac{d^2}{dt^2} \sum_i m_i \mathbf{r}_i = \left(\sum_i m_i \right) \ddot{\mathbf{R}}, \quad (1.31)$$

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

$$\mathbf{R} = \frac{\sum_i m_i \mathbf{r}_i}{\sum_i m_i} \quad (1.32)$$

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 i th particle, we have

$$\mathbf{r}_i \times \mathbf{F}_i = \mathbf{r}_i \times \mathbf{F}_{i,ext} + \mathbf{r}_i \times \sum_j \mathbf{F}_{j \rightarrow i} = \frac{dL_i}{dt}. \quad (1.33)$$

Summing over all particles,

$$\sum_i (\mathbf{r}_i \times \mathbf{F}_{i,ext}) + \sum_i (\mathbf{r}_i \times \sum_j \mathbf{F}_{j \rightarrow i}) = \boldsymbol{\tau}_{ext} + \sum_i ((\mathbf{r}_i - \mathbf{r}_j) \times \mathbf{F}_{j \rightarrow i}) = \boldsymbol{\tau}_{ext} = \sum_i \frac{dL_i}{dt} = \dot{\mathbf{L}}_{tot}. \quad (1.34)$$

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

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

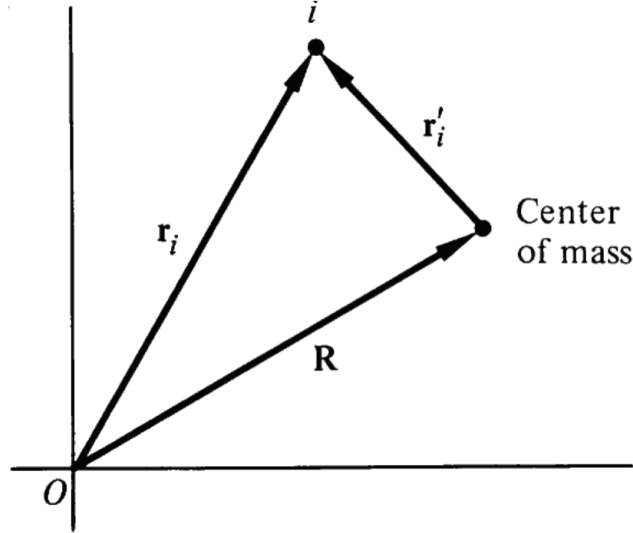


Figure 1.1

, thus $\mathbf{p}_i = m_i \dot{\mathbf{r}}'_i + m_i \dot{\mathbf{R}}$ and the total angular momentum becomes

$$\mathbf{L}_{tot} = \sum_i \mathbf{r}_i \times \mathbf{p}_i = \sum_i ((\mathbf{r}'_i + \mathbf{R}) \times (m_i \dot{\mathbf{r}}'_i + m_i \dot{\mathbf{R}})) = (\sum_i m_i) \mathbf{R} \times \dot{\mathbf{R}} + \sum_i m_i (\mathbf{r}'_i \times \dot{\mathbf{r}}'_i) \quad (1.35)$$

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

$$T = \sum_i \frac{1}{2} m_i v_i^2 = \frac{1}{2} \sum_i m_i (\dot{\mathbf{r}}'_i + \dot{\mathbf{R}})^2 = \frac{1}{2} \sum_i m_i \dot{\mathbf{r}}'^2_i + \frac{1}{2} (\sum_i m_i) \dot{\mathbf{R}}^2 \quad (1.36)$$

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

⁴Another 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.

⁵Yet another interesting and useful fact is that if the motion of the body is planar (*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

$$\Delta\boldsymbol{\theta} \stackrel{?}{=} \Delta\theta_x\hat{\mathbf{x}} + \Delta\theta_y\hat{\mathbf{y}} \neq \Delta\theta_y\hat{\mathbf{y}} + \Delta\theta_x\hat{\mathbf{x}}. \quad (1.37)$$

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

$$\delta\boldsymbol{\theta} = \delta\theta_x\hat{\mathbf{x}} + \delta\theta_y\hat{\mathbf{y}} = \delta\theta_y\hat{\mathbf{y}} + \delta\theta_x\hat{\mathbf{x}}. \quad (1.38)$$

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.

⁶From 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.

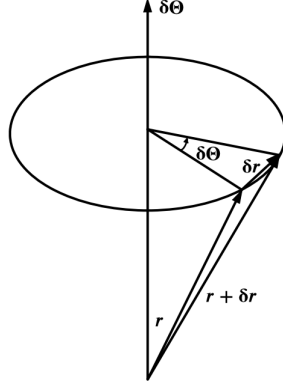


Figure 1.2

From it, it is clear that

$$\delta \mathbf{r} = \delta \boldsymbol{\theta} \times \mathbf{r}. \quad (1.39)$$

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

$$\delta \mathbf{r}_{12} = \delta \boldsymbol{\theta}_1 \times \mathbf{r} + \delta \boldsymbol{\theta}_2 \times (\mathbf{r} + \delta \mathbf{r}) = (\delta \boldsymbol{\theta}_1 + \delta \boldsymbol{\theta}_2) \times \mathbf{r} = \delta \mathbf{r}_{21} \quad (1.40)$$

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

Dividing eq. (1.39) by δt , we have

$$\mathbf{v} = \frac{\delta \mathbf{r}}{\delta t} = \frac{\delta \boldsymbol{\theta}}{\delta t} \times \mathbf{r}. \quad (1.41)$$

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

$$\mathbf{v}_P = \mathbf{v}_O + \boldsymbol{\omega} \times \mathbf{r}_{O \rightarrow P} \quad (1.42)$$

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, $\mathbf{v}_P = \mathbf{v}_{c.m.} + \boldsymbol{\omega} \times \mathbf{r}_{c.m. \rightarrow P}$ and $\mathbf{v}_O = \mathbf{v}_{c.m.} + \boldsymbol{\omega} \times \mathbf{r}_{c.m. \rightarrow O}$ yields the general result. In fact, eq. (1.42) 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

$$\boldsymbol{\Omega}_{1\text{rel.}2} = \boldsymbol{\Omega}_{1\text{rel.}3} - \boldsymbol{\Omega}_{2\text{rel.}3}. \quad (1.43)$$

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 $\Delta \rightarrow \delta$ and it is harder to visualize the entire process.

1.5 Tensor of Inertia

1.5.1 Angular Momentum and Energy

Now we return to eq. (1.35) 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

$$\frac{d\mathbf{r}}{dt} = \boldsymbol{\omega} \times \mathbf{r}. \quad (1.44)$$

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

$$\mathbf{L}_{rot} = \sum_i m_i (\mathbf{r}'_i \times \dot{\mathbf{r}}'_i) = \sum_i m_i (\mathbf{r}'_i \times (\boldsymbol{\omega} \times \mathbf{r}'_i)) = \sum_i m_i (|\mathbf{r}'_i|^2 \boldsymbol{\omega} - \mathbf{r}'_i (\mathbf{r}'_i \cdot \boldsymbol{\omega})). \quad (1.45)$$

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

$$\begin{aligned} L_{rot,x} &= \sum_i m_i ((x_i'^2 + y_i'^2 + z_i'^2) \omega_x - x_i' (x_i' \omega_x + y_i' \omega_y + z_i' \omega_z)) \\ &= \sum_i m_i ((y_i'^2 + z_i'^2) \omega_x - (x_i' y_i') \omega_y - (x_i' z_i') \omega_z), \\ L_{rot,y} &= \sum_i m_i (x_i' (x_i'^2 + z_i'^2) \omega_y - (y_i' z_i') \omega_z - (x_i' y_i') \omega_x), \\ \text{and } L_{rot,z} &= \sum_i m_i ((x_i'^2 + y_i'^2) \omega_z - (x_i' z_i') \omega_x - (y_i' z_i') \omega_y). \end{aligned} \quad (1.46)$$

In matrix form,

$$\mathbf{L}_{rot} = \begin{pmatrix} L_{rot,x} \\ L_{rot,y} \\ L_{rot,z} \end{pmatrix} = \begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix} \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix} = \tilde{\mathbf{I}}\boldsymbol{\omega}. \quad (1.47)$$

Similarly, for the abstract sum for the kinetic energy in eq. (1.36) due to the rotational motion, it now becomes

$$T_{rot} = \frac{1}{2} \sum_i m_i \dot{\mathbf{r}}_i'^2 = \frac{1}{2} \sum_i m_i (\boldsymbol{\omega} \times \mathbf{r}_i') \cdot (\boldsymbol{\omega} \times \mathbf{r}_i') = \frac{1}{2} \boldsymbol{\omega} \cdot \sum_i m_i \mathbf{r}_i' \times (\boldsymbol{\omega} \times \mathbf{r}_i') = \frac{1}{2} \boldsymbol{\omega} \cdot \mathbf{L} \quad (1.48)$$

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. (1.48), as it now becomes

$$T_{rot} = \frac{1}{2} I_{xx} \omega_x^2 + \frac{1}{2} I_{yy} \omega_y^2 + \frac{1}{2} I_{zz} \omega_z^2. \quad (1.49)$$

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

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

$$\begin{aligned} I_{xx,P} &= \sum_i m_i (y_{i,P}'^2 + z_{i,P}'^2) = \sum_i m_i ((y'_{i,c.m.} - Y)^2 + (z'_{i,c.m.} - Z)^2) \\ &= \sum_i m_i ((y_{i,c.m.}'^2 + z_{i,c.m.}'^2) + (Y^2 + Z^2) - 2(y_{i,c.m.}Y + z_{i,c.m.}Z)) \\ &= I_{xx,c.m.} + \sum_i m_i (Y^2 + Z^2) \end{aligned} \quad (1.50)$$

$$\begin{aligned} \text{and } I_{xy,P} &= - \sum_i m_i (x_{i,P} y_{i,P}') = \sum_i m_i ((x_{i,c.m.} - X)(y_{i,c.m.} - Y)) \\ &= \sum_i m_i ((x_{i,c.m.} y_{i,c.m.}) - XY - (X y_{i,c.m.} + x_{i,c.m.} Y)) \\ &= I_{xy,c.m.} - \sum_i m_i XY \text{ etc.} \end{aligned}$$

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

1.5.3 Euler's Equations

With all the prerequisites explained, we are now ready to tackle the seemingly simple differential equation $\boldsymbol{\tau} = \frac{d\mathbf{L}}{dt}$. 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 $\boldsymbol{\omega}dt$ 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

$$\begin{aligned}\boldsymbol{\tau} = \frac{d\mathbf{L}}{dt} &= \frac{d}{dt}(L_1\hat{\mathbf{1}} + L_2\hat{\mathbf{2}} + L_3\hat{\mathbf{3}}) = \frac{dL_1}{dt}\hat{\mathbf{1}} + L_1\frac{d\hat{\mathbf{1}}}{dt} + \frac{dL_2}{dt}\hat{\mathbf{2}} + L_2\frac{d\hat{\mathbf{2}}}{dt} + \frac{dL_3}{dt}\hat{\mathbf{3}} + L_3\frac{d\hat{\mathbf{3}}}{dt} \\ &= \frac{dL_1}{dt}\hat{\mathbf{1}} + \frac{dL_2}{dt}\hat{\mathbf{2}} + \frac{dL_3}{dt}\hat{\mathbf{3}} + (\boldsymbol{\omega} \times \hat{\mathbf{1}})L_1 + (\boldsymbol{\omega} \times \hat{\mathbf{2}})L_2 + (\boldsymbol{\omega} \times \hat{\mathbf{3}})L_3\end{aligned}\tag{1.51}$$

where $i = 1, 2$ and 3 and we used eq. (1.44) 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

$$\begin{aligned}\tau_1 &= I_1\dot{\omega}_1 + \omega_2\omega_3(I_3 - I_2) \\ \tau_2 &= I_2\dot{\omega}_2 + \omega_3\omega_1(I_1 - I_3) \\ \tau_3 &= I_3\dot{\omega}_3 + \omega_1\omega_2(I_2 - I_1).\end{aligned}\tag{1.52}$$

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 $\omega_1 = \text{constant}$ and

$$\frac{d^2\omega_2}{dt^2} + \left(\frac{(I_1 - I_2)(I_1 - I_3)}{I_2 I_3} \omega_1^2 \right) \omega_2 = 0 \quad (1.53)$$

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

^aFor 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.

