

Principles of Dynamics

Ed Copeland¹ and Antonio Padilla²

School of Physics and Astronomy, University of Nottingham, Nottingham NG72RD, UK

Abstract

These lecture notes cover the Principles of Dynamics second year optional module (F32OT1). They provide an introduction to Lagrangian mechanics, including Calculus of Variations, Hamilton's principle of least action, and generalised coordinates. The ideas are applied to various systems including planetary orbits and the motion of rigid bodies. Constrained systems and the use of Lagrange multipliers is also discussed. The dynamics of Charged particles in an electromagnetic field are obtained. The Hamiltonian is introduced and Hamilton's equations derived. The equations are applied to solve a symmetric top or toy gyroscope. This involves the introduction to Euler angles, angular velocity, angular momentum and kinetic energy for the system. In particular the precession of the spinning top is obtained and the general motion of the top obtained.

¹ed.copeland@nottingham.ac.uk

²antonio.padilla@nottingham.ac.uk

Useful resources

- H. Goldstein, C. Poole and J. Safko, *Classical mechanics*
- T. W. B. Kibble and F. H. Berkshire, *Classical mechanics*
- L. Hand and J. Finch, *Analytical mechanics*
- C. T. Whelan, Lecture Notes: <http://www.damtp.cam.ac.uk/user/studrep/ftp/podv0.1.pdf>
- D. Tong, Lecture Notes: <http://www.damtp.cam.ac.uk/user/tong/dynamics.htm>

Contents

1	Mathematical Preliminaries	5
1.1	Cartesian coordinates and index notation	5
1.2	Products of vectors	6
1.3	Cylindrical polar coordinates	8
1.4	Spherical polar coordinates	9
1.5	Position, velocity and acceleration	11
1.6	Vector calculus	12
2	A brief history of classical mechanics	14
2.1	Aristotle	14
2.2	Ptolemaeus	15
2.3	Kepler	16
2.4	Galileo	17
2.5	Newton	17
2.5.1	Newton's Laws of Motion	18
2.5.2	Newton's Laws of Universal Gravitation (NLG)	18
2.5.3	Is all mass the same?	18
2.6	After Newton	19
3	From Newton to Kepler	20
3.1	Derivation of Kepler's Laws	26
4	Lagrangians and Calculus of Variations	28
4.1	Calculus of Variations in one dimension	29
4.2	Calculus of Variations in more than one dimension	31
4.3	The physical Lagrangian	33
5	Generalised Coordinates and Hamilton's Principle	34
5.1	Many particle systems	35
6	The two body problem	38
7	Rigid bodies	40
7.1	Euler's theorem	41
7.2	Angular velocity about a fixed point	41
7.3	Angular momentum about a fixed point	42
7.4	The moment of inertia tensor	44
7.5	Kinetic energy of a rigid body about a fixed point	48
7.6	Gravitational potential energy of a rigid body	49
8	The compound pendulum	50

9	More on rigid bodies	52
9.1	Chasles' theorem	52
9.2	The importance of centre of mass	52
10	The double pendulum	53
10.1	Normal modes and normal frequencies	57
11	Constraints	58
11.1	Holonomic constraints	60
11.2	Non-holonomic constraints	62
11.3	Other tricks involving Lagrangians (<i>not examinable</i>)	62
12	Hamilton's principle –again	65
12.1	Constant Generalised Momentum	65
12.2	Constant Pseudo-energy – known as the Hamiltonian	67
12.3	Resonance	71
13	Charged particle in an electromagnetic field	75
14	Hamilton's Equations	79
14.1	Hamiltonian Function – $H(q, p)$	81
15	The symmetric top – or toy gyroscope	84
15.1	Euler angles	85
15.2	Angular velocity, angular momentum and kinetic energy	86
15.3	Precession of a symmetric top	88
15.4	General motion of the symmetric top	91
15.5	Describing the motion of the top	92
15.6	Stability of a vertical top	93

*"Find me in my field of grass – Mother Nature's son
Swaying daises sing a lazy song beneath the sun.
Mother Nature's son."*

The Beatles, *Mother Nature's son*

1 Mathematical Preliminaries

According to Wikipedia, Theoretical Physics is “a branch of physics which employs mathematical models and abstractions of physics in an attempt to explain natural phenomena.” Since this course discusses important concepts that form part of the bedrock of modern Theoretical Physics, it is inevitable that we will make use of plenty of mathematics. With this in mind we kick things off with a brief recap.

1.1 Cartesian coordinates and index notation

Most of what we need will involve the manipulation of vectors and vector operators. Recall that these can be expressed in terms of coordinate systems. For example, in a Cartesian coordinates system (x, y, z) , a 3 dimensional vector \underline{v} is usually written as

$$\underline{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = v_1 \hat{i} + v_2 \hat{j} + v_3 \hat{k} \quad (1.1)$$

where \hat{i} , \hat{j} and \hat{k} are the unit vectors along the x , y and z axes respectively, as shown in Fig. 1.1. The vector \underline{v} has components v_1 , v_2 and v_3 along these orthogonal directions. Note

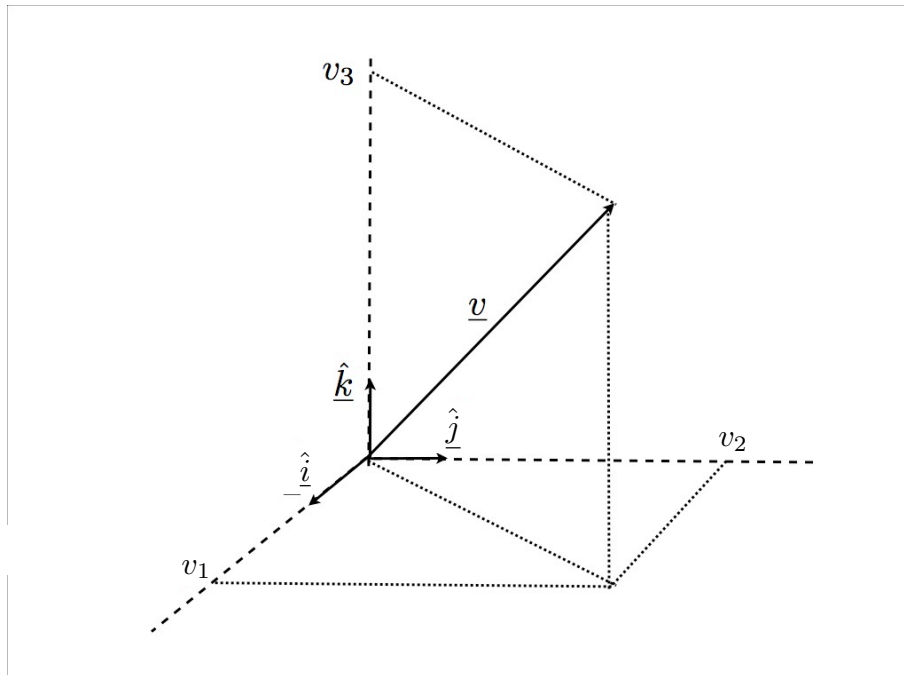


Figure 1.1: The Cartesian coordinate system

that \hat{i} , \hat{j} and \hat{k} are sometimes written as \hat{e}_1 , \hat{e}_2 and \hat{e}_3 , so that the vector can be written

succinctly as

$$\underline{v} = \sum_{i=1}^3 v_i \hat{e}_i \quad (1.2)$$

Similarly, we will sometimes prefer to refer to our coordinates as (x_1, x_2, x_3) as opposed to (x, y, z) . These adjustment have two advantages: (i) they make the generalisation to arbitrary dimension straight forward¹, and (ii) they bring us naturally on to *index notation*.

Index notation is not particularly deep, but it is computationally very powerful. Basically, we refer to the vector \underline{v} in terms of its components v_i along the Cartesian axes. We can also extend this to matrices as follows. We simply refer to the matrix

$$M = \begin{pmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{pmatrix} \quad (1.3)$$

in terms of its components M_{ij} .

Exercise: Verify that the matrix equation $\underline{u} = M\underline{v}$ can be written in index notation as $u_i = \sum_{j=1}^3 M_{ij}v_j$.

An example of such a matrix is of course the identity

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (1.4)$$

which has components given by the Kronecker delta symbol, δ_{ij} . We can also imagine objects labelled by more than 2 indices, such as the totally antisymmetric Levi Civita tensor, which in 3 dimensions has components, ϵ_{ijk} , where

$$\begin{aligned} \epsilon_{123} &= \epsilon_{231} = \epsilon_{312} = 1 \\ \epsilon_{213} &= \epsilon_{321} = \epsilon_{132} = -1 \\ \epsilon_{ijk} &= 0 \text{ otherwise} \end{aligned} \quad (1.5)$$

In general, as tensor of rank n will be labelled by n indices, $M_{i_1 \dots i_n}$, but we won't need to worry too much about that. We will only really consider vectors (rank 1 tensors), matrices (rank 2 tensors), and the antisymmetric Levi Civita tensor (which is an example of a rank 3 tensor in 3 dimensions).

1.2 Products of vectors

There are two types of product we will be interested in: scalar (aka dot) products, and vector (aka cross) products. The scalar product is given by

$$\underline{u} \cdot \underline{v} = |\underline{u}| |\underline{v}| \cos \alpha \quad (1.6)$$

¹In D dimensions, we have D axes, along which we have the D orthogonal unit vectors $\hat{e}_1, \dots, \hat{e}_D$. Then the D dimensional vector \underline{v} can be written as $\underline{v} = \sum_{i=1}^D v_i \hat{e}_i$.

where $|\underline{u}|$ and $|\underline{v}|$ are the length (or norm) of the two vectors, and α is the angle between them. Two things follow immediately from this: (i) we can calculate the norm of a vector using $|\underline{u}| = \sqrt{\underline{u} \cdot \underline{u}}$ and (ii) if \underline{u} and \underline{v} are orthogonal then $\underline{u} \cdot \underline{v} = 0$.

Exercise: Show that $\hat{e}_i \cdot \hat{e}_j = \delta_{ij}$

Using this we can derive the following familiar result

$$\underline{u} \cdot \underline{v} = \left(\sum_{i=1}^3 u_i \hat{e}_i \right) \cdot \left(\sum_{j=1}^3 v_j \hat{e}_j \right) = \sum_{i=1}^3 \sum_{j=1}^3 u_i v_j \hat{e}_i \cdot \hat{e}_j = \sum_{i=1}^3 \sum_{j=1}^3 u_i v_j \delta_{ij} = \sum_{i=1}^3 u_i v_i \quad (1.7)$$

These results for the scalar product can be generalised to arbitrary dimension. In contrast, the vector product is only defined in three dimensions². It is given by

$$\underline{u} \times \underline{v} = (|\underline{u}||\underline{v}| \sin \alpha) \hat{n} \quad (1.8)$$

where we have introduced the unit vector \hat{n} orthogonal to both \underline{u} and \underline{v} . The direction of $\underline{u} \times \underline{v}$, and, by association, the direction of \hat{n} is given by the well known *right-hand rule* (see Fig. 1.2). Note that if \underline{u} and \underline{v} are parallel then $\underline{u} \times \underline{v} = 0$.