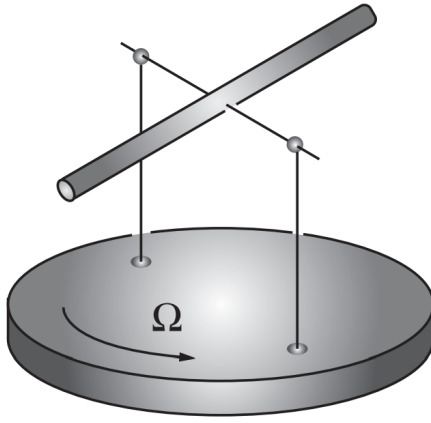
Solution: Referring to the figures, we have $\omega_1 = \dot{\theta}$, $\omega_2 = \Omega \sin \theta$ and $\omega_3 = \Omega \cos \theta$. Substituting them into the Euler's equations and leveraging the small angle approximation $\sin \theta \approx \theta$ gives

$$\ddot{\theta} + \left(\frac{I_3 - I_2}{I_1} \right) \Omega^2 \theta = 0. \quad (1.54)$$

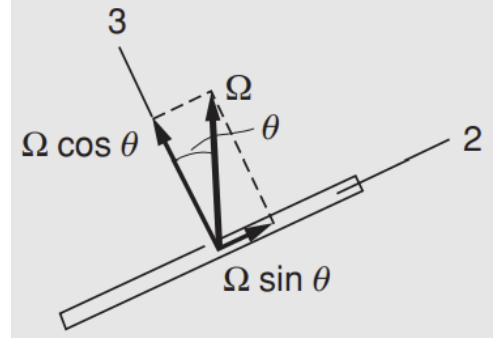
So we conclude that θ undergo simple harmonic motion with angular frequency $\gamma = \sqrt{\frac{I_3 - I_2}{I_1}} \Omega$.

1.6 Torque-free Precession

One of the most classic applications of Euler's equations is a torque-free precession. 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



(a)



(b)

Figure 1.3

$$\frac{d^2\omega_2}{dt^2} + \left(\frac{I_1 - I_\perp}{I_\perp}\right)^2 \omega_s^2 \omega_2 = 0. \quad (1.55)$$

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

$$\omega_2 = \omega_\perp \cos \gamma t \quad (1.56)$$

where ω_\perp depends on the initial condition .

Further calculation would give that

$$\omega_3 = \pm \omega_\perp \sin \gamma t \quad (1.57)$$

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 .

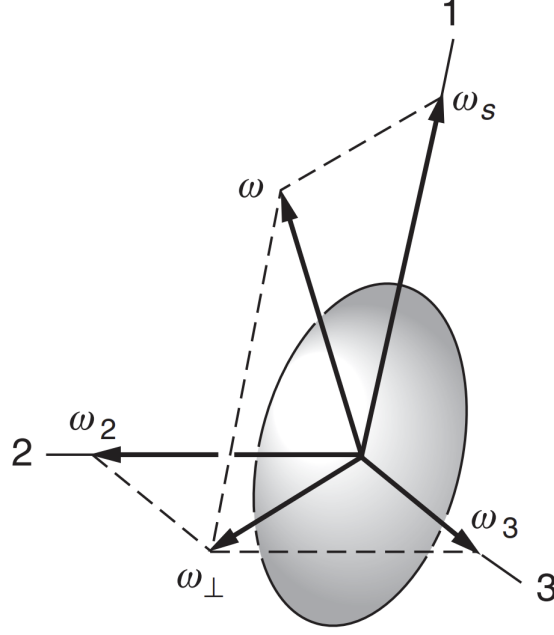


Figure 1.4

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

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

⁷This 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 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.

⁸The 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 $\omega_1 = \omega_s = \text{constant}$ and ω_\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. (1.48)) must be constant since there is no external work done, so α must be constant.

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 $\hat{\mathbf{l}}$ 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, $\hat{\mathbf{l}}$ 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 $\hat{\mathbf{l}}$ 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. (1.43), 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

$$\boldsymbol{\Omega}_{\boldsymbol{\omega} \text{ rel. lab}} = \boldsymbol{\Omega}_{\boldsymbol{\omega} \text{ rel. body}} - \boldsymbol{\Omega}_{\text{lab rel. body}}. \quad (1.58)$$

or

$$\Omega_p \hat{\mathbf{z}} = \gamma \hat{\mathbf{l}} - (-\boldsymbol{\omega}). \quad (1.59)$$

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

$$\begin{aligned} \Omega_p \cos \alpha &= \gamma + \omega_s \\ \Omega_p &= \frac{I_1 \omega_s}{I_\perp \cos \alpha}. \end{aligned} \quad (1.60)$$

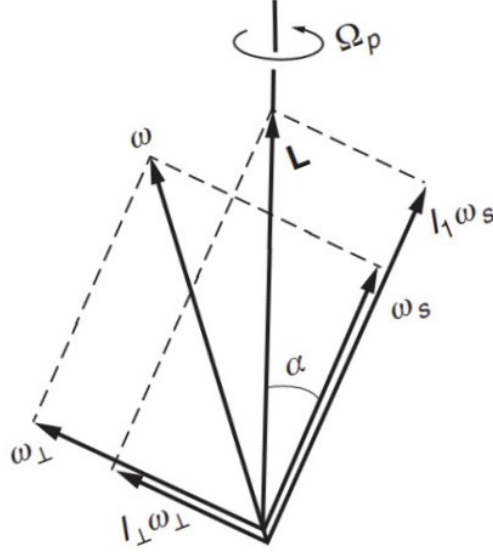


Figure 1.5

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

Firstly, as mentioned, $\omega_1 = \omega_s = \text{constant}$ 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

$$(\hat{\mathbf{1}}, \hat{\mathbf{2}}, \hat{\mathbf{3}}) = \boldsymbol{\lambda}(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}). \quad (1.61)$$

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.

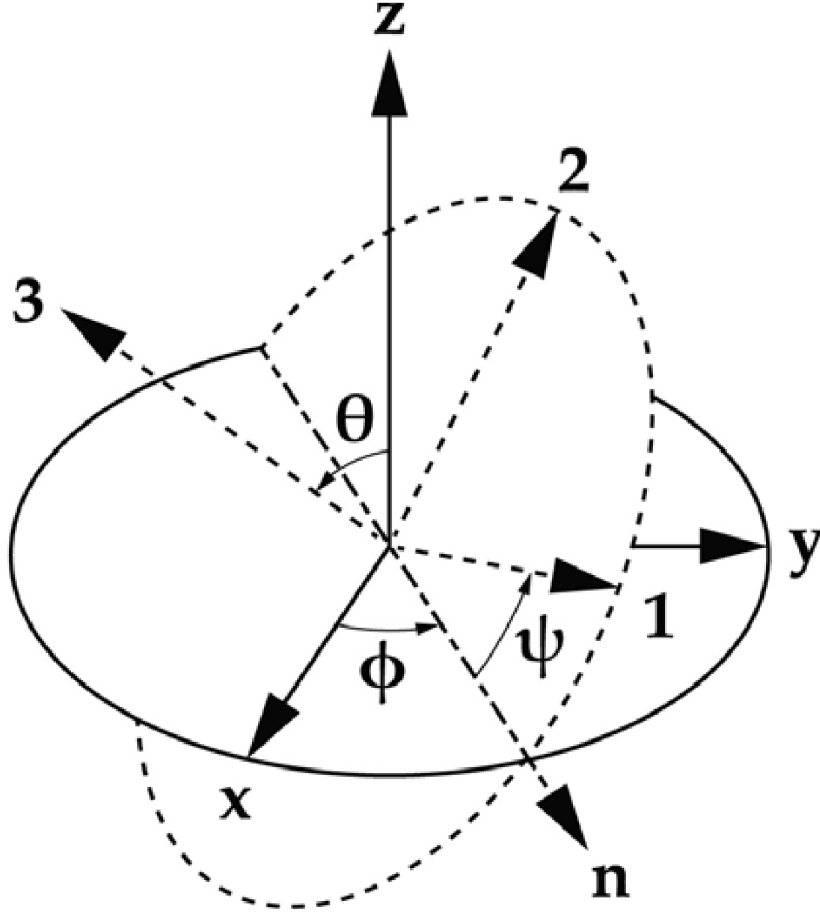


Figure 1.6

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 $\hat{\mathbf{z}}$ unchanged, $\hat{\mathbf{x}}$ 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

$$\boldsymbol{\lambda}_\phi = \begin{pmatrix} \cos \phi & \sin \phi & 0 \\ -\sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \boldsymbol{\lambda}_\theta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & \sin \theta \\ 0 & -\sin \theta & \cos \theta \end{pmatrix}, \quad \text{and} \quad \boldsymbol{\lambda}_\psi = \begin{pmatrix} \cos \psi & \sin \psi & 0 \\ -\sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (1.62)$$

Therefore the total rotational matrix is

$$\boldsymbol{\lambda} = \boldsymbol{\lambda}_\phi \boldsymbol{\lambda}_\theta \boldsymbol{\lambda}_\psi = \begin{pmatrix} \cos \phi \cos \psi - \sin \phi \cos \theta \sin \psi & \sin \phi \cos \psi + \cos \phi \cos \theta \sin \psi & \sin \theta \sin \psi \\ -\cos \phi \sin \psi - \sin \phi \cos \theta \cos \psi & -\sin \phi \sin \psi + \cos \phi \cos \theta \cos \psi & \sin \theta \cos \psi \\ \sin \phi \sin \theta & -\cos \phi \sin \theta & \cos \theta \end{pmatrix}. \quad (1.63)$$