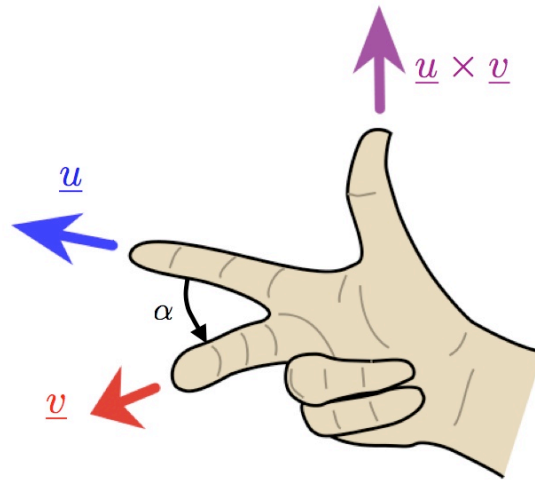


Figure 1.2: The right hand rule for cross products

²Higher dimensional generalisations do exist but are way beyond the scope of this course.

Exercise: Using the fact that $\hat{e}_i \times \hat{e}_j = \sum_{k=1}^3 \epsilon_{ijk} \hat{e}_k$, show that $\underline{w} = \underline{u} \times \underline{v}$ has components $w_i = \sum_{j,k} \epsilon_{ijk} u_j v_k$. Hence verify the familiar result

$$\underline{w} = \begin{pmatrix} u_2 v_3 - u_3 v_2 \\ u_3 v_1 - u_1 v_3 \\ u_1 v_2 - u_2 v_1 \end{pmatrix}$$

Having considered the basic products of vectors, we now consider products of products.

Exercise: Explain why $\underline{u} \cdot (\underline{u} \times \underline{v}) = 0$ for any 3 dimensional vectors \underline{u} and \underline{v} .

In addition to this rather trivial result, we have a number of less trivial, but extremely useful identities, including

$$\underline{a} \cdot (\underline{b} \times \underline{c}) = \underline{c} \cdot (\underline{a} \times \underline{b}) = \underline{b} \cdot (\underline{c} \times \underline{a}) \quad (1.9)$$

$$\underline{a} \times (\underline{b} \times \underline{c}) = (\underline{a} \cdot \underline{c})\underline{b} - (\underline{a} \cdot \underline{b})\underline{c} \quad (1.10)$$

$$(\underline{a} \times \underline{b}) \cdot (\underline{c} \times \underline{d}) = (\underline{a} \cdot \underline{c})(\underline{b} \cdot \underline{d}) - (\underline{a} \cdot \underline{d})(\underline{c} \cdot \underline{b}) \quad (1.11)$$

Exercise: Prove each of the identities (1.9) to (1.11). *Hint:* in some cases you will need to make use of the identity $\sum_{k=1}^3 \epsilon_{kij} \epsilon_{klm} = \delta_{il} \delta_{jm} - \delta_{im} \delta_{jl}$.

Whilst you will have proven these identities using a Cartesian coordinate system, they are in fact true in any coordinate system. That is one of the great things about the index notation: it enables you to derive identities fairly easily using Cartesian coordinates, even though you may want to apply them in more complicated coordinate systems. The mathematical reason for this comes from the fact that we are dealing with tensors in 3 dimensional Euclidean space, but you probably shouldn't be worrying about that here!

1.3 Cylindrical polar coordinates

Often the symmetries of a dynamical system make it convenient to work in an alternative coordinate system. An example of this is cylindrical polar coordinates (ρ, θ, z) . These are related to the usual Cartesian coordinates (x, y, z) as follows

$$x = \rho \cos \theta, \quad y = \rho \sin \theta, \quad z = z \quad (1.12)$$

Note that the limits of ρ , θ , and z are

$$0 \leq \rho < \infty, \quad 0 \leq \theta \leq 2\pi, \quad -\infty < z < \infty \quad (1.13)$$

In this new coordinate system, vectors are written in terms of their components along the r , θ and z directions. In other words

$$\underline{v} = v_\rho \hat{e}_\rho + v_\theta \hat{e}_\theta + v_z \hat{e}_z \quad (1.14)$$

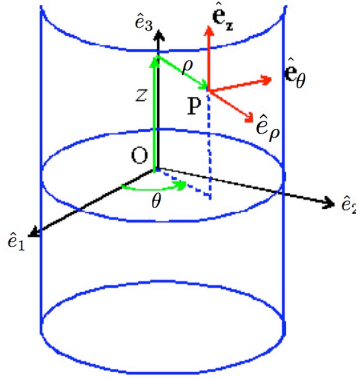


Figure 1.3: Cylindrical polar coordinates

What we have done here is replace the set of unit vectors in Cartesian coordinates, $\{\hat{e}_1, \hat{e}_2, \hat{e}_3\}$, with the new set of unit vectors in cylindrical polar coordinates $\{\hat{e}_\rho, \hat{e}_\theta, \hat{e}_z\}$. These are often referred to as the basis vectors for the relevant coordinate system. The basis vectors for cylindrical polars are shown in Fig. 1.3. We say that \hat{e}_ρ is the unit vector along the ρ direction, \hat{e}_θ is the unit vector along the θ direction and \hat{e}_z is the unit vector along the z direction. It is important to note that although each of $\{\hat{e}_\rho, \hat{e}_\theta, \hat{e}_z\}$ have fixed length, they do not have fixed direction, in contrast to $\{\hat{e}_1, \hat{e}_2, \hat{e}_3\}$. In fact, we have that $\frac{\partial}{\partial \theta} \hat{e}_\rho \neq 0$, $\frac{\partial}{\partial \theta} \hat{e}_\theta \neq 0$. We can see this after a quick glance at Fig. 1.3 from which we infer the following

$$\hat{e}_\rho = \hat{e}_1 \cos \theta + \hat{e}_2 \sin \theta, \quad \hat{e}_\theta = -\hat{e}_1 \sin \theta + \hat{e}_2 \cos \theta, \quad \hat{e}_z = \hat{e}_3 \quad (1.15)$$

Finally, we note that just as $\{\hat{e}_1, \hat{e}_2, \hat{e}_3\}$ are sometimes referred to as $\{\hat{i}, \hat{j}, \hat{k}\}$, we will sometimes refer to $\{\hat{e}_\rho, \hat{e}_\theta, \hat{e}_z\}$ as $\{\hat{\rho}, \hat{\theta}, \hat{z}\}$.

1.4 Spherical polar coordinates

Another well known and useful coordinate system is spherical polars (r, θ, ϕ) . These are related to the Cartesian coordinates (x, y, z) as follows

$$x = r \sin \phi \cos \theta, \quad y = r \sin \phi \sin \theta, \quad z = r \cos \phi \quad (1.16)$$

The limits of r , θ and ϕ are

$$0 \leq r < \infty, \quad 0 \leq \theta \leq 2\pi, \quad 0 \leq \phi < \pi \quad (1.17)$$

In this coordinate system, vectors are written in terms of their components along the r , θ and ϕ directions. In other words

$$\underline{v} = v_r \hat{e}_r + v_\theta \hat{e}_\theta + v_\phi \hat{e}_\phi \quad (1.18)$$

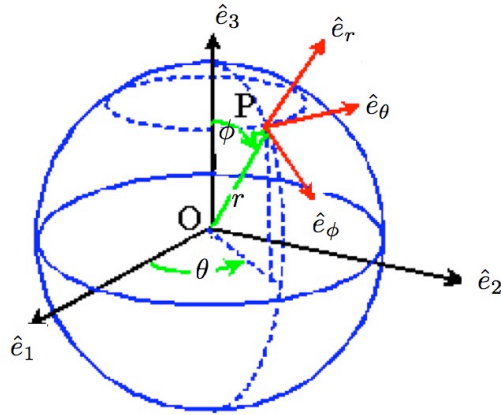


Figure 1.4: Spherical polar coordinates

where the new basis vectors $\{\hat{e}_r, \hat{e}_\theta, \hat{e}_\phi\}$ are shown in Fig. 1.4. We say that \hat{e}_r is the unit vector along the radial direction, \hat{e}_θ is the unit vector along the θ direction and \hat{e}_ϕ is the unit vector along the ϕ direction. Again, although each of the $\{\hat{e}_r, \hat{e}_\theta, \hat{e}_\phi\}$ has fixed length, they do not have fixed direction. Indeed, one should convince oneself that

$$\begin{aligned}
 \hat{e}_r &= \hat{e}_1 \sin \phi \cos \theta + \hat{e}_2 \sin \phi \sin \theta + \hat{e}_3 \cos \phi \\
 \hat{e}_\theta &= -\hat{e}_1 \sin \theta + \hat{e}_2 \cos \theta \\
 \hat{e}_\phi &= \hat{e}_1 \cos \phi \cos \theta + \hat{e}_2 \cos \phi \sin \theta - \hat{e}_3 \sin \phi
 \end{aligned} \tag{1.19}$$

Note that we will sometimes refer to $\{\hat{e}_r, \hat{e}_\theta, \hat{e}_\phi\}$ as $\{\hat{r}, \hat{\theta}, \hat{\phi}\}$.

In certain very special cases, spherical polars and cylindrical polars are interchangeable. This occurs when the motion is confined to the xy plane, or equivalently the equatorial plane where $\phi = \pi/2$. When this happens (and only when this happens!), \hat{r} defined for spherical polars is indistinguishable from $\hat{\rho}$ defined for cylindrical polars, whereas the vectors \hat{e}_z and \hat{e}_ϕ become obsolete.

Finally, we note that in more than 3 dimensions one can generalise spherical polars by introducing new angular directions. An interested student may like to refer to the following page on hyperspherical coordinates on PlanetMath.org:

<http://planetmath.org/encyclopedia/HypersphericalCoordinates.html>

1.5 Position, velocity and acceleration

In nearly everything we do in this course, we will be required to define some notion of position, $\underline{r}(t)$, in space³. In Cartesian coordinates, the position vector is given by

$$\underline{r}(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} = \sum_{i=1}^3 x_i(t) \hat{e}_i \quad (1.20)$$

It is important that we appreciate the distinction between the position $\underline{r}(t)$, which is a vector, and the distance from the origin, $r(t)$, which is a scalar. Indeed, the distance from the origin is given by

$$r(t) = |\underline{r}(t)| = \sqrt{x_1(t)^2 + x_2(t)^2 + x_3(t)^2} \quad (1.21)$$

Throughout this course it will be easy to distinguish between scalars and vectors — vectors are underlined, scalars are not!

Now velocity is defined as the rate of change of position. Velocity is a vector, and is given by

$$\underline{v}(t) = \frac{d}{dt} \underline{r}(t) = \dot{\underline{r}} = \begin{pmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \end{pmatrix} = \sum_{i=1}^3 \dot{x}_i(t) \hat{e}_i \quad (1.22)$$

Here we have introduced the “dot” notation, corresponding to differentiation with respect to time. Many of you will already be familiar with this – for those that aren’t, be rest assured that it will make everything a lot neater! Note also that we have used the fact that the basis vectors, \hat{e}_i , are fixed and do not change with time.