The angular velocity will be

$$\boldsymbol{\omega} = \dot{\phi}\hat{\mathbf{z}} + \dot{\theta}\hat{\mathbf{n}} + \psi\hat{\mathbf{z}}. \quad (1.64)$$

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

$$\begin{aligned} \hat{\mathbf{z}} &= \sin \theta \sin \psi \hat{\mathbf{1}} + \sin \theta \cos \psi \hat{\mathbf{2}} + \cos \theta \hat{\mathbf{3}} \\ \hat{\mathbf{n}} &= \cos \psi \hat{\mathbf{1}} - \sin \psi \hat{\mathbf{2}}. \end{aligned} \quad (1.65)$$

So

$$\boldsymbol{\omega} = (\dot{\phi} \sin \theta \sin \psi + \dot{\theta} \cos \psi) \hat{\mathbf{1}} + (\dot{\phi} \sin \theta \cos \psi - \dot{\theta} \sin \psi) \hat{\mathbf{2}} + (\dot{\psi} + \dot{\phi} \cos \theta) \hat{\mathbf{3}}. \quad (1.66)$$

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

$$\boldsymbol{\omega} = (\dot{\theta} \cos \phi + \dot{\psi} \sin \theta \sin \phi) \hat{\mathbf{x}} + (\dot{\theta} \sin \phi - \dot{\psi} \sin \theta \cos \phi) \hat{\mathbf{y}} + (\dot{\phi} + \dot{\psi} \cos \theta) \hat{\mathbf{z}}. \quad (1.67)$$

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

$$\boldsymbol{\omega} \cdot \boldsymbol{\omega} = \omega_1^2 + \omega_2^2 + \omega_3^2 = \omega_x^2 + \omega_y^2 + \omega_z^2 = \dot{\phi}^2 + \dot{\theta}^2 + \dot{\psi}^2 + 2\dot{\phi}\dot{\psi} \cos \theta \quad (1.68)$$

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

$$T_{rot} = \frac{1}{2}(I_1(\dot{\phi} \sin \theta \sin \psi + \dot{\theta} \cos \psi)^2 + I_2(\dot{\phi} \sin \theta \cos \psi - \dot{\theta} \sin \psi)^2 + I_3(\dot{\phi} \cos \theta + \dot{\psi})^2). \quad (1.69)$$

2.1 Normal Modes

2.1.1 Basics

We start with the simple case with the equation

$$m\ddot{\mathbf{x}} = -K\mathbf{x}, \quad (2.1)$$

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 .

Substituting $\mathbf{x} = P\mathbf{q}$ into the equation, where $P = (\mathbf{v}_1, \dots, \mathbf{v}_n)$, we get

$$m\ddot{\mathbf{q}} = -K'\mathbf{q}, \quad (2.2)$$

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

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

Therefore the solution for \mathbf{x} is

$$\mathbf{x} = \mathbf{v}_1 \left(A_1 e^{i\sqrt{\frac{\lambda_1}{m}}t} + B_1 e^{-i\sqrt{\frac{\lambda_1}{m}}t} \right) + \dots + \mathbf{v}_n \left(A_n e^{i\sqrt{\frac{\lambda_n}{m}}t} + B_n e^{-i\sqrt{\frac{\lambda_n}{m}}t} \right). \quad (2.4)$$

Alternatively, we can guess $\mathbf{x} = \mathbf{v}e^{i\omega t}$ to get¹

$$m\omega^2\mathbf{v} = K\mathbf{v}, \quad (2.5)$$

¹This 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.

2.1.2 Unequal Masses

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

$$M\ddot{\mathbf{x}} = -K\mathbf{x}. \quad (2.6)$$

The generalized eigenvalues and eigenvectors can be found by solving

$$\det(K - \lambda_i M) = 0, \quad (2.7)$$

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

$$\mathbf{x} = \mathbf{v}_1 \left(A_1 e^{i\sqrt{\lambda_1}t} + B_1 e^{-i\sqrt{\lambda_1}t} \right) + \dots + \mathbf{v}_n \left(A_n e^{i\sqrt{\lambda_n}t} + B_n e^{-i\sqrt{\lambda_n}t} \right). \quad (2.8)$$

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

2.1.3 Damped and Forced Oscillators

If there is damping, then the equation becomes

$$m\ddot{\mathbf{x}} = -\gamma\dot{\mathbf{x}} - K\mathbf{x}, \quad (2.9)$$

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

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

$$m\ddot{\mathbf{x}} = -\gamma\dot{\mathbf{x}} - K\mathbf{x} + \mathbf{F}e^{i\omega t}, \quad (2.10)$$

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

$$\mathbf{C} = (K + (i\gamma\omega - \omega^2)\mathbb{I})^{-1} \mathbf{F} = (P(K' + (i\gamma\omega - \omega^2)\mathbb{I})P^{-1})^{-1} \mathbf{F} = PGP^{-1}\mathbf{F}, \quad (2.11)$$

where $G = \text{diag}((\lambda_i - \omega^2 + i\gamma\omega))^{-1}$. 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

$$\mathbf{C} \approx \mathbf{v}_1 \frac{\mathbf{v}_1^T \mathbf{F}}{i\gamma\omega} \implies \mathbf{x}_P \approx \mathbf{v}_1 \frac{\mathbf{v}_1^T \mathbf{F}}{\gamma\omega} \sin(\omega t) \quad (2.12)$$

2.1.4 Energy Conservation

Multiplying both sides of eq. (2.6) with $\dot{\mathbf{x}}^T$, we get

$$\frac{d}{dt} \left(\frac{\mathbf{x}^T M \mathbf{x}}{2} \right) = -\dot{\mathbf{x}}^T K \mathbf{x} = -\frac{d}{dt} \left(\frac{\mathbf{x}^T K \mathbf{x}}{2} \right), \quad (2.13)$$

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

$$E = \frac{1}{2}(\mathbf{x}^T M \mathbf{x}) + \frac{1}{2}(\mathbf{x}^T K \mathbf{x}) = \frac{1}{2}(\mathbf{q}^T M' \mathbf{q}) + \frac{1}{2}(\mathbf{q}^T K' \mathbf{q}). \quad (2.14)$$

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,

$$E = \frac{1}{2} \sum_{i=1}^n (m \dot{q}_i^2 + \lambda_i q_i^2). \quad (2.15)$$

2.1.5 N Masses

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

$$m \frac{d^2}{dt^2} \begin{pmatrix} \vdots \\ x_{n-1} \\ x_n \\ x_{n+1} \\ \vdots \end{pmatrix} = \begin{pmatrix} \vdots & & & & \\ \cdots & k & -2k & k & \\ & & k & -2k & k \\ & & & k & -2k & k & \cdots \\ & & & & \vdots & & \end{pmatrix} \begin{pmatrix} \vdots \\ x_{n-1} \\ x_n \\ x_{n+1} \\ \vdots \end{pmatrix}. \quad (2.16)$$

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

$$-\omega^2 v_n = \omega_0^2 (v_{n-1} - 2v_n + v_{n+1}), \quad (2.17)$$

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, and to accomodate for $v_0 = v_{n+1} = 0$, we require $B = 0$ and $\theta = nm\pi/(N+1)$, where $m = 1, \dots, N$. Substituting into the above equation, we have

$$\omega_m = 2\omega_0 \sin\left(\frac{m\pi}{2(N+1)}\right) \quad \text{and} \quad v_n = A \sin(nm\pi/(N+1)). \quad (2.18)$$

When $N \rightarrow \infty$, we have

$$\omega_m = \frac{m\pi}{L} \sqrt{\frac{T}{\rho}} = m\omega_1, \quad m \ll N. \quad (2.19)$$

The cases where $m = 1, 2, 3, 4, 5, N$ are shown in fig. 2.1.

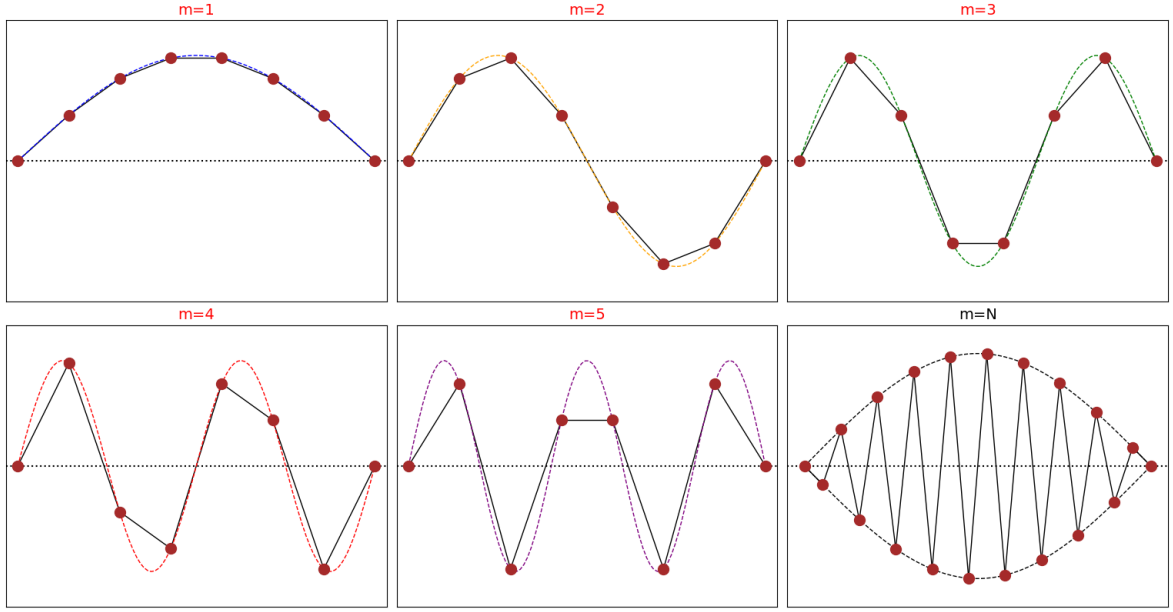


Figure 2.1

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 linear 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 at regular interval. On the other hand, the frequencies of each normal mode can be found by dividing a quarter circle with radius $2\omega_0$ into $N+1$ equal intervals, and finding the values of the resulting points. The case for $N = 3$ is illustrated in fig. 2.2.