The speed is given by the magnitude of the velocity, $v = |\dot{\underline{r}}|$, and it is very important to realise that in general $|\dot{\underline{r}}| \neq \dot{r}$. This is easy enough to see just by evaluating both terms explicitly. From equation (1.22), we have

$$|\dot{\underline{r}}| = \sqrt{\dot{x}_1(t)^2 + \dot{x}_2(t)^2 + \dot{x}_3(t)^2} \quad (1.23)$$

whereas by differentiating equation (1.21) we have

$$\dot{r} = \frac{x_1 \dot{x}_1 + x_2 \dot{x}_2 + x_3 \dot{x}_3}{\sqrt{x_1(t)^2 + x_2(t)^2 + x_3(t)^2}} \quad (1.24)$$

Intuitively, it is easy to understand this result by considering motion in a circle around the origin. The distance from the origin is fixed and so $\dot{r} = 0$, but that does not mean we have vanishing speed. In other words, for motion in a circle, $|\dot{\underline{r}}| \neq 0 = \dot{r}$.

Finally, we introduce acceleration, defined as the rate of change of velocity. This is also a vector and is given by

$$\underline{a}(t) = \frac{d}{dt} \underline{v}(t) = \frac{d^2}{dt^2} \underline{r}(t) = \ddot{\underline{r}} = \begin{pmatrix} \ddot{x}_1(t) \\ \ddot{x}_2(t) \\ \ddot{x}_3(t) \end{pmatrix} = \sum_{i=1}^3 \ddot{x}_i(t) \hat{e}_i \quad (1.25)$$

³Space won’t always mean real 3 dimensional space in this course – it might refer to some abstract configuration space of any dimensionality, but let’s not worry about that just yet!

Note the “double dot” notation, corresponding to twice differentiating with respect to time.

Exercise: Is the magnitude of acceleration $|\ddot{\underline{r}}|$ the same as \ddot{r} ?

Of course, it may be convenient to use an alternative coordinate system to describe our position, etc. For example, in spherical polar coordinates, we have position

$$\underline{r}(t) = r(t)\hat{\underline{r}} \quad (1.26)$$

Working with this directly, we find that the velocity

$$\dot{\underline{r}} = \frac{d}{dt}(r(t)\hat{\underline{r}}) = \frac{dr}{dt}\hat{\underline{r}} + r\frac{d}{dt}\hat{\underline{r}} = \dot{r}\hat{\underline{r}} + r\frac{d}{dt}\hat{\underline{r}} \quad (1.27)$$

Now, the unit vector along the radial direction is not fixed in time, $\frac{d}{dt}\hat{\underline{r}} \neq 0$. Thus equation (1.27) clearly demonstrates that in general the velocity has a component along the $\frac{d}{dt}\hat{\underline{r}}$ direction, and so, as we saw earlier, the speed $|\dot{\underline{r}}| \neq \dot{r}$.

Exercise: Show that $\dot{r} = \hat{\underline{r}} \cdot \dot{\underline{r}}$. *Hint:* use the fact that $\hat{\underline{r}} \cdot \hat{\underline{r}} = 1$ to prove that $\hat{\underline{r}} \cdot \frac{d}{dt}\hat{\underline{r}} = 0$. Note that the product rule works in the usual way with scalar products.

One can also compute the acceleration in this coordinate system. We will not go into detail just now, simply reminding the reader that the basis vectors $\{\hat{\underline{r}}, \hat{\underline{\theta}}, \hat{\underline{\phi}}\}$ are not necessarily fixed in time, in contrast to the Cartesian basis vectors.

1.6 Vector calculus

In most dynamical systems, we are interested in functions of many variables, each variable corresponding to a coordinate in the system. An example of this might be the gravitational potential which depends on one’s location relative to the massive objects that influence the gravitational field. We will often need to know how such quantities change with a change in position. Indeed, the change of gravitational potential with respect to position is precisely what we mean by the gravitational force! In any event, we need to know how to take derivatives with respect to many variables, and this brings us on to vector calculus.

Let us work in a Cartesian coordinate system for the time being. The most important object in vector calculus is the gradient operator, $\underline{\nabla}$. In index notation this has components $\frac{\partial}{\partial x_i}$. For scalar valued functions of space, $f(x_1, x_2, x_3)$, we define the *gradient* (or *grad*) of f as

$$\underline{\nabla}f = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \frac{\partial f}{\partial x_3} \end{pmatrix} \quad (1.28)$$

Here the operator maps the scalar valued function to a vector. However, one can also perform vector calculus on vector valued functions, $\underline{F}(x_1, x_2, x_3)$. The *divergence* (or *div*) of

\underline{F} is defined as

$$\underline{\nabla} \cdot \underline{F} = \frac{\partial F_1}{\partial x_1} + \frac{\partial F_2}{\partial x_2} + \frac{\partial F_3}{\partial x_3} = \sum_{i=1}^3 \frac{\partial}{\partial x_i} F_i \quad (1.29)$$

and maps a vector to a scalar. One can think of the divergence as being like the dot product of the gradient operator with vector valued function.

Finally, the *curl* of \underline{F} is defined as

$$\underline{\nabla} \times \underline{F} = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \frac{\partial}{\partial x_3} \end{pmatrix} \times \begin{pmatrix} F_1 \\ F_2 \\ F_3 \end{pmatrix} = \sum_{i=1}^3 \hat{e}_i (\epsilon_{ijk} \frac{\partial}{\partial x_j} F_k) \quad (1.30)$$

and maps a vector to a vector. One can think of the curl as being like the cross product of the gradient operator with vector valued function.

Exercise: Show that for any scalar function f , we have $\underline{\nabla} \times \underline{\nabla} f = 0$.

Of course, we can generalise vector calculus to different coordinate systems, although we have to take some care since the gradient operator can act non-trivially on the new basis vectors. For example, in spherical polars $\underline{\nabla} \cdot \hat{e}_r = 2/r$. This makes the precise form of grad, div and curl quite complicated in principle. Often it is easiest to use index notation throughout your working, and convert the answer to the relevant coordinate system at the end of calculation.

2 A brief history of classical mechanics

Our current understanding of the physical laws that govern Nature is built upon over 2000 years of observation and imagination. Let us embark on a journey through time to see how this important branch of Physics developed throughout the ages. We will start in the philosophical melting pot that was ancient Greece.

2.1 Aristotle

In around 330 BC Aristotle published his book “Physica” (meaning “natural”), which described the general principles of motion. He introduced 5 elements: Earth, Water, Air, Fire and the mysterious Aether, each having their “natural” place, as shown in Fig. 2.1. He

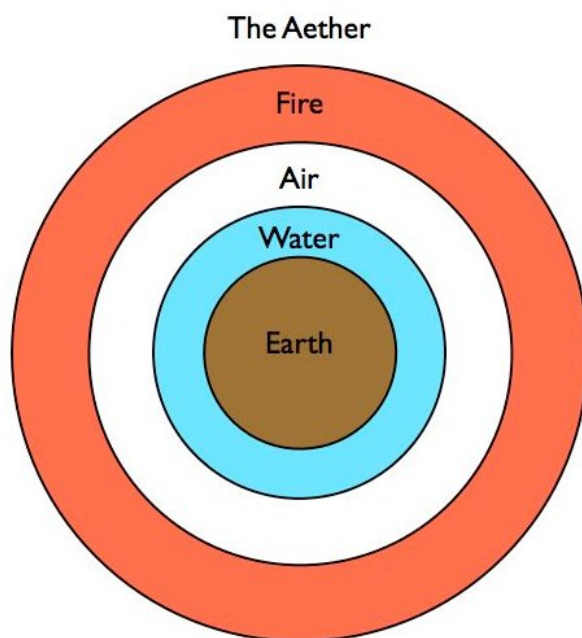


Figure 2.1: Aristotle’s elements in their natural place.

claimed that objects will flow to their natural place, depending on what they are made of. For example

- earth falls through water
- fire rises through air
- rain falls through air

- air bubbles rise in water

For what it worth, the "aether" was understood to be the divine substance making up the heavenly bodies. According to Aristotle, these bodies moved in perpetual circular motion.

2.2 Ptolemaeus

In around 150 AD Ptolemaeus proposed a sophisticated mathematical model describing the motion of celestial bodies. A modest chap, he published his work in a book which he entitled "Almagest", meaning "Great Book". The Ptolemaic universe involved three important concepts: the deferent, the epicycle and the equant. The deferent is a large circle

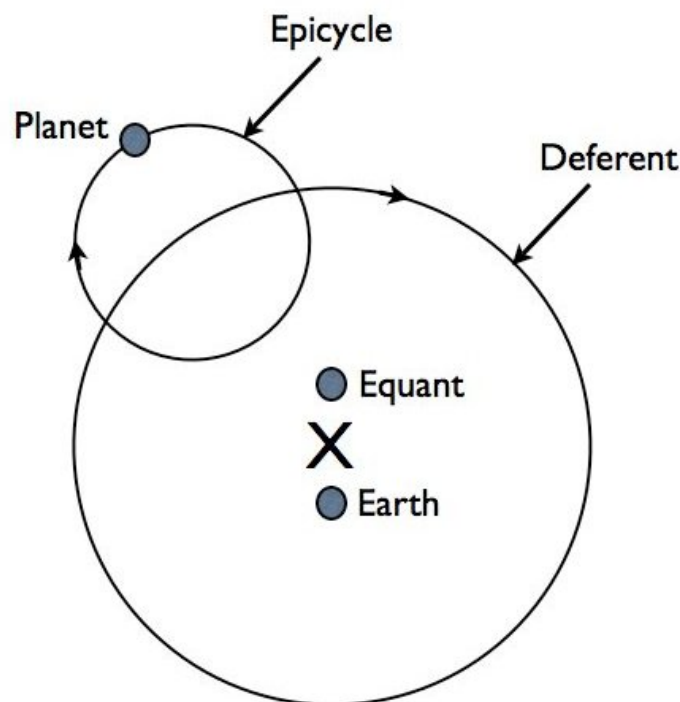


Figure 2.2: The Ptolemaic Universe

centred at some point X, lying midway between the earth and the equant. The epicycle is a smaller circle whose centre moves along the deferent, with constant speed *relative to the equant*. The celestial body moves around the epicycle. In some ways the equant is an ugly addition to the model, but was necessary to account for the anomalous motion of the planets. In analogy with many models of dark matter and dark energy popular today, it was introduced solely to "fit the data", without a sound theoretical understanding. In Physics

this is often a sign that your theory, if you can call it that, isn't quite right, and you should go back and think again.

Nevertheless, Ptolemaeus' was certainly entitled to claim his book was "great", because this remained the preferred model of celestial mechanics until the renaissance. Ironically, given the nature of Astronomy, things didn't really look up for another 1500 years. With this in mind, let us fast forward to Shakespearean time, although we will be more interested in the work of a German astronomer than that of an English playwright.

2.3 Kepler

In 1605 Kepler challenged the Ptolemaic view in the light of some crucial observations made by Tycho Brahe. In last year's "Newton to Einstein" course you learnt about *Keplers Laws of planetary motion*, which is slightly ironic given that Kepler predated Newton! Not to worry, Keplers laws state:

K1L: the orbit of every planet is an ellipse, with the Sun at a focus.