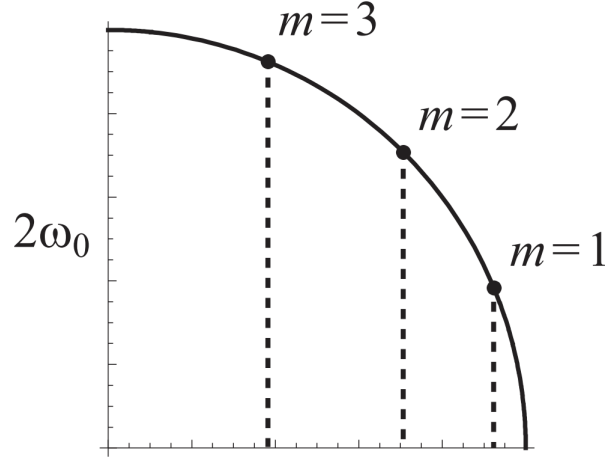


Figure 2.2

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 . 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 $N \rightarrow \infty$, then the equations become the wave equation

$$\rho \frac{\partial^2 \xi}{\partial t^2} = E \frac{\partial^2 \xi}{\partial x^2}, \quad (2.20)$$

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

$$\rho A \delta x \frac{\partial^2 \xi}{\partial t^2} = \delta F, \quad F = EA \frac{\partial \xi}{\partial x}, \quad (2.21)$$

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

2.2.2 Transverse Oscillations 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

$$\mu \frac{\partial^2 \psi}{\partial t^2} dx = T \frac{\partial^2 \psi}{\partial x^2}. \quad (2.22)$$

2.2.3 Acoustic Waves

Refer to fig. 2.3, from mass conservation we have

$$(\rho_0 + \rho_1)A(x + dx + \xi(x + dx, t) - (x + \xi(x, t))) = \rho_0 A dx \implies \frac{\rho_1}{\rho_0} = \frac{1}{1 + \frac{\partial \xi}{\partial x}} - 1 \approx -\frac{\partial \xi}{\partial x}. \quad (2.23)$$

where the relevant force is now

$$F = A(p_0 + p_1), \quad p_1 = \left(\frac{\rho_0 + \rho_1}{\rho_0} \right)^\gamma p_0 \approx \frac{\gamma \rho_1}{\rho_0} p_0 \approx -\gamma p_0 \frac{\partial \xi}{\partial x}. \quad (2.24)$$

Newton's second law thus gives

$$\rho_0 \frac{\partial^2 \xi}{\partial t^2} = \gamma p_0 \frac{\partial^2 \xi}{\partial x^2}. \quad (2.25)$$

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

$$F = -\kappa A \frac{\partial \xi}{\partial x} \implies \kappa = -V \frac{\partial p}{\partial V}. \quad (2.26)$$

so the wave equation can also be written as

$$\rho \frac{\partial^2 \xi}{\partial t^2} = \kappa \frac{\partial^2 \xi}{\partial x^2}. \quad (2.27)$$

For example, if the expansion is isothermal instead of adiabatic, then we have

$$\kappa = -V \left(-\frac{RT}{V^2} \right) = p \implies v = \sqrt{\frac{p_0}{\rho_0}}. \quad (2.28)$$

The velocity of the adiabatic compression case can also be derived generally since

$$\kappa = -V \left(-\frac{pV^\gamma \gamma}{V^{\gamma+1}} \right) = \gamma p \implies v = \sqrt{\frac{\gamma p_0}{\rho_0}}. \quad (2.29)$$

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

The characteristic impedance is

$$Z = \frac{p_1}{\frac{\partial \xi}{\partial t}} = \frac{\kappa \frac{\partial \xi}{\partial x}}{\frac{\partial \xi}{\partial t}} = \frac{\kappa k}{\omega} = \sqrt{\rho \kappa}. \quad (2.30)$$

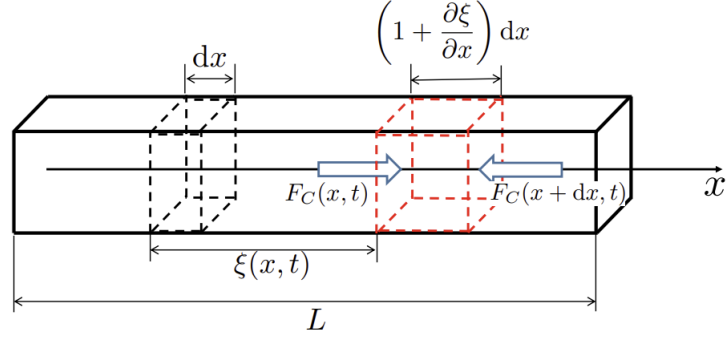


Figure 2.3

2.3 Solution to the Wave Equation

2.3.1 d'Alembert's Solution

General Solution

We start with the wave equation

$$\frac{\partial^2 y}{\partial t^2} = c^2 \frac{\partial^2 y}{\partial x^2}, \quad (2.31)$$

subject to the initial conditions

$$y(x, t = 0) = y_0(x) \quad \text{and} \quad \frac{\partial y}{\partial t}(x, t = 0) = \dot{y}_0(x). \quad (2.32)$$

The general solution of the wave equation is given by

$$y(x, t) = f(x - ct) + g(x + ct), \quad (2.33)$$

with velocity

$$\frac{\partial y(x, t)}{\partial t} = \dot{y}(x, t) = \frac{\partial f(x - ct)}{\partial t} + \frac{\partial g(x + ct)}{\partial t} = c(g'(x + ct) - f'(x - ct)). \quad (2.34)$$

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

$$f(x) + g(x) = y_0(x) \quad \text{and} \quad c(g'(x) - f'(x)) = \dot{y}_0(x). \quad (2.35)$$

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

$$f(x) - g(x) = -\frac{1}{c} \int^x \dot{y}_0(s) ds + C. \quad (2.36)$$

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

$$y(x, t) = \frac{1}{2} \left(y_0(x - ct) + y_0(x + ct) + \frac{1}{c} \int_{x-ct}^{x+ct} \dot{y}_0(s) ds \right). \quad (2.37)$$

Infinite String

Consider the initial condition

$$y_0(x) = \begin{cases} a \left(1 + \frac{x}{L}\right) & \text{for } -L \leq x < 0, \\ a \left(1 - \frac{x}{L}\right) & \text{for } 0 \leq x < L, \\ 0 & \text{otherwise.} \end{cases} \quad \text{and } \dot{y}_0(x) = 0. \quad (2.38)$$

We find that the solution is

$$y(x, t) = \frac{1}{2} (y_0(x - ct) + y_0(x + ct)), \quad (2.39)$$

as indicated in fig. 2.4. The red line is the actual displacement while the blue and purple lines are the right- and left-traveling waves respectively.

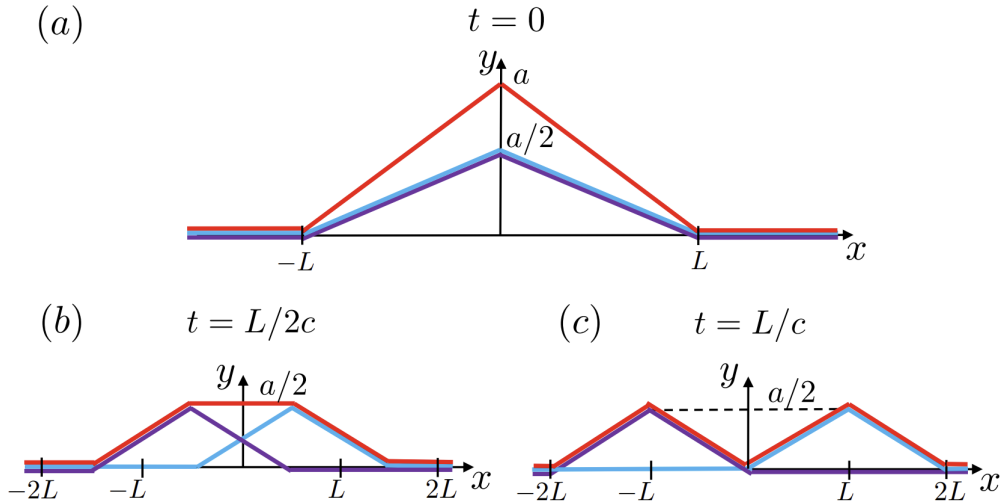


Figure 2.4

If the initial condition is given by

$$\dot{y}_0(x) = \begin{cases} V & \text{for } -L \leq x \leq L, \\ 0 & \text{otherwise,} \end{cases} \quad (2.40)$$

then the solution is

$$y(x, t) = \frac{1}{2c} \int_{x-ct}^{x+ct} \dot{y}_0(s) ds = \begin{cases} -VL/2c & \text{for } x < -L, \\ Vx/2c & \text{for } -L \leq x < L, \\ VL/2c & \text{for } x \geq L, \end{cases} \quad (2.41)$$

as indicated in fig. 2.5.

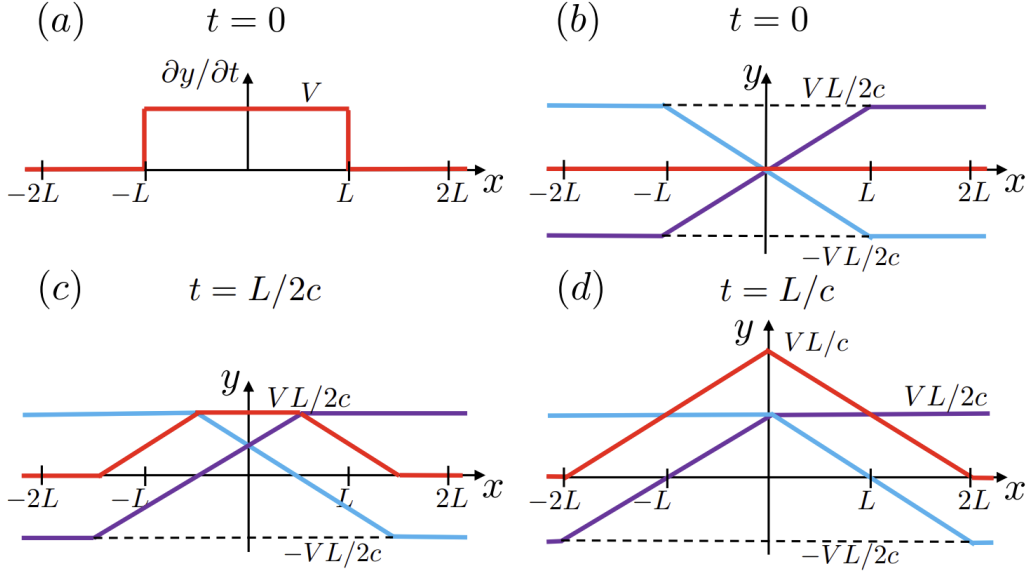


Figure 2.5

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

$$y(x, t) = f(x - ct) = y_0(x - ct) \implies \dot{y}_0(x) = \left. \frac{\partial y}{\partial t} \right|_{(x,0)} = -cy'_0(x). \quad (2.42)$$

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 $y \rightarrow 0$ as $x \rightarrow \pm\infty$.

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 net force acting on an infinitesimal mass element.