K2L: the line joining a planet and the Sun sweeps out the same area in a given time.

K3L: if T is the orbital period, and a is the semi-major axis of the orbit, then $T^2 \propto a^3$.

To get a feel for Kepler's second and third laws, check out Figs 2.3 and 2.4 respectively.

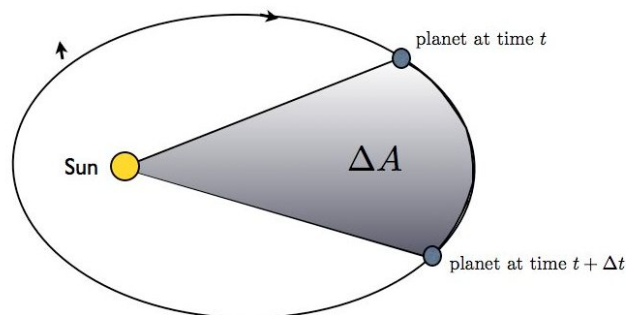


Figure 2.3: Kepler's 2nd Law: the area swept out over an interval Δt is $\Delta A = (\text{constant})\Delta t$.

Interestingly, Kepler also thought about why there are three dimensions of space. He concluded that it reflected the 3-fold nature of the Holy Trinity. We still don't really have any idea why there are three large spatial dimensions in the Universe. Actually, the subject of religion brings us nicely onto our next big player: someone whose work led to him being excommunicated from the Church and imprisoned as a heretic by the Spanish Inquisition (which he probably never expected!).

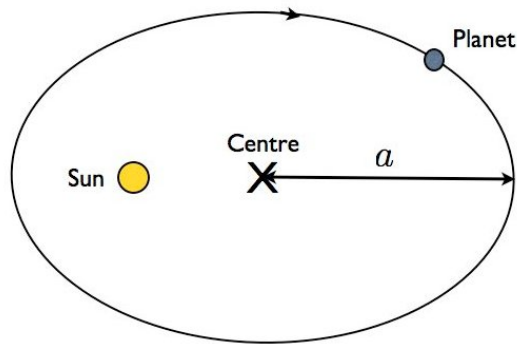


Figure 2.4: Kepler's 3rd Law: $(\text{period of orbit})^2 = T^2 \propto a^3$.

2.4 Galileo

In or around 1638 Galileo is said to have performed a famous experiment whereby he dropped a cannon ball and a lighter wooden ball from the leaning tower of Pisa, demonstrating that both objects fall at the same rate. This was a radical departure from the old Aristotlean view. Actually there is some doubt that Galileo ever performed this experiment, but others certainly did, including his student in an attempt to convince the doubting Aristotlean scholars of the day. A similar experiment was also performed by astronauts on the moon, with the same results.

In any case, in his "Discorsi", Galileo postulated the following

- A moving body falls with uniform acceleration, independently of its mass or composition, so long as the resistance through the medium is negligible (an early version of Einstein's *Principle of Equivalence*).
- A body moving on a level surface will continue in the same direction at constant speed, unless disturbed (the *Principle of Inertia*).

When a certain Physicist, who we will visit shortly, said "If I can see further than anyone else, it is only because I am standing on the shoulders of giants", he was certainly talking, in part, about Galileo. However, the man responsible for that famous quote was arguably the greatest of them all. If Elvis was the King who revolutionised music with rock 'n' roll, then Newton was the King who revolutionised Physics with his *Principia*.

2.5 Newton

In 1687 Newton published what many believe to be the greatest scientific publication of all time: *Principia Mathematica*. In it he laid down his 3 laws of motion and his law of Universal Gravitation. Whereas Kepler's Laws were a useful mathematical model, Newton

proposed a *theory* from which Kepler's Laws may be derived (as we shall see later). Let us remind ourselves of those laws:

2.5.1 Newton's Laws of Motion

N1L: A body at rest remains at rest, and a body in linear motion remains in motion with constant velocity unless an external force is applied on it. (Galileo)

N2L: The force applied on a body is equal to the rate of change of its momentum

$$\underline{F} = \frac{d\underline{p}}{dt}$$

(For a body of constant inertial mass m and velocity \underline{v} , the momentum $\underline{p} = m\underline{v} \Rightarrow \underline{F} = m\frac{d\underline{v}}{dt} = m\underline{a}$, where \underline{a} is the acceleration. However, mass isn't always constant, in which case one should use $\underline{F} = \frac{d\underline{p}}{dt}$, eg. a rocket burning up its fuel.)

N3L: If a body A exerts a force \underline{F} on body B, then body B exerts an equal and opposite force $-\underline{F}$ on body A.

2.5.2 Newton's Laws of Universal Gravitation (NLG)

Every point mass attracts every other point mass by a force along the line joining the two points. The size of the force is proportional to the product of the two masses, and inversely proportional to the square of their separation. That is

$$F = G_N \frac{m_A m_B}{r^2}$$

for two masses m_A and m_B separated by a distance r . The constant of proportionality is given by G_N (Newton's constant). Check out Fig. 2.5.

2.5.3 Is all mass the same?

Up until now we have been referring to the mass of a particle without worrying too much about the role that mass is playing. When we do start to worry we realise that there are, in principle at least, 3 different types of mass. We call them *inertial* mass, *passive gravitational* mass, and *active gravitational* mass.

Inertial mass: measures the resistance to change in motion.

Passive gravitational mass: measures the reaction to a gravitational field.

Active gravitational mass: measures the ability to source a gravitational field.

The inertial mass is essentially the mass that appears in N2L (the m in $\underline{F} = m\underline{a}$). To get a feel for the gravitational masses, refer again to Fig. 2.5, only now say that body A has active mass m_A and passive mass \bar{m}_A . We similarly introduce m_B, \bar{m}_B . The force of body A on body B is should now be given by

$$F_{AB} = G_N \frac{m_A \bar{m}_B}{r^2}$$

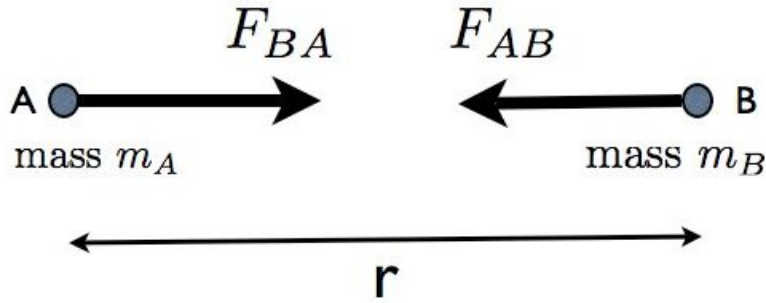


Figure 2.5: Newton’s Law of Gravitation: in the diagram the forces have magnitude $F_{AB} = F_{BA} = G_N \frac{m_A m_B}{r^2}$, where G_N is Newton’s constant.

since A is active, and B is passive. Likewise

$$F_{BA} = G_N \frac{\bar{m}_A m_B}{r^2}$$

since now B is active and A is passive.

However, by N3L, these forces should be equal in magnitude, which is consistent with the assumption that active and passive gravitational masses are the same ($m_A = \bar{m}_A$). Indeed, from now on we will assume that *all three types of mass are the same*, in accordance with the principle of equivalence. Nevertheless, it is worth being aware of the fact that this is an assumption for which we have some experimental evidence, but no proof.

2.6 After Newton

Not everyone was entirely happy with Newton’s theory, including Newton himself. He was particularly upset about the “action at a distance”—the idea that two very distance bodies could “act” upon each other instantaneously. Indeed, in 1692 he wrote:

“That one body may act upon another at a distance through a vacuum without the mediation of anything else, by and through which their action and force may be conveyed from one another, is to me so great an absurdity that, I believe, no man who has in philosophic matters a competent faculty of thinking could ever fall into it.”

It would take Einstein, over 300 years later, to resolve this conundrum.

Others, meanwhile, were keen to understand the underlying reasons behind Newton’s Laws. Leibnitz, Newton’s arch rival, felt that gravity hadn’t really been explained by “simply” postulating the existence of a gravitational force. Where had it come from?

Maupertuis came closest to the truth. In around 1750 he suggested that the Laws of Nature were such as to expend the least possible “action”—an economy of effort on the part of the Creator. He, and Euler, argued that a particle travelling between points \underline{x}_1 and \underline{x}_2

does so along a path, $\underline{x}(s)$, that minimizes the “action”,

$$S = \int_{s_1}^{s_2} mvd s, \quad (2.1)$$

with boundary conditions $\underline{x}(s_1) = \underline{x}_1$ and $\underline{x}(s_2) = \underline{x}_2$. Here m is mass, v is speed, and s measures the distance along the path. This economy of effort by Nature, was, according to Maupertuis, proof of God’s existence.

Maupertuis’ action requires the additional assumption of energy conservation to recover Newton’s Laws upon minimization. *Hamilton’s principle* (1834) goes one better and does not require the assumption of energy conservation. Hamilton replaced Maupertuis’ action with “Hamilton’s principal function”. As we will see in the rest of this course, minimizing this function yields Newton’s Laws in all their glory. The act of “minimizing” uses techniques developed by Euler and Lagrange known as *Calculus of Variations*. Hamilton and Lagrange’s work are the key ingredients in our modern approach to classical dynamics. They are also vital to understanding and developing aspects of quantum mechanics, electromagnetism, General Relativity, and even string theory.

Although the *principle of least action* historically refers to Maupertuis’ principle, in modern parlance we associate it with Hamilton’s principle, simply referring to his principal function as “the action”. Hamilton and Lagrange’s ideas are the inspiration for this course.

3 From Newton to Kepler

Newton’s Laws are truly incredible. We can use them to accurately predict the motions of particles in many scenarios. They can tell you the path of a ball kicked by a Stevie G freekick, a fact which seems lost on most premiership goalkeepers. On larger scales they can even reproduce the orbital motion of the planets. Consider this for a moment. With just a few simple laws we can predict the motions of planets around the Sun to a remarkable degree of accuracy.

Perhaps the best example of the power of Newtonian physics lies in the discovery of the planet Neptune. Neptune’s existence was actually *predicted* some 20 years before its discovery in 1846. This foresight was based on the orbit of the planet Uranus, which didn’t quite tally with the Newtonian prediction. Rather than lose faith in Newton, Bouvard suggested that there must be a very massive unseen body disrupting Uranus’ orbit. He was right—the massive body was Neptune.

Today the same reasoning has led us to predict the existence of dark matter. In this case the orbits of the outer stars in galaxies don’t matter the Newtonian prediction based on the amount of visible matter seen in the stars and gases that make up the galaxy. A word of warning however: Newtonian physics also failed to account for the perihelion precession (a kind of wobble) in the orbit of Mercury. Again, many kept faith in Newton and postulated the existence of a “dark planet” called Vulcan. Einstein didn’t. He invented General Relativity.

In this section we will show how Kepler’s Laws describing planetary motion can be derived from Newtonian theory. To this end, consider a heavy object (the Sun) of mass M fixed at

the origin. We will model a planet as a test particle of mass m . This approximation works well as long as the scale of the planet is far smaller than the scale of its orbit. Our aim is to derive the path of the planet $\underline{r} = \underline{r}(t)$ under the influence of the Sun's gravitational field, see Fig. 3.1.