For example, for the initial condition

$$y_0(x) = \begin{cases} \frac{ax}{L} & \text{for } 0 \leq x < L, \\ a \left(2 - \frac{x}{L}\right) & \text{for } L \leq x < 2L, \\ 0 & \text{for } x \geq 2L, \end{cases} \quad (2.43)$$

and

$$\dot{y}_0(x) = cy'_0(x) = \begin{cases} \frac{ca}{L} & \text{for } 0 \leq x < L, \\ -\frac{ca}{L} & \text{for } L \leq x < 2L, \\ 0 & \text{for } x \geq 2L, \end{cases} \quad (2.44)$$

which is sketched in fig. 2.6, we would obtain for $f(u)$ and $g(v)$

$$f(u) = \begin{cases} ?? & \text{for } u < 0, \\ 0 & \text{for } u \geq 0. \end{cases} \quad \text{and} \quad g(v) = \begin{cases} ?? & \text{for } v < 0, \\ y_0(v) & \text{for } v \geq 0. \end{cases} \quad (2.45)$$

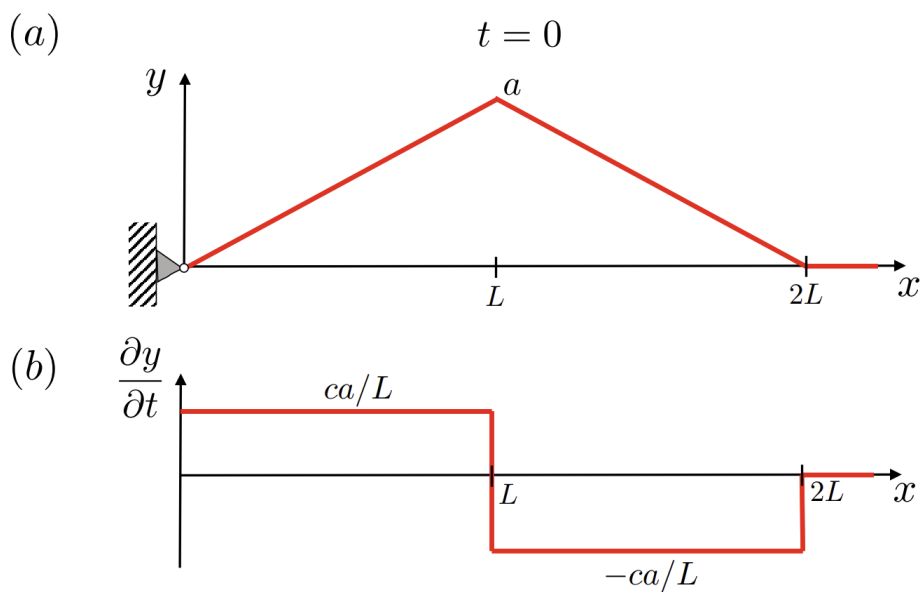


Figure 2.6

It is not until we impose the boundary condition $y = 0$ or $\partial y/\partial x = 0$ at $x = 0$ do we get $f(u) = -g(-u)$ or $f(u) = g(-u)$ and the solution is therefore determined, as showned in fig. 2.7 for the former case and fig. 2.8 for the latter case

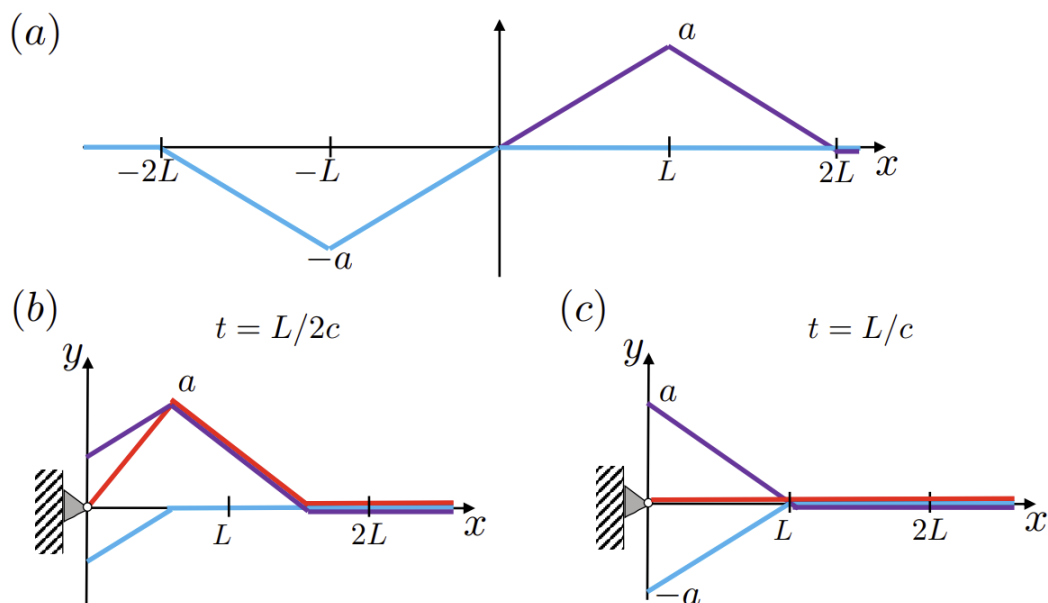


Figure 2.7

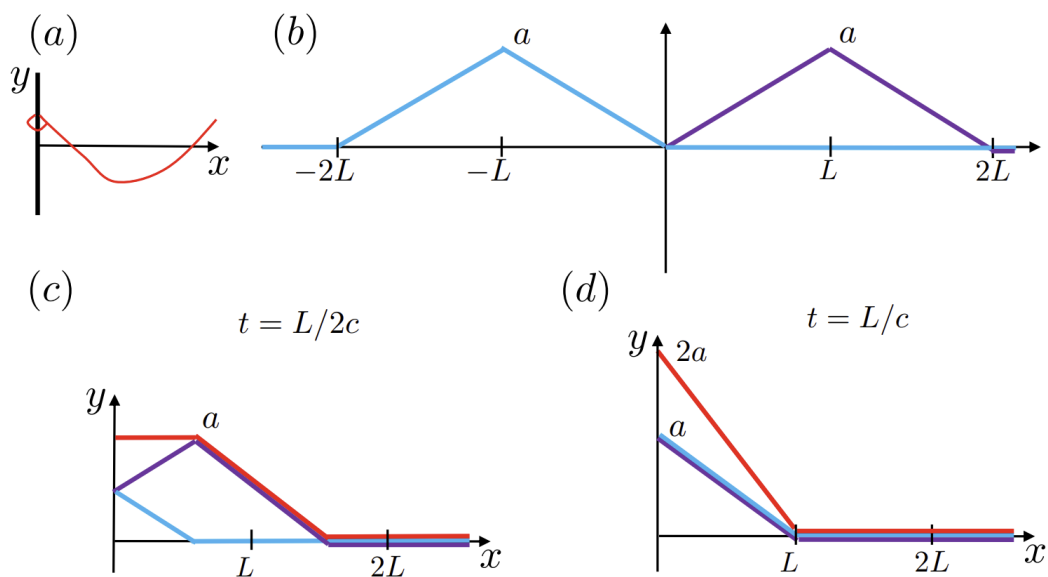


Figure 2.8

Finite String

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

$$y_0(x) = \begin{cases} ax/L & \text{for } 0 \leq x < L, \\ a(2 - x/L) & \text{for } L \leq x < 2L. \end{cases} \quad \text{and} \quad y_0'(x) = \begin{cases} ac/L & \text{for } 0 \leq x < L, \\ -ac/L & \text{for } L \leq x < 2L. \end{cases} \quad (2.46)$$

Using the d'Alembert's solution, we get

$$f(u) = \begin{cases} ?? & \text{for } u < 0, \\ 0 & \text{for } 0 \leq u \leq 2L. \end{cases} \quad \text{and} \quad g(v) = \begin{cases} y_0(v) & \text{for } 0 \leq v < 2L, \\ ?? & \text{for } v \geq 2L. \end{cases} \quad (2.47)$$

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

$$f(u) = \begin{cases} ?? & \text{for } u < 2L, \\ -y_0(-u) & \text{for } -2L \leq u < 0, \\ 0 & \text{for } 0 \leq u \leq 2L. \end{cases} \quad \text{and} \quad g(v) = \begin{cases} y_0(v) & \text{for } 0 \leq v < 2L, \\ 0 & \text{for } 2L \leq v < 4L, \\ ?? & \text{for } v \geq 4L. \end{cases} \quad (2.48)$$

The complete solution can thus be constructed step by step as shown in fig. 2.9.

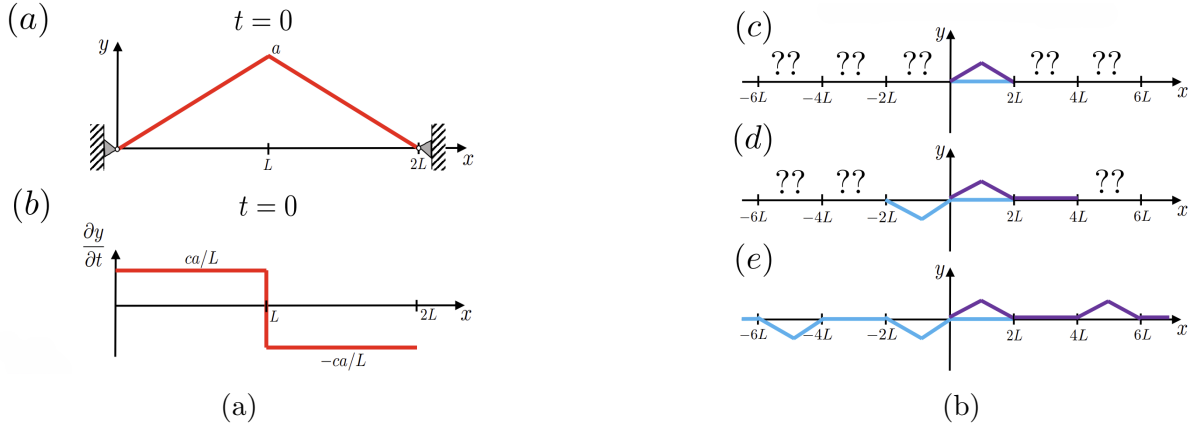


Figure 2.9

2.3.2 Separation of Variables

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

$$m\ddot{\mathbf{x}} = -K\mathbf{x}, \quad (2.49)$$

we guess $y = X(x)T(t)$ to the equation

$$\frac{\partial^2 y}{\partial t^2} = c^2 \frac{\partial^2 y}{\partial x^2}, \quad (2.50)$$

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

$$\frac{1}{c^2 T} \frac{d^2 T}{dt^2} = \frac{1}{X} \frac{d^2 X}{dx^2}. \quad (2.51)$$

Since the LHS and RHS are functions of different variables, they can only be equal if they equals to the same constant $-\Lambda^2$, where we have require the constant to be negative

because we expect oscillating behaviour but not exponentials, which in some case is what we expect such as in evanescent waves or in matter waves. Substituting, we have

$$\frac{d^2 T}{dt^2} = -(\Lambda c)^2 T \quad \text{and} \quad \frac{d^2 X}{dx^2} = -\Lambda^2 X, \quad (2.52)$$

which has the general solution

$$y(x, t) = T(t)X(x) = (Ae^{i\Lambda ct} + Be^{-i\Lambda ct}) (Ce^{i\Lambda x} + De^{-i\Lambda x}). \quad (2.53)$$

Solving the wave equation now becomes a matter of finding the coefficients A, B, C, D to satisfy the boundary conditions and initial conditions. The complete solution is built up of these stationary waves with different wavenumbers and frequencies.