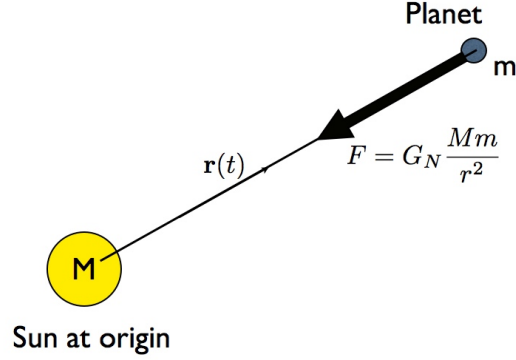


Figure 3.1: The Sun exerts a force on the planet of size $F = G_N \frac{Mm}{r^2}$, where $r = |\underline{r}(t)|$, directed towards the Sun at the origin.

The planet has velocity $\dot{\underline{r}}(t) = \frac{d\underline{r}(t)}{dt}$ and acceleration $\ddot{\underline{r}}(t) = \frac{d^2\underline{r}(t)}{dt^2}$. By NLG it experiences a force of size $F = G_N \frac{Mm}{r^2}$, directed towards the origin. Applying N2L, we have

$$m\ddot{\underline{r}} = G_N \frac{Mm}{r^2} (-\hat{\underline{r}}), \quad \hat{\underline{r}} = \frac{\underline{r}}{r} \quad (3.1)$$

$$\Rightarrow \ddot{\underline{r}} = -\frac{G_N M}{r^3} \underline{r} \quad (3.2)$$

Note that since the force acts towards the origin it has direction $-\hat{\underline{r}}$.

Conservation of energy

Dotting both sides of Eq. 3.2 with $\dot{\underline{r}}$, we find

$$\dot{\underline{r}} \cdot \ddot{\underline{r}} = -\frac{G_N M}{r^3} \underline{r} \cdot \dot{\underline{r}} \quad (3.3)$$

$$\Rightarrow \frac{1}{2} \frac{d}{dt} (\dot{\underline{r}} \cdot \dot{\underline{r}}) = -\frac{G_N M}{r^3} \frac{1}{2} \frac{d}{dt} (\underline{r} \cdot \underline{r}) = -\frac{G_N M}{r^3} \frac{1}{2} \frac{d}{dt} (r^2) = -\frac{G_N M}{r^2} \frac{dr}{dt} \quad (3.4)$$

$$\Rightarrow \frac{1}{2} |\dot{\underline{r}}|^2 - \frac{G_N M}{r} = E, \text{ constant} \quad (3.5)$$

Here $\frac{1}{2}|\dot{\underline{r}}|^2$ is the kinetic energy per unit mass, $-\frac{G_N M}{r}$ is the potential energy per unit mass, and E is the total energy per unit mass. Since E is a constant we see that this is just a statement of conservation of energy.

Conservation of Angular Momentum

We now cross both sides of Eq. 3.2 with \underline{r}

$$\underline{r} \times \ddot{\underline{r}} = -\frac{G_N M}{r^3} \underline{r} \times \underline{r} = 0 \quad (3.6)$$

$$\Rightarrow \frac{d}{dt}(\underline{r} \times \dot{\underline{r}}) = 0 \quad (3.7)$$

$$\Rightarrow \underline{r} \times \dot{\underline{r}} = \underline{h}, \text{ constant} \quad (3.8)$$

Recall that angular momentum is given by $\underline{r} \times m\dot{\underline{r}}$, so \underline{h} is just the angular momentum per unit mass. Since \underline{h} is constant this is just a statement of conservation of angular momentum.

It follows that the motion is always perpendicular to the constant vector \underline{h} , since

$$\underline{r} \cdot \underline{h} = \dot{\underline{r}} \cdot \underline{h} = 0 \quad (3.9)$$

Now, without loss of generality (WLOG), we can choose our axes such that \underline{h} points along the z -direction

$$\underline{h} = \begin{pmatrix} 0 \\ 0 \\ h \end{pmatrix} \quad (3.10)$$

This means the planet's motion lies in the xy plane. It is convenient to make use of polar coordinates so that

$$\underline{r}(t) = \begin{pmatrix} x(t) \\ y(t) \\ 0 \end{pmatrix} \quad (3.11)$$

where $x(t) = r \cos \theta$, $y(t) = r \sin \theta$. It follows that

$$\dot{\underline{r}}(t) = \dot{r}\hat{\underline{r}} + r\dot{\theta}\hat{\underline{\theta}}, \quad \hat{\underline{r}} = \begin{pmatrix} \cos \theta \\ \sin \theta \\ 0 \end{pmatrix}, \quad \hat{\underline{\theta}} = \begin{pmatrix} -\sin \theta \\ \cos \theta \\ 0 \end{pmatrix} \quad (3.12)$$

Here $\hat{\underline{r}}$ and $\hat{\underline{\theta}}$ are the unit vectors along the radial and angular directions respectively (see Fig. 3.2). Note that $\dot{r} = \frac{d|r|}{dt} \neq |\dot{\underline{r}}|$. Now, from conservation of angular momentum

$$\underline{h} = \underline{r} \times \dot{\underline{r}} = r\hat{\underline{r}} \times (\dot{r}\hat{\underline{r}} + r\dot{\theta}\hat{\underline{\theta}}) = r^2\dot{\theta}\hat{\underline{r}} \times \hat{\underline{\theta}} = \begin{pmatrix} 0 \\ 0 \\ r^2\dot{\theta} \end{pmatrix} \quad (3.13)$$

$$\Rightarrow h = r^2\dot{\theta} \quad (3.14)$$

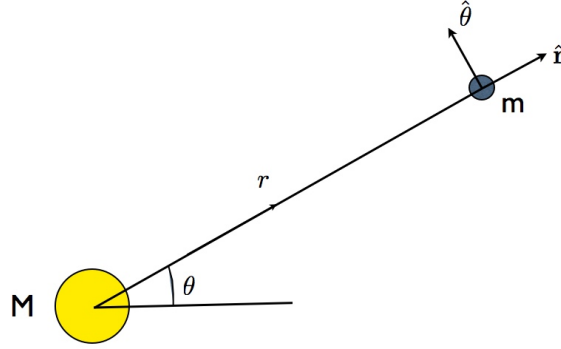


Figure 3.2: Polar coordinates in the plane of motion.

Consider now the energy of the particle. This requires us to calculate

$$|\dot{\underline{r}}|^2 = \dot{\underline{r}} \cdot \dot{\underline{r}} = (\dot{r}\hat{r} + r\dot{\theta}\hat{\theta}) \cdot (\dot{r}\hat{r} + r\dot{\theta}\hat{\theta}) \quad (3.15)$$

$$= \dot{r}^2 + (r\dot{\theta})^2 \quad \text{since } |\hat{r}| = |\hat{\theta}| = 1, \hat{r} \cdot \hat{\theta} = 0 \quad (3.16)$$

$$= \dot{\theta}^2 \left[\left(\frac{dr}{d\theta} \right)^2 + r^2 \right] \quad \text{since } \frac{dr}{d\theta} = \frac{\dot{r}}{\dot{\theta}}. \quad (3.17)$$

$$= \left(\frac{h}{r^2} \right)^2 \left[\left(\frac{dr}{d\theta} \right)^2 + r^2 \right] \quad \text{by Eq. 12.10} \quad (3.18)$$

Plugging this expression into our formula for energy conservation (Eq. 3.5) , we find

$$2E = \left(\frac{h}{r^2} \right)^2 \left[\left(\frac{dr}{d\theta} \right)^2 + r^2 \right] - \frac{2G_N M}{r} \quad (3.19)$$

This equation governs the motion of the planet. It is a bit difficult to solve in its current form so we perform a few tricks. First note that r typically appears with inverse powers. With this in mind we introduce $u = 1/r$, from which we obtain

$$\frac{dr}{d\theta} = -\frac{1}{u^2} \frac{du}{d\theta} \quad (3.20)$$

Substituting this into Eq. 3.19, we find

$$2E = (hu^2)^2 \left[\frac{1}{u^4} \left(\frac{du}{d\theta} \right)^2 + \frac{1}{u^2} \right] - 2G_N M u \quad (3.21)$$

$$\Rightarrow \left(\frac{du}{d\theta} \right)^2 = \frac{2E}{h^2} + \left(\frac{G_N M}{h^2} \right)^2 - \left(u - \frac{G_N M}{h^2} \right)^2 \quad (3.22)$$

We now let $u = \frac{G_N M}{h^2} + \alpha \cos \beta(\theta)$, where α is a constant to be chosen shortly. Since

$$\frac{du}{d\theta} = -\alpha \sin \beta \frac{d\beta}{d\theta} \quad (3.23)$$

Eq. 3.22 takes the form

$$\alpha^2 \sin^2 \beta \left(\frac{d\beta}{d\theta} \right)^2 = \frac{2E}{h^2} + \left(\frac{G_N M}{h^2} \right)^2 - \alpha^2 \cos^2 \beta \quad (3.24)$$

Choosing

$$\alpha = \sqrt{\frac{2E}{h^2} + \left(\frac{G_N M}{h^2} \right)^2} \quad (3.25)$$

ensures that the right hand side of Eq. 3.24 becomes $\alpha^2 \sin^2 \beta$, and so

$$\left(\frac{d\beta}{d\theta} \right)^2 = 1 \quad \Rightarrow \quad \beta = \pm(\theta - \theta_0) \quad \Rightarrow \quad u = \frac{G_N M}{h^2} + \alpha \cos(\theta - \theta_0) \quad (3.26)$$

where θ_0 is a constant. The planet's orbit can now be expressed in polar coordinates as

$$r = \frac{p}{1 + \epsilon \cos(\theta - \theta_0)} \quad (3.27)$$

where $p = \frac{h^2}{G_N M}$ and $\epsilon = \sqrt{1 + 2E \left(\frac{h}{G_N M} \right)^2}$.

Exercise: Derive Eq. 3.27 by applying N2L directly in the plane of motion.

Hint: Show that the planet's acceleration is given by

$$\ddot{\underline{r}} = (\ddot{r} - r\dot{\theta}^2)\hat{\underline{r}} + (r\ddot{\theta} + 2\dot{r}\dot{\theta})\hat{\underline{\theta}} \quad (3.28)$$

and apply N2L along both the radial and angular directions.

We can classify the orbits according to whether or not they are bounded – in other words, whether or not they can escape the gravitational field of the source. Mathematically, this corresponds to asking whether or not there exists some angle θ_∞ such that $r \rightarrow \infty$ and $\theta \rightarrow \theta_\infty$.

If $\epsilon < 1$, there exists no such θ_∞ , and the orbit is said to be *bound*. This corresponds to $E < 0$, and physically means that the planet has insufficient kinetic energy to escape from the potential well created by the gravitational field of the massive source (in this case, the Sun).

If $\epsilon \geq 1$, then $\theta_\infty = \theta_0 + \cos^{-1}(-1/\epsilon)$. Such orbits are said to be *unbound*, since $E \geq 0$, and the planet has enough kinetic energy to climb out of the potential well.

To picture the shapes of the orbits, we set $\theta_0 = 0$, WLOG, and recall that

$$x = r \cos \theta, \quad y = r \sin \theta \quad (3.29)$$

Now from Eq. 3.27, we have $p = r + \epsilon r \cos \theta = r + \epsilon x$, and so

$$x^2 + y^2 = r^2 = (p - \epsilon x)^2 \quad (3.30)$$

$$\Rightarrow x^2(1 - \epsilon^2) + 2p\epsilon x + y^2 = p^2 \quad (3.31)$$

$$\Rightarrow \frac{(x + c)^2}{a^2} + \frac{y^2}{b^2} = 1 \quad \text{for } \epsilon \neq 1 \quad (3.32)$$

where

$$a = \frac{p}{1 - \epsilon^2}, \quad b = \frac{p}{\sqrt{1 - \epsilon^2}}, \quad c = \frac{p\epsilon}{1 - \epsilon^2} \quad (3.33)$$

For $\epsilon < 1$ the orbit is elliptical (see Fig. 3.3). The ellipse has centre $(-c, 0)$, semi major axis a and semi minor axis $b < a$. The minimum radius of the orbit occurs when $\theta = 0$ so that $r_{min} = \frac{p}{1 + \epsilon}$. The maximum radius occurs when $\theta = \frac{\pi}{2}$ so that $r_{max} = \frac{p}{1 - \epsilon}$. The quantity ϵ is called the ‘‘eccentricity’’ of the ellipse. For $\epsilon > 1$ the orbit is hyperbolic (see Fig. 3.4).

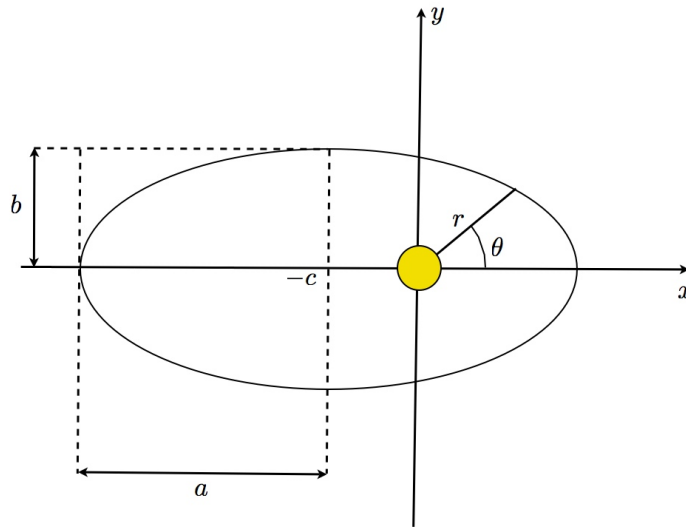


Figure 3.3: The bound elliptic orbit for $\epsilon < 1$.

For $\epsilon = 1$, we have $x = \frac{p}{2} - \frac{y^2}{2p}$, corresponding to a parabolic orbit (see Fig. 3.5). In this marginal case, the energy $E = 0$, so the particle has just enough kinetic energy to climb out of the potential well.

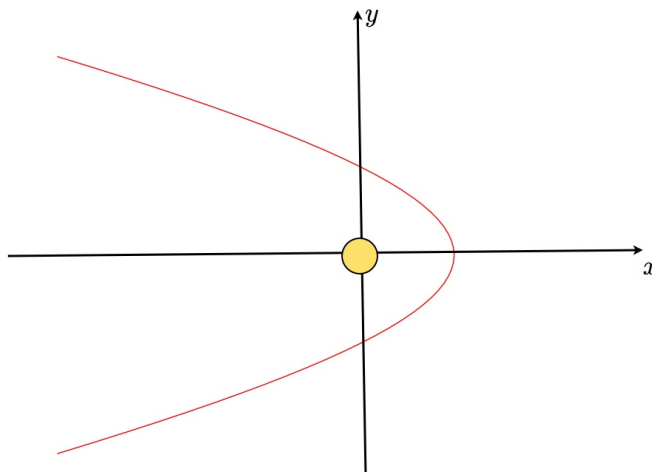


Figure 3.4: The unbound hyperbolic orbit for $\epsilon > 1$.

3.1 Derivation of Kepler's Laws

We are now in a position to derive Kepler's Laws. Let us look at them individually, beginning with Kepler's First Law.

K1L: the orbit of every planet is an ellipse, with the Sun at a focus.

Proof: the planetary orbits are clearly bound, and by Newton's Laws we have shown that the orbit is given by Eq. 3.27 with $\epsilon < 1$. This is the formula for an ellipse, written in polar coordinates, about a focus at the origin.

Exercise: An ellipse can be defined as the locus of all points, P , in the plane such that

$$PF_1 + PF_2 = \text{constant}$$

where F_1 and F_2 are two fixed points, and PF_1 is the distance from P to F_1 (similarly, PF_2). F_1 and F_2 are known as the foci. By placing F_1 at the origin, and F_2 at $x = -\frac{2p\epsilon}{1-\epsilon^2}$, show that the ellipse satisfies Eq. 3.27 when we choose the "constant" = $\frac{2p}{1-\epsilon^2}$.

Hint: Draw a picture and use the cos rule.

We now move on to Kepler's Second Law.

K2L: the line joining a planet and the Sun sweeps out the same area in a given time.

Proof: Consider the area δA swept out in an infinitesimal time δt . It is given by

$$\delta A \approx \frac{1}{2} r^2 \delta \theta$$

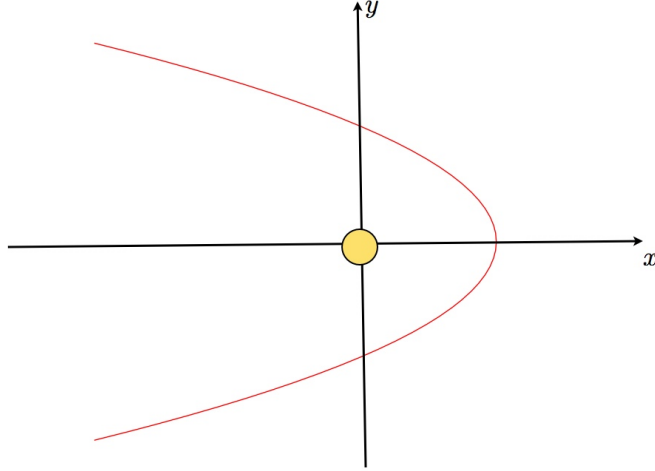


Figure 3.5: The unbound parabolic orbit for $\epsilon = 1$.

where $\delta\theta \approx \dot{\theta}\delta t$. It follows that

$$\frac{\delta A}{\delta t} \approx \frac{1}{2}r^2\dot{\theta}$$

Taking $\delta t \rightarrow 0$, we find that $\frac{dA}{dt} = \frac{1}{2}r^2\dot{\theta} = \frac{h}{2}$, which is constant by conservation of angular momentum.

Finally, we prove Kepler's Third Law.

K3L: if T is the orbital period, and a is the semi-major axis of the orbit, then $T^2 \propto a^3$.

Proof: Given that $\frac{dA}{dt} = \frac{h}{2}$, it is easy to see that the total area of the ellipse is given by

$$\mathcal{A} = \int_0^T \frac{dA}{dt} dt = \frac{1}{2}hT$$

However, it is well known that the area of the ellipse is $\mathcal{A} = \pi ab$. From Eq. 3.33, we see that $b = a\sqrt{1 - \epsilon^2}$, and so bringing everything together we find

$$T = \frac{2\pi}{h}a^2\sqrt{1 - \epsilon^2} \quad \Rightarrow \quad \frac{T^2}{a^3} = (2\pi)^2 \left[\frac{a(1 - \epsilon^2)}{h^2} \right]$$

Recall that $a = \frac{p}{1 - \epsilon^2}$ where $p = \frac{h^2}{G_N M}$. It follows that

$$\frac{a(1 - \epsilon^2)}{h^2} = \frac{1}{G_N M} \quad \Rightarrow \quad \frac{T^2}{a^3} = \frac{(2\pi)^2}{G_N M} = \text{constant}$$

The power of Newton's Laws is clearly evident, enabling us to derive Kepler's Laws from a set of fundamental principles. But we can go one better. We can derive Newton's Laws themselves from a single principle – Hamilton's principle. This formulation will enable us to easily expand Newtonian theory to accommodate many particle systems, rigid bodies of finite size and constrained dynamical systems.

In the following sections we will look deeply into Newton's Laws and understand them in terms of Hamilton's principle of least action – Nature's economy of effort, alluded to earlier. We will be taking on plenty of new ideas, but this won't be a problem, since Nottingham's students are not the sort to follow Nature's dubious example!

4 Lagrangians and Calculus of Variations

Recall Maupertuis' dream of "least action" – Nature's economy of effort. How do we describe the "action" mathematically?

Consider the following example. We have a particle of mass m moving under the influence of some potential. The particle starts at position \underline{x}_0 at some initial time t_0 , and ends up at position \underline{x}_1 at some final time t_1 . We would like to know how the particle moves from \underline{x}_0 to \underline{x}_1 . In other words, we want to know the position $\underline{x}(t)$ describing the particle path for $t_0 \leq t \leq t_1$.

Let us forget Newton's Laws for the moment, and focus on clarifying what we mean by the "action". Consider the motion of the particle over an infinitesimal time δt between times t and $t + \delta t$. Clearly the corresponding "action", $\delta S \propto \delta t$, where the "constant" of proportionality is called the *Lagrangian*. The Lagrangian usually has the symbol, \mathcal{L} , so $\delta S = \mathcal{L}\delta t$

What can \mathcal{L} depend on? Well, the only things available to us are time t , the position $\underline{x}(t)$, and its derivatives, $\dot{\underline{x}}(t), \ddot{\underline{x}}(t) \dots$ That is,

$$\mathcal{L} = \mathcal{L}(t, \underline{x}(t), \dot{\underline{x}}(t), \ddot{\underline{x}}(t), \dots) \quad (4.1)$$

In a large number of cases, the Lagrangian will only depend on time, t , position $\underline{x}(t)$ and velocity $\dot{\underline{x}}(t)$, so for the most part let us restrict attention to the case $\mathcal{L} = \mathcal{L}(t, \underline{x}(t), \dot{\underline{x}}(t))$.

To get the full action describing the motion between times t_0 and t_1 , we sum up all the infinitesimal actions using integration, so that

$$S[\underline{x}] = \int_{t_0}^{t_1} \mathcal{L}(t, \underline{x}(t), \dot{\underline{x}}(t)) dt \quad (4.2)$$

The full action $S[\underline{x}]$ is a functional of the function $\underline{x}(t)$ describing the path. Note that the action is subject to the boundary conditions

$$\underline{x}(t_0) = \underline{x}_0, \quad \underline{x}(t_1) = \underline{x}_1 \quad (4.3)$$

The action is essentially a measure of the average Lagrangian over the relevant time interval.