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

$$C = -D \quad \text{and} \quad \sin(\Lambda L) = m\pi, \quad m = 1, 2, \dots \quad (2.54)$$

where negative m is not considered as we can combine the coefficients of negative m and positive m , while $k = 0$ is not considered since then the general solution given in eq. (2.53) 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-trivial then we would have to take this solution into account as well.

The general solution of the wave equation is therefore

$$y(x, t) = \sum_{m=1}^{\infty} \left(A \cos\left(\frac{m\pi ct}{L}\right) + B \sin\left(\frac{m\pi ct}{L}\right) \right) \sin\left(\frac{m\pi x}{L}\right), \quad (2.55)$$

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

$$y_0(x) = \sum_{m=1}^{\infty} Y_m \sin\left(\frac{m\pi x}{L}\right) \quad \text{and} \quad \dot{y}_0(x) = \sum_{m=1}^{\infty} Y'_m \sin\left(\frac{m\pi x}{L}\right), \quad (2.56)$$

and compare them with A and B to get

$$A_m = Y_m \quad \text{and} \quad \frac{m\pi c}{L} B_m = Y'_m. \quad (2.57)$$

kx-
omet/-
kx+omet
for for-
ward
and vice
versa

Example: A General Boundary Condition.

Question: Consider a stretched string attached to a vertical spring at $x = 0$ with spring constant K . Find the general solution to the wave equation.

Solution: The boundary condition is

$$T \frac{\partial y}{\partial x}(0, t) - Ky(0, t) = 0. \quad (2.58)$$

Substitute

$$y = Ae^{i(kx+\omega t)} + A^*e^{-i(kx+\omega t)} + Be^{i(kx-\omega t)} + B^*e^{-i(kx-\omega t)}, \quad (2.59)$$

we have

$$\frac{A}{B} = e^{i\varphi}, \quad \varphi = -\pi + 2 \tan^{-1} \left(\frac{kT}{K} \right). \quad (2.60)$$

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

$$y_1(0, t) = y_2(0, t) \quad \text{and} \quad \frac{\partial y_1}{\partial x}(0, t) = \frac{\partial y_2}{\partial x}(0, t). \quad (2.61)$$

We substitute

$$\begin{aligned} y_1 &= Ae^{i(k_1x+\omega_1t)} + A^*e^{-i(k_1x+\omega_1t)} + Be^{i(k_1x-\omega_1t)} + B^*e^{-i(k_1x-\omega_1t)}, \\ y_2 &= Ce^{i(k_2x-\omega_2t)} + C^*e^{-i(k_2x-\omega_2t)}, \end{aligned} \quad (2.62)$$

where y_2 has two less terms since there is no left-travelling wave on string 2. Alternatively one can just substitute, for example, $y_2 = Ce^{i(k_2x-\omega_2t)}$, where we take the real part of this solution at the end. Imposing the boundary conditions we get

$$\omega_1 = \omega_2, \quad A + B = C \quad \text{and} \quad k_1(A - B) = k_2C. \quad (2.63)$$

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

$$r = \frac{k_1 - k_2}{k_1 + k_2} \quad \text{and} \quad t = \frac{2k_1}{k_1 + k_2}. \quad (2.64)$$

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

$$y_1(0, t) = y_2(0, t) \quad \text{and} \quad T_1 \frac{\partial y_1}{\partial x}(0, t) = T_2 \frac{\partial y_2}{\partial x}(0, t). \quad (2.65)$$

and the reflection and transmission coefficients becomes

$$r = \frac{k_1 T_1 - k_2 T_2}{k_1 T_1 + k_2 T_2} \quad \text{and} \quad t = \frac{2k_1 T_1}{k_1 T_1 + k_2 T_2}. \quad (2.66)$$

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

$$F_y = T_2 \frac{\partial y_2}{\partial x}(0, t) = -\frac{T_2}{v_2} \frac{\partial y_2}{\partial t}(0, t) = -Z_2 v_y. \quad (2.67)$$

Example: Stretched Strings connecte by a Mass.

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

Solution: The boundary conditions now becomes

$$y_1(0, t) = y_2(0, t) \quad \text{and} \quad T \frac{\partial y_2}{\partial x}(0, t) - T \frac{\partial y_1}{\partial x}(0, t) = M \frac{\partial y_1}{\partial t}(0, t) = M \frac{\partial y_2}{\partial t}(0, t). \quad (2.68)$$

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

$$r = \frac{(k_1 - k_2)T - i\omega^2 M}{(k_1 + k_2)T + i\omega^2 M} = |r|e^{i\theta} \quad \text{and} \quad t = \frac{2k_1 T}{(k_1 + k_2)T + i\omega^2 M} = |t|e^{i\phi}. \quad (2.69)$$

The energy of the system is conserved since

$$k_1 |r|^2 + k_2 |t|^2 = k_1. \quad (2.70)$$

Example: Impedances in Transmission lines.

Question: Refer to fig. 2.10, 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

$$\delta V = \frac{1}{C} \delta Q = -\frac{1}{C} \frac{\partial I}{\partial x} \delta t. \quad (2.71)$$

From the definition of inductance we also have

$$\delta V = -L \delta x \frac{\partial I}{\partial t}. \quad (2.72)$$

Combining the two equations yield the wave equation

$$\frac{\partial^2 V}{\partial x^2} = LC \frac{\partial^2 V}{\partial t^2} \quad \text{and} \quad \frac{\partial^2 I}{\partial x^2} = LC \frac{\partial^2 I}{\partial t^2}. \quad (2.73)$$

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

$$Z_0 = \frac{V_0}{I_0} = \frac{\omega L}{k} = \sqrt{\frac{L}{C}}, \quad (2.74)$$

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

$$V(0, t) = Z_T I(0, t). \quad (2.75)$$

Substituting

$$\begin{aligned} V(x, t) &= A e^{i(\omega t - kx)} + A^* e^{-i(\omega t - kx)} + B e^{i(\omega t + kx)} + B^* e^{-i(\omega t + kx)}, \\ Z_0 I(x, t) &= A e^{i(\omega t - kx)} + A^* e^{-i(\omega t - kx)} - B e^{i(\omega t + kx)} - B^* e^{-i(\omega t + kx)}, \end{aligned} \quad (2.76)$$

we get the reflection coefficient

$$r = \frac{Z_T - Z_0}{Z_T + Z_0}. \quad (2.77)$$

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

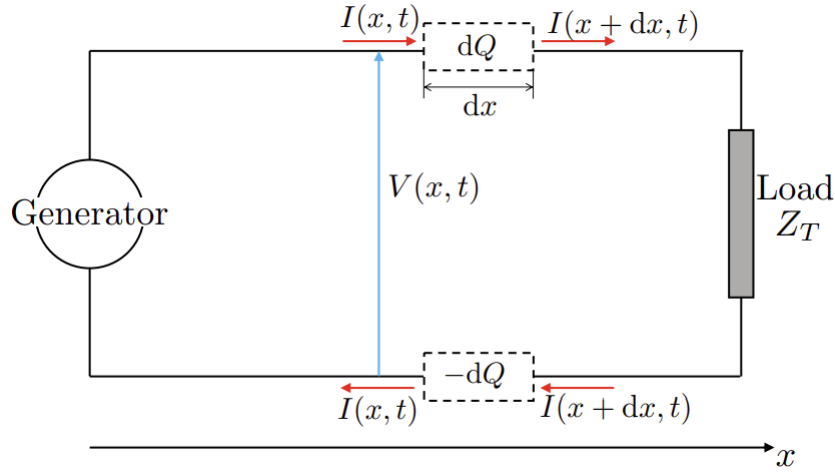


Figure 2.10

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. (2.53) 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

$$dl = \sqrt{dx^2 + d\psi^2} \approx dx + \frac{1}{2} \left(\frac{d\psi}{dx} \right)^2 dx, \quad (2.78)$$

so the energy density of the string is

$$\epsilon = \frac{1}{2} \mu \left(\frac{\partial \psi}{\partial t} \right)^2 + \frac{1}{2} T \left(\frac{\partial \psi}{\partial x} \right)^2. \quad (2.79)$$

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

$$\epsilon_{\text{avg}} = \frac{1}{4} \mu A^2 \omega^2 + \frac{1}{4} T A^2 k^2 = \frac{1}{2} \mu A^2 \omega^2 = \frac{1}{2} T A^2 k^2. \quad (2.80)$$

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

$$\frac{\partial \epsilon}{\partial t} = \mu \frac{\partial y}{\partial t} \frac{\partial^2 y}{\partial t^2} + T \frac{\partial y}{\partial x} \frac{\partial^2 y}{\partial t \partial x} = T \left(\frac{\partial y}{\partial t} \frac{\partial^2 y}{\partial x^2} + \frac{\partial y}{\partial x} \frac{\partial^2 y}{\partial t \partial x} \right) = \frac{\partial}{\partial x} \left(T \frac{\partial y}{\partial x} \frac{\partial y}{\partial t} \right). \quad (2.81)$$

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

$$\frac{dE}{dt} = \left(T \frac{\partial y}{\partial x} \frac{\partial y}{\partial t} \right) \Big|_{x=0}^{x=L} = (F_y v_y) \Big|_{x=0}^{x=L} \quad (2.82)$$

Thus at each point Fv_y is the power flowing from the left to the right.

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

$$\epsilon = \frac{1}{2} \left(\mu c (g' - f')^2 + T(f' + g')^2 \right) = T(f'^2 + g'^2) = \epsilon_f + \epsilon_g, \quad (2.83)$$

and its rate of change as

$$\frac{\partial \epsilon}{\partial t} = Tc (g'^2 - f'^2) = c(\epsilon_g - \epsilon_f). \quad (2.84)$$

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

$$\frac{dE}{dt} = c (\epsilon_g - \epsilon_f) \Big|_{x=0}^{x=L}. \quad (2.85)$$

One can refer to the sketches in fig. 2.11 to visualize the terms.