Remember Nature is lazy and chooses the path of least action. To find the path we need to know how to minimize the functional $S[\underline{x}]$ with respect to the function $\underline{x}(t)$. This requires *Calculus of Variations*, introduced by Euler and Lagrange. Forget functionals for a second. Recall how we find the location of the minimum/maximum of a ordinary function, $F(t)$. The minimum/maximum occurs at $t = t_*$, where $F'(t_*) = 0$. This is the same as saying

$$\lim_{\epsilon \rightarrow 0} \frac{F(t_* + \epsilon) - F(t_*)}{\epsilon} = 0 \quad (4.4)$$

In other words, a small displacement from the minimum/maximum yields no effect to first order.

Now let's apply the same logic to the functional $S[\underline{x}]$. We take the path corresponding to the minimum to be given by $\underline{x}_*(t)$, with action $S[\underline{x}_*]$. Now consider a small virtual displacement around that path $\underline{x}_*(t) + \epsilon \underline{f}(t)$, where the vector valued function $\underline{f}(t)$ is completely *arbitrary*, up to satisfying the boundary conditions $\underline{f}(t_0) = \underline{f}(t_1) = \underline{0}$. These boundary conditions on $\underline{f}(t)$ ensure that the displaced path satisfies the same boundary conditions as $\underline{x}_*(t)$, given by Eq. 4.3. Since $\underline{x}_*(t)$ is the location of the minimum, we have

$$\lim_{\epsilon \rightarrow 0} \frac{S[\underline{x}_* + \epsilon \underline{f}] - S[\underline{x}_*]}{\epsilon} = 0 \quad (4.5)$$

4.1 Calculus of Variations in one dimension

To develop this further, we start with the simplest case of a particle moving in one dimension. The action has the form

$$S[x] = \int_{t_0}^{t_1} \mathcal{L}(t, x(t), \dot{x}(t)) dt \quad (4.6)$$

where the path is a scalar valued function $x(t)$ satisfying the boundary conditions $x(t_0) = x_0$, $x(t_1) = x_1$. To find the path $x_*(t)$ that minimizes the action, we introduce an arbitrary function, $f(t)$, satisfying $f(t_0) = f(t_1) = 0$ and set

$$\lim_{\epsilon \rightarrow 0} \frac{S[x_* + \epsilon f] - S[x_*]}{\epsilon} = 0 \quad (4.7)$$

However

$$S[x_* + \epsilon f] = \int_{t_0}^{t_1} \mathcal{L}(t, x_*(t) + \epsilon f(t), \dot{x}_*(t) + \epsilon \dot{f}(t)) dt \quad (4.8)$$

$$= \int_{t_0}^{t_1} \left[\mathcal{L}(t, x_*(t), \dot{x}_*(t)) + \epsilon f \frac{\partial \mathcal{L}}{\partial x} \Big|_{x=x_*} + \epsilon \dot{f} \frac{\partial \mathcal{L}}{\partial \dot{x}} \Big|_{x=x_*} + \mathcal{O}(\epsilon^2) \right] dt \quad (4.9)$$

$$= S[x_*] + \epsilon \int_{t_0}^{t_1} \left[f \frac{\partial \mathcal{L}}{\partial x} \Big|_{x=x_*} + \dot{f} \frac{\partial \mathcal{L}}{\partial \dot{x}} \Big|_{x=x_*} \right] dt + \mathcal{O}(\epsilon^2) \quad (4.10)$$

After an integration by parts, it follows that

$$\lim_{\epsilon \rightarrow 0} \frac{S[x_* + \epsilon f] - S[x_*]}{\epsilon} = \left[f \frac{\partial \mathcal{L}}{\partial \dot{x}} \right]_{t_0}^{t_1} + \int_{t_0}^{t_1} f(t) \left[\frac{\partial \mathcal{L}}{\partial x} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}} \right) \right] dt \quad (4.11)$$

The first term on the right hand side of Eq. 4.11 vanishes since $f(t_0) = f(t_1) = 0$. Since the left hand side must also vanish by Eq. 4.7, we obtain the following equation, valid for any choice of the function $f(t)$ satisfying the aforementioned boundary conditions.

$$\int_{t_0}^{t_1} f(t) \left[\frac{\partial \mathcal{L}}{\partial x} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}} \right) \right] dt = 0 \quad (4.12)$$

Since $f(t)$ is indeed arbitrary at generic values of t , it follows that

$$\frac{\partial \mathcal{L}}{\partial x} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}} \right) = 0 \quad (4.13)$$

This equation of motion is satisfied by the path of least action $x_*(t)$. It is known as the *Euler-Lagrange equation*.

Example: Consider a Lagrangian of the form $\mathcal{L} = \frac{1}{2}m\dot{x}^2 - V(x)$. Then

$$S[x] = \int_{t_0}^{t_1} \frac{1}{2}m\dot{x}^2 - V(x) dt \quad (4.14)$$

We introduce the arbitrary function $f(t)$ satisfying $f(t_0) = f(t_1) = 0$, and set

$$\lim_{\epsilon \rightarrow 0} \frac{S[x + \epsilon f] - S[x]}{\epsilon} = 0 \quad (4.15)$$

However,

$$S[x + \epsilon f] = \int_{t_0}^{t_1} \left[\frac{1}{2}m(\dot{x} + \epsilon \dot{f})^2 - V(x + \epsilon f) \right] dt \quad (4.16)$$

$$= \int_{t_0}^{t_1} \left[\frac{1}{2}m\dot{x}^2 + m\epsilon \dot{x} \dot{f} - V(x) - \epsilon f V'(x) + \mathcal{O}(\epsilon^2) \right] dt \quad (4.17)$$

$$= S[x] + \epsilon \int_{t_0}^{t_1} \left[m\dot{x} \dot{f} - f V'(x) \right] dt + \mathcal{O}(\epsilon^2) \quad (4.18)$$

After integrating by parts we find that

$$\lim_{\epsilon \rightarrow 0} \frac{S[x + \epsilon f] - S[x]}{\epsilon} = [m\dot{x}f]_{t_0}^{t_1} - \int_{t_0}^{t_1} f(t) [m\ddot{x} + V'(x)] dt \quad (4.19)$$

The first term on the right hand side of this equation vanishes due to the boundary conditions on f . Since we want the left hand side to be zero, for any function $f(t)$ satisfying $f(t_0) = f(t_1) = 0$, it follows that

$$m\ddot{x} + V'(x) = 0 \quad (4.20)$$

Exercise: Check that this equation is the same as the Euler-Lagrange equation for the Lagrangian $\mathcal{L} = \frac{1}{2}m\dot{x}^2 - V(x)$.

Interestingly, note that Eq. 4.20 can be written as

$$m\ddot{x} = -V'(x) \quad (4.21)$$

which is just N2L for a particle of mass m moving in one dimension, under the action of a conservative force $F = -V'(x)$.

Exercise: Show that the Euler-Lagrange equation for a Lagrangian $\mathcal{L} = \mathcal{L}(\underline{x}(t), \dot{\underline{x}}(t), \ddot{\underline{x}}(t))$ is given by

$$\frac{\partial \mathcal{L}}{\partial \underline{x}} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{\underline{x}}} \right) + \frac{d^2}{dt^2} \left(\frac{\partial \mathcal{L}}{\partial \ddot{\underline{x}}} \right) = 0$$

Hint: You will need to perform integration by parts more than once.

Why do we need to impose the dynamical boundary conditions $\left. \frac{\partial \mathcal{L}}{\partial \ddot{\underline{x}}} \right|_{t_0} = \left. \frac{\partial \mathcal{L}}{\partial \ddot{\underline{x}}} \right|_{t_1} = 0$?

We have now learnt the basics of Calculus of Variations, and used it to minimize a general action. Strictly speaking the solutions to the Euler-Lagrange equations find extrema of the action, that is minima *and* maxima. In this sense, the principle of least action should be more accurately described as the principle of stationary action. In any case case, we have already seen how N2L may be recovered from these techniques, at least in the example we chose. We will continue to develop this further.

4.2 Calculus of Variations in more than one dimension

We will now return to the case where the particle moves in more than one dimension (eg. three), so its path is described by a vector valued function $\underline{x}(t)$. Let us start with an example, where the Lagrangian is given by

$$\mathcal{L}(\underline{x}(t), \dot{\underline{x}}(t)) = \frac{1}{2}m|\dot{\underline{x}}|^2 - V(\underline{x}) \quad (4.22)$$

The action is

$$S[\underline{x}] = \int_{t_0}^{t_1} \frac{1}{2}m|\dot{\underline{x}}|^2 - V(\underline{x}) dt \quad (4.23)$$

with the boundary conditions $\underline{x}(t_0) = \underline{x}_0$ $\underline{x}(t_1) = \underline{x}_1$. To minimize the action we introduce the arbitrary vector valued function $\underline{f}(t)$ satisfying $\underline{f}(t_0) = \underline{f}(t_1) = 0$, and set

$$\lim_{\epsilon \rightarrow 0} \frac{S[\underline{x} + \epsilon \underline{f}] - S[\underline{x}]}{\epsilon} = 0 \quad (4.24)$$

Now

$$S[\underline{x} + \epsilon \underline{f}] = \int_{t_0}^{t_1} \frac{1}{2}m|\dot{\underline{x}} + \epsilon \dot{\underline{f}}|^2 - V(\underline{x} + \epsilon \underline{f}) dt \quad (4.25)$$

where

$$|\underline{\dot{x}} + \epsilon \underline{\dot{f}}|^2 = (\underline{\dot{x}} + \epsilon \underline{\dot{f}}) \cdot (\underline{\dot{x}} + \epsilon \underline{\dot{f}}) = |\underline{\dot{x}}|^2 + 2\epsilon \underline{\dot{x}} \cdot \underline{\dot{f}} + \epsilon^2 |\underline{\dot{f}}|^2 \quad (4.26)$$

$$V(\underline{x} + \epsilon \underline{f}) = V(\underline{x}) + \epsilon \underline{f} \cdot \underline{\nabla} V(\underline{x}) + \mathcal{O}(\epsilon^2) \quad (4.27)$$

It follows that

$$S[\underline{x} + \epsilon \underline{f}] = S[\underline{x}] + \epsilon \int_{t_0}^{t_1} m \underline{\dot{x}} \cdot \underline{\dot{f}} - \underline{f} \cdot \underline{\nabla} V dt + \mathcal{O}(\epsilon^2) \quad (4.28)$$

$$\Rightarrow \lim_{\epsilon \rightarrow 0} \frac{S[\underline{x} + \epsilon \underline{f}] - S[\underline{x}]}{\epsilon} = [m \underline{\dot{x}} \cdot \underline{\dot{f}}]_{t_0}^{t_1} - \int_{t_0}^{t_1} \underline{f} \cdot [m \underline{\ddot{x}} + \underline{\nabla} V] dt \quad (4.29)$$

where we have performed an integration by parts. The first term on the right hand side of Eq. 4.29 vanishes due to the boundary conditions on \underline{f} . Since we want the left hand side to be zero, for any function $\underline{f}(t)$ satisfying $\underline{f}(t_0) = \underline{f}(t_1) = 0$, it follows that

$$m \underline{\ddot{x}} + \underline{\nabla} V = 0 \quad (4.30)$$

This generalises our previous result, since it corresponds to N2L for a conservative force $\underline{F} = -\underline{\nabla} V$.