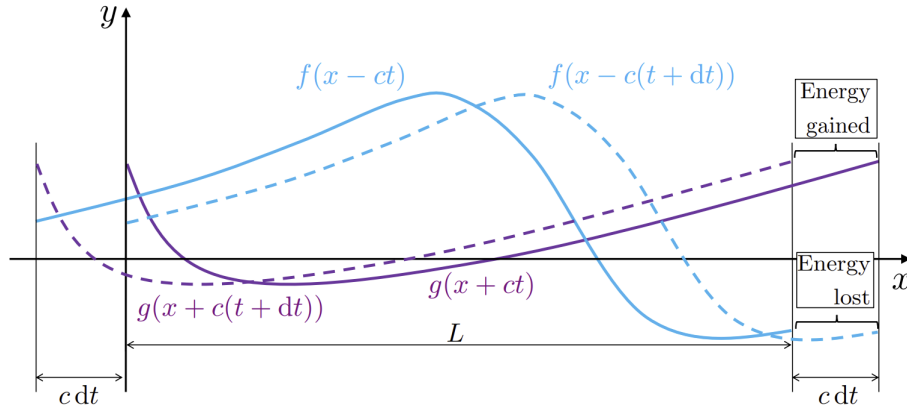


Figure 2.11

Note that the wave equation can be rewritten as

$$\frac{\partial \epsilon_f}{\partial t} + c \frac{\partial \epsilon_f}{\partial x} = 0, \quad (2.86)$$

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

$$y(x, t) = A_0 + A_1 t + \sum_{k=1}^{\infty} C_k \cos \left(\sqrt{-\Lambda_k} ct - \varphi_k \right) Q_k(x), \quad (2.87)$$

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

The energy of a string is

$$\begin{aligned} E &= \frac{1}{2} \left(\mu \int_0^L \left(\frac{\partial y}{\partial t} \right)^2 dx + T \int_0^L \left(\frac{\partial y}{\partial x} \right)^2 dx \right) = \frac{1}{2} \int_0^L \left(\mu \left(\frac{\partial y}{\partial t} \right)^2 + T y \left(\frac{\partial^2 y}{\partial x^2} \right) \right) dx \\ &= \frac{\mu}{2} \int_0^L \left(A_1 - c \sum_{k=1}^{\infty} \sqrt{-\Lambda_k} C_k \sin \left(\sqrt{-\Lambda_k} ct - \varphi_k \right) Q_k(x) \right)^2 dx \\ &\quad - \frac{T}{2} \int_0^L \left(\sum_{k=1}^{\infty} C_k \cos \left(\sqrt{-\Lambda_k} ct - \varphi_k \right) Q_k(x) \right) \left(\sum_{l=1}^{\infty} \Lambda_l C_l \cos \left(\sqrt{-\Lambda_l} ct - \varphi_l \right) Q_l(x) \right) dx. \end{aligned} \quad (2.88)$$

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 $C_k C_l \int_0^L Q_k(x) Q_l(x) dx$, but we will show that all the integrals are zero except for integrals in the form $C_k^2 \int_0^L Q_k(x)^2 dx$.

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

$$\int_0^L Q_l \frac{d^2 Q_k}{dx^2} dx = - \int_0^L \frac{dQ_l}{dx} \frac{dQ_k}{dx} dx = \Lambda_k \int_0^L Q_l Q_k dx, \quad (2.89)$$

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

$$\int_0^L Q_l Q_k dx = 0, \quad k \neq l. \quad (2.90)$$

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

$$E = \frac{1}{2} \left(\mu L A_1^2 - T \sum_{k=1}^{\infty} \Lambda_k C_k^2 \int_0^L Q_k^2(x) dx \right), \quad (2.91)$$

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

$$E = \frac{\mu L A_1^2}{2} - \frac{TL}{4} \sum_{k=1}^{\infty} \Lambda_k C_k^2. \quad (2.92)$$

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

$$P_1 = T \frac{\partial y_1}{\partial x} \frac{\partial y_1}{\partial t} = -T(-k_1 A + k_1 r A)(\omega A + \omega r A) = T k_1 \omega (A^2 - r^2 A^2) = \frac{4\omega^2 A^2 \rho_1 \sqrt{\rho_2} T}{(\sqrt{\rho_1} + \sqrt{\rho_2})^2}, \quad (2.93)$$

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

$$P_2 = T \frac{\partial y_2}{\partial x} \frac{\partial y_2}{\partial t} = T(k_2 t A)(\omega t A) = \frac{4\omega^2 A^2 \rho_1 \sqrt{\rho_2} T}{(\sqrt{\rho_1} + \sqrt{\rho_2})^2}. \quad (2.94)$$

Thus energy is conserved.

2.5 Dispersive Waves

Waves that does not satisfy the wave equation eq. (2.31) are dispersive waves, mainly because the relations between the frequencies ω and wavenumbers k are not so simple anymore.

The genearl 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

$$v_p(k) = \frac{\omega(k)}{k} \quad (2.95)$$

is the velocity in which the wave with wavenumber k is moving. 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. (2.18)

$$\omega_m = 2\omega_0 \sin\left(\frac{m\pi}{2(N+1)}\right) = 2\omega_0 \sin\left(\frac{kl}{2}\right), \quad (2.96)$$

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

2.5.2 Group Velocity

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 $p(x, y, t) = p_a - \rho gy + p_1(x, y, t)$, where p_1 is the excess pressure.

Mass conservation gives

$$dxdydz = \left(\left(1 + \frac{\partial \xi}{\partial x}\right) \left(1 + \frac{\partial \eta}{\partial y}\right) - \frac{\partial \xi}{\partial y} \frac{\partial \eta}{\partial x} \right) dxdydz \implies \frac{\partial \xi}{\partial x} + \frac{\partial \eta}{\partial y} = 0 \quad (2.97)$$

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

$$\rho \frac{d^2 \xi}{dt^2} = -\frac{\partial p_1}{\partial x} \quad \text{and} \quad \rho \frac{\partial^2 \eta}{\partial t^2} = \frac{\partial p_1}{\partial y} - \rho g. \quad (2.98)$$

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

$$\rho \frac{\partial^2}{\partial t^2} = \left(\frac{\partial \xi}{\partial x} + \frac{\partial \eta}{\partial y} \right) = -\frac{\partial^2 p_1}{\partial x^2} - \frac{\partial^2 p_1}{\partial y^2} \implies \frac{\partial^2 p_1}{\partial x^2} + \frac{\partial^2 p_1}{\partial y^2} = 0. \quad (2.99)$$

Assuming $h(x, t) = \eta(x, y = 0, t) \propto e^{i(kx - \omega t)}$ for typical waves, we have from eq. (2.101) that $p_1(x, y, t) = P_1(y)e^{i(kx - \omega t)}$, which upon substitution gives

$$\frac{d^2 P_1}{dy^2} = k^2 P_1 \implies p_1(x, y, t) = p_1(x, y = 0, t)e^{|k|y}, \quad (2.100)$$

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

Refer to fig. 2.12, by balancing the vertical forces acting on an infinitesimal volume of water on the interface, we have

Lagrangian mechanics

3.1 Euler-Lagrange Equation

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

$$S = \int_{t_1}^{t_2} \mathcal{L}(q(t), \dot{q}(t), t) dt \quad (3.1)$$

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

$$\delta q(t) = q(\epsilon, t) - q(0, t) = \epsilon \eta(t) \implies \delta \dot{q}(t) = \epsilon \dot{\eta}(t), \quad (3.2)$$

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

$$\eta(t_1) = \eta(t_2) = 0. \quad (3.3)$$

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

$$\delta \mathcal{L} = \frac{\partial \mathcal{L}}{\partial q} \delta q + \frac{\partial \mathcal{L}}{\partial \dot{q}} \delta \dot{q} = \epsilon \eta(t) \frac{\partial \mathcal{L}}{\partial q} + \epsilon \dot{\eta} \frac{\partial \mathcal{L}}{\partial \dot{q}} \quad (3.4)$$

The corresponding variation in the action S is

$$\delta S = \epsilon \int_{t_1}^{t_2} \left(\frac{\partial \mathcal{L}}{\partial q} q(t) + \frac{\partial \mathcal{L}}{\partial \dot{q}} \dot{q}(t) \right) dt = \epsilon \int_{t_1}^{t_2} \left(\frac{\partial \mathcal{L}}{\partial q} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}} \right) \right) \eta(t) dt, \quad (3.5)$$

where in the last equality we used integration by part and the boundary term vanishes due to the eq. (3.3).

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

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

$$\frac{\partial q}{\partial t} = \frac{d}{dt} \frac{\partial \dot{q}}{\partial t}, \quad (3.6)$$

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

If the Lagrangian \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.

To derive the Beltrami identity when $\frac{\partial \mathcal{L}}{\partial q} = 0$

$$\mathcal{L} - \dot{q} \frac{\partial \mathcal{L}}{\partial \dot{q}} = \text{constant}, \quad (3.7)$$

we note that

$$\frac{d\mathcal{L}}{dt} = \frac{\partial \mathcal{L}}{\partial t} + \frac{\partial \mathcal{L}}{\partial q} \frac{dq}{dt} + \frac{\partial \mathcal{L}}{\partial \dot{q}} \frac{d\dot{q}}{dt} \quad \text{and} \quad \frac{d}{dt} \left(\dot{q} \frac{\partial \mathcal{L}}{\partial \dot{q}} \right) = \frac{\partial \mathcal{L}}{\partial \dot{q}} \frac{d\dot{q}}{dt} + \dot{q} \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}} \right), \quad (3.8)$$

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 equations gives

$$\frac{d}{dt} \left(\dot{q} \frac{\partial \mathcal{L}}{\partial \dot{q}} \right) = \frac{d\mathcal{L}}{dt} - \frac{\partial \mathcal{L}}{\partial t} + \dot{q} \left(\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}} - \frac{\partial \mathcal{L}}{\partial q} \right) \implies \frac{\partial \mathcal{L}}{\partial t} = \frac{d}{dt} \left(\mathcal{L} - \dot{q} \frac{\partial \mathcal{L}}{\partial \dot{q}} \right). \quad (3.9)$$

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

$$L = \int_{x_1}^{x_2} \sqrt{1 + \left(\frac{dy}{dx} \right)^2} dx, \quad (3.10)$$

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

$$\frac{\partial \mathcal{L}}{\partial y} = 0 = \frac{\partial \mathcal{L}}{\partial y'} \implies y = ax + b. \quad (3.11)$$

Example: Brachistochrone Problem.

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

Solution: We want to minimize the transit time

$$t = \int_{x_1}^{x_2} \frac{ds}{v} = \int_{y_1}^{y_2} \frac{\sqrt{1+x'^2}}{2gy} dy, \quad (3.12)$$

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

$$\frac{\partial t}{\partial x} = 0 = \frac{d}{dy} \frac{\partial t}{\partial x'}. \quad (3.13)$$

Solving gives

$$x = a(\theta - \sin \theta) \quad \text{and} \quad y = a(1 - \cos \theta), \quad (3.14)$$

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 minimize the flying cost

$$C = \int_{-a}^a e^{-\kappa z} ds = \int_{-a}^a e^{-\kappa z} \sqrt{1+z'^2} dx. \quad (3.15)$$

Applying the Euler-Lagrange equation gives

$$\frac{\partial \mathcal{L}}{\partial z} = -\kappa e^{-\kappa z} \sqrt{1+z'^2} = \frac{d}{dx} \frac{\partial \mathcal{L}}{\partial z'} = \frac{z'' e^{-\kappa z}}{\sqrt{1+z'^2}} - \frac{\kappa z'^2 e^{-\kappa z}}{\sqrt{1+z'^2}} - \frac{z'' z'^2 e^{-\kappa z}}{(1+z'^2)^{3/2}}. \quad (3.16)$$

Solving gives

$$z(x) = \frac{1}{\kappa} \ln \frac{\cos(\kappa x)}{\cos(\kappa a)}. \quad (3.17)$$

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

$$S = 2\pi \int \rho \sqrt{dz^2 + d\rho^2} = 2\pi \int_{z_1}^{z_2} \rho \sqrt{1 + \left(\frac{d\rho}{dz}\right)^2} dz = 2\pi \int_{\rho_1}^{\rho_2} \rho \sqrt{1 + \left(\frac{dz}{d\rho}\right)^2} d\rho. \quad (3.18)$$

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

$$\frac{\partial \mathcal{L}}{\partial \rho} = \sqrt{1 + \rho'^2} = \frac{d}{dz} \frac{\partial \mathcal{L}}{\partial \rho'} = \frac{d}{dz} \left(\frac{\rho \rho'}{\sqrt{1 + \rho'^2}} \right), \quad (3.19)$$

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

$$\frac{\partial \mathcal{L}}{\partial z} = 0 = \frac{d}{d\rho} \frac{\partial \mathcal{L}}{\partial z'} = \frac{d}{d\rho}. \quad (3.20)$$

Solving gives

$$\rho = a \cosh \left(\frac{z - b}{a} \right), \quad (3.21)$$

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

Example: Fermat's Principle

Question: Refer to fig. 3.1 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

$$t = \int_{y_1}^{y_2} \frac{ds}{v} = \int_{y_1}^{y_2} \frac{1}{c} n(x, y, z) \sqrt{1 + x'^2 + z'^2} dy \quad (3.22)$$

The Euler-Lagrange equation for z reads

$$0 + \frac{d}{dy} \left(\frac{1}{c} \left(\frac{n_1 z'}{\sqrt{1 + x'^2 + z'^2}} + \frac{n_2 z'}{\sqrt{1 + x'^2 + z'^2}} \right) \right) = 0 \implies z' = 0, \quad (3.23)$$

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

$$0 + \frac{d}{dy} \left(\frac{1}{c} \left(\frac{n_1 \tan \theta_1}{\sqrt{1 + \tan^2 \theta_1}} - \frac{n_2 \tan \theta_2}{\sqrt{1 + \tan^2 \theta_2}} \right) \right) = \frac{d}{dy} \left(\frac{1}{c} (n_1 \sin \theta_1 - n_2 \sin \theta_2) \right) = 0. \quad (3.24)$$

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-Lagrange equation is faster. However due to the simplicity of this approach it is not illustrated here.

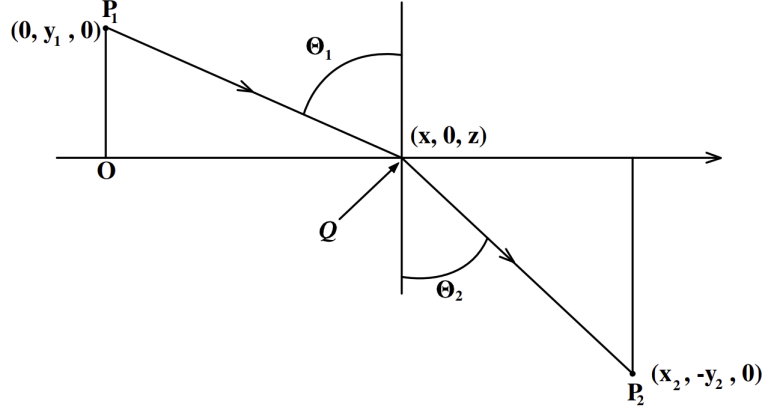


Figure 3.1

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

$$J = \frac{1}{V} \int (\nabla\phi)^2 dV = \frac{1}{V} \int \left(\left(\frac{\partial\phi}{\partial x} \right)^2 + \left(\frac{\partial\phi}{\partial y} \right)^2 + \left(\frac{\partial\phi}{\partial z} \right)^2 \right). \quad (3.25)$$

The Euler-Lagrange equation for the three coordinates read

$$\frac{\partial\mathcal{L}}{\partial\phi} = 0 = \frac{\partial}{\partial x_i} \left(\frac{\partial\mathcal{L}}{\partial\phi'} \right) \Rightarrow \frac{\partial^2\phi}{\partial x_i^2} = 0 \Rightarrow \sum_{i=1}^3 \frac{\partial^2\phi}{\partial x_i^2} = 0 \Rightarrow \nabla^2\phi = 0, \quad (3.26)$$

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

3.2 Constraints

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

$$f(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots, t) = 0, \quad (3.27)$$

For example, the constraint on a rigid body is that the distance between 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 equations. If there are k holonomic constraints, then we have $3N + k$ equations can be used to solve for the $3N$ (now dependent) variables as well as the k constraint forces. In other words, we can use the k constraint equations to eliminate k dependent variables

and introduce $3N - k$ independent variables (also known as generalized coordinates²) $q_1, q_2, \dots, q_{3N-k}$ in terms of the dependent variables r_1, r_2, \dots, r_{3N} and time t .

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, referring to fig. 3.2, 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 .

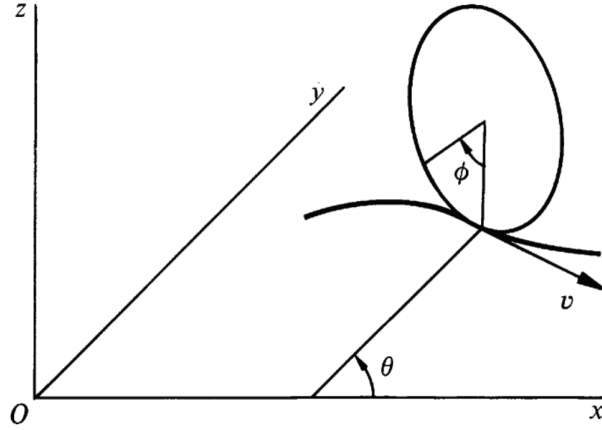


Figure 3.2

The angular velocity of the disc is

$$\boldsymbol{\omega} = \dot{\theta} \hat{\mathbf{z}} + \dot{\phi} (\cos \theta \hat{\mathbf{x}} + \sin \theta \hat{\mathbf{y}}), \quad (3.28)$$

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.*,

$$\begin{aligned} \mathbf{v}_{\text{c.m.}} + \boldsymbol{\omega} \times \mathbf{r}_{\text{contact}} &= 0 \\ (x\hat{\mathbf{x}} + y\hat{\mathbf{y}}) + \left(\dot{\theta}\hat{\mathbf{z}} + \dot{\phi}(\cos \theta \hat{\mathbf{x}} + \sin \theta \hat{\mathbf{y}}) \right) \times (-a\hat{\mathbf{z}}) &= 0 \\ (x\hat{\mathbf{x}} + y\hat{\mathbf{y}}) + a\dot{\phi} \cos \theta \hat{\mathbf{y}} - a\dot{\phi} \sin \theta \hat{\mathbf{x}} &= 0 \\ (x - a\dot{\phi} \sin \theta)\hat{\mathbf{x}} + (y + a\dot{\phi} \cos \theta)\hat{\mathbf{y}} &= 0 \\ \mathbf{x} = a\dot{\phi} \sin \theta \quad \text{and} \quad \mathbf{y} = -a\dot{\phi} \cos \theta. \end{aligned} \quad (3.29)$$

²In some sense, the number of generalized coordinates is the minimum number of variables that can fully describe the state of the system.

However, these equations of constraint 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 constraint is a particle being put on the surface of a sphere of radius a . In this case the constraint appears in the form of an inequality $x^2 + y^2 + z^2 > a$.

The constraint 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 .

3.3 Lagrange Multipliers for Holonomic Constraints.

Beside using generalized coordinates to solve for the equation of motions with presence of constraint forces as detailed in section 3.2, we can use the method of Lagrange multipliers.

Appendices

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.*,

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

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

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

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

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

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

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

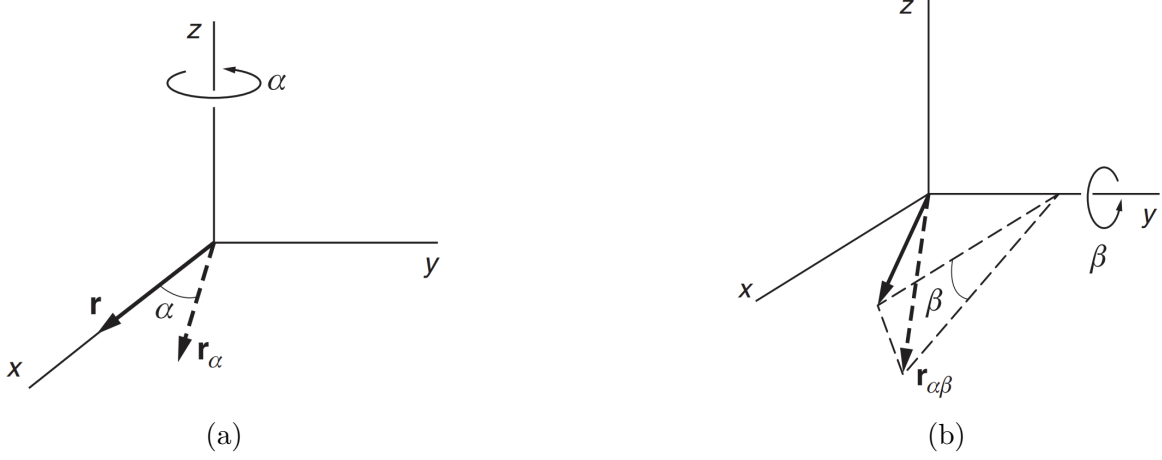


Figure A.1

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

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

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

The linear velocity will then be

$$\mathbf{v} = \lim_{\Delta t \rightarrow 0} \frac{\Delta \mathbf{r}}{\Delta t} = \lim_{\Delta t \rightarrow 0} \frac{\Delta \boldsymbol{\theta} \times \mathbf{r}}{\Delta t} = \boldsymbol{\omega} \times \mathbf{r}. \quad (\text{A.6})$$

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

$$\boldsymbol{\omega} = \lim_{\Delta t \rightarrow 0} \frac{\Delta \boldsymbol{\theta}}{\Delta t} \quad (\text{A.7})$$

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