To study the general case, it is convenient to abandon vector notation, and write \underline{x} in terms of its components, x_i . That is

$$\underline{x} = (x_1, x_2, \dots, x_D)$$

where D is the number of dimensions. The Lagrangian $\mathcal{L}(\underline{x}(t), \underline{\dot{x}}(t))$ is often written in the form $\mathcal{L}(x_i(t), \dot{x}_i(t))$, where it is understood that we should include all components $i = 1, 2, \dots, D$. Now a general action takes the form

$$S[\underline{x}] = \int_{t_0}^{t_1} \mathcal{L}(t, x_i(t), \dot{x}_i(t)) dt \quad (4.31)$$

To minimize this we introduce the arbitrary function $\underline{f}(t)$, with components $f_i(t)$, satisfying the boundary conditions $f_i(t_0) = f_i(t_1) = 0$, and impose Eq. 4.24. Now

$$S[\underline{x} + \epsilon \underline{f}] = \int_{t_0}^{t_1} \mathcal{L}(t, x_i(t) + \epsilon f_i(t), \dot{x}_i(t) + \epsilon \dot{f}_i(t)) dt \quad (4.32)$$

where

$$\mathcal{L}(t, x_i(t) + \epsilon f_i(t), \dot{x}_i(t) + \epsilon \dot{f}_i(t)) = \mathcal{L}(t, x_i(t), \dot{x}_i(t)) + \sum_{i=1}^D \epsilon f_i \frac{\partial \mathcal{L}}{\partial x_i} + \epsilon \dot{f}_i \frac{\partial \mathcal{L}}{\partial \dot{x}_i} + \mathcal{O}(\epsilon^2) \quad (4.33)$$

It follows that

$$S[\underline{x} + \epsilon \underline{f}] = S[\underline{x}] + \epsilon \int_{t_0}^{t_1} \left[\sum_{i=1}^D f_i \frac{\partial \mathcal{L}}{\partial x_i} + \dot{f}_i \frac{\partial \mathcal{L}}{\partial \dot{x}_i} \right] dt + \mathcal{O}(\epsilon^2) \quad (4.34)$$

$$\Rightarrow \lim_{\epsilon \rightarrow 0} \frac{S[\underline{x} + \epsilon \underline{f}] - S[\underline{x}]}{\epsilon} = \left[\sum_{i=1}^D f_i \frac{\partial \mathcal{L}}{\partial \dot{x}_i} \right]_{t_0}^{t_1} + \sum_{i=1}^D \int_{t_0}^{t_1} f_i \left[\frac{\partial \mathcal{L}}{\partial x_i} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}_i} \right) \right] dt \quad (4.35)$$

where we have performed an integration by parts. The first term on the right hand side of Eq. 4.35 vanishes due to the boundary conditions on the f_i . Since we want the left hand side to be zero, for all functions $f_i(t)$ satisfying $f_i(t_0) = f_i(t_1) = 0$, it follows that

$$\frac{\partial \mathcal{L}}{\partial x_i} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}_i} \right) = 0, \quad i = 1, 2, \dots, D \quad (4.36)$$

This represents a family of D equations, corresponding to the Euler-Lagrange equations in D dimensions.

Check: Let us check that Eq. 4.36 yields Eq. 4.30 for the Lagrangian

$$\mathcal{L}(\underline{x}(t), \dot{\underline{x}}(t)) = \frac{1}{2}m|\dot{\underline{x}}|^2 - V(\underline{x})$$

In component form, this is written as

$$\begin{aligned} \mathcal{L}(x_i, \dot{x}_i) &= \frac{1}{2}m \left(\sum_{i=1}^D \dot{x}_i^2 \right) - V(x_i) \\ \Rightarrow \frac{\partial \mathcal{L}}{\partial x_i} &= -\frac{\partial V}{\partial x_i}, \quad \frac{\partial \mathcal{L}}{\partial \dot{x}_i} = m\dot{x}_i \end{aligned}$$

Plugging these expressions into the Euler-Lagrange Eq. 4.36 yields

$$m\ddot{x}_i + \frac{\partial V}{\partial x_i} = 0$$

which is indeed equivalent to Eq. 4.30.

4.3 The physical Lagrangian

In the examples given we were able to recover N2L for a particle moving under the action of a conservative force using the Lagrangian,

$$\mathcal{L}(\underline{x}(t), \dot{\underline{x}}(t)) = \frac{1}{2}m|\dot{\underline{x}}|^2 - V(\underline{x})$$

Note that the kinetic energy of the particle is

$$T = \frac{1}{2}m|\dot{\underline{x}}|^2$$

whereas the potential energy of the particle is

$$V = V(\underline{x})$$

This means the Lagrangian takes the form

$$\mathcal{L} = T - V \quad (4.37)$$

We take this to be the definition of the Lagrangian in dynamical systems. As we have seen, Newton's Laws can be derived by simply minimizing the average of this quantity, and yet it is nothing more than the difference between the kinetic and potential energy of the system.

5 Generalised Coordinates and Hamilton's Principle

Up until now we have been studying a single particle system described by its position $\underline{x} = (x_1, \dots, x_D)$, written in terms of Cartesian coordinates in D dimensions. What happens when there is more than one particle, or if we wish to use an alternative coordinate system such as spherical polars. One of the cute things about the Lagrangian formulation we have been discussing is that it copes easily with many particles and alternative coordinate systems via the notion of *generalised coordinates*. To get a feel for generalised coordinates, we stick, for the moment at least, to a single particle system. Suppose we want to switch from our Cartesian coordinates, $\underline{x}(t) = (x_1(t), \dots, x_D(t))$, to another coordinate system $\underline{q}(t) = (q_1(t), \dots, q_D(t))$. The new coordinates can be defined in terms of the original ones by means of a coordinate transformation

$$\underline{q}(t) = \underline{q}(\underline{x}(t), t) = \underline{q}(x_1(t), \dots, x_D(t), t)$$

In component language this can be written more compactly as

$$q_i(t) = q_i(x_j(t), t)$$

In words, this notation reflects the fact that the i th component of $\underline{q}(t)$ can be regarded as a function of t and each of the $x_j(t)$, where $j = 1, \dots, D$.

Now we further assume that we can invert the transformation so that $x_j(t) = x_j(q_i(t), t)$. In the new coordinates we introduce the *generalised velocity*

$$\dot{q}_i = \frac{d}{dt}q_i(x_j(t), t) = \frac{\partial q_i}{\partial t} + \sum_{j=1}^D \frac{\partial q_i}{\partial x_j} \dot{x}_j \quad (5.1)$$

where we have used the Chain Rule. Similarly we have

$$\dot{x}_j = \frac{d}{dt}x_j(q_i(t), t) = \frac{\partial x_j}{\partial t} + \sum_{i=1}^D \frac{\partial x_j}{\partial q_i} \dot{q}_i \quad (5.2)$$

Now consider the Lagrangian $\mathcal{L}(t, x_i, \dot{x}_i)$. We know that the Euler-Lagrange equations are

$$\frac{\partial \mathcal{L}}{\partial x_i} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}_i} \right) = 0$$

Now lets rewrite the Lagrangian in terms of the new coordinates

$$\mathcal{L}(t, x_j(q_i, t), \dot{x}_j(q_i, \dot{q}_i, t)) = \mathcal{L}'(t, q_i, \dot{q}_i)$$

where we recall the formula for \dot{x}_j given by Eq. 5.2. Now, making repeated use of the chain rule, we obtain

$$\begin{aligned} \frac{\partial \mathcal{L}'}{\partial q_i} &= \sum_j \frac{\partial \mathcal{L}}{\partial x_j} \frac{\partial x_j}{\partial q_i} + \frac{\partial \mathcal{L}}{\partial \dot{x}_j} \frac{\partial \dot{x}_j}{\partial q_i} \\ \frac{\partial \mathcal{L}'}{\partial \dot{q}_i} &= \sum_j \frac{\partial \mathcal{L}}{\partial \dot{x}_j} \frac{\partial \dot{x}_j}{\partial \dot{q}_i} \end{aligned}$$

But by Eq. 5.2, we have $\frac{\partial \dot{x}_j}{\partial \dot{q}_i} = \frac{\partial x_j}{\partial q_i}$ and so

$$\begin{aligned} \frac{\partial \mathcal{L}'}{\partial \dot{q}_i} &= \sum_j \frac{\partial \mathcal{L}}{\partial \dot{x}_j} \frac{\partial x_j}{\partial q_i} \\ \Rightarrow \frac{d}{dt} \left(\frac{\partial \mathcal{L}'}{\partial \dot{q}_i} \right) &= \sum_j \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}_j} \right) \frac{\partial x_j}{\partial q_i} + \frac{\partial \mathcal{L}}{\partial \dot{x}_j} \frac{d}{dt} \left(\frac{\partial x_j}{\partial q_i} \right) = \sum_j \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}_j} \right) \frac{\partial x_j}{\partial q_i} + \frac{\partial \mathcal{L}}{\partial \dot{x}_j} \frac{\partial \dot{x}_j}{\partial q_i} \end{aligned}$$

Consider now the analogue of the Euler-Lagrange equation in the new coordinate system. We can see that it holds, as expected since,

$$\begin{aligned} \frac{\partial \mathcal{L}'}{\partial q_i} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}'}{\partial \dot{q}_i} \right) &= \sum_j \left[\frac{\partial \mathcal{L}}{\partial x_i} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}_i} \right) \right] \frac{\partial x_j}{\partial q_i} \\ \Rightarrow \frac{\partial \mathcal{L}'}{\partial q_i} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}'}{\partial \dot{q}_i} \right) &= 0 \iff \frac{\partial \mathcal{L}}{\partial x_i} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}_i} \right) = 0 \end{aligned}$$

In other words, if the Euler-Lagrange equation holds in one coordinate system, it holds in the other.

Exercise: Consider a particle of mass m moving in the xy plane under the influence of a potential $V = -G_N M m / \sqrt{x^2 + y^2}$. If the position of the particle is given by $(x(t), y(t))$, show that the Lagrangian is given by

$$\mathcal{L} = \frac{1}{2} m (\dot{x}^2 + \dot{y}^2) + \frac{G_N M m}{\sqrt{x^2 + y^2}}$$

and derive the equations of motion.

Now change coordinates to polar coordinates

$$x = r \cos \theta, \quad y = r \sin \theta.$$

Calculate the Lagrangian and derive the equations of motion in the new coordinate system. By working directly with the equations of motion, show that they are equivalent in the two coordinate systems.

5.1 Many particle systems

The power of generalised coordinates really comes into play in understanding systems of many particles. Consider a system of N particles moving in D dimensions. If the k -th particle has position $\underline{x}^{(k)}(t)$, then the Lagrangian takes the form

$$\mathcal{L}(t, \underline{x}^{(1)}(t), \dots, \underline{x}^{(N)}(t), \dot{\underline{x}}^{(1)}(t), \dots, \dot{\underline{x}}^{(N)}(t))$$

which is sometimes written in the more compact form $\mathcal{L}(t, \underline{x}_i^{(k)}(t), \dot{\underline{x}}_i^{(k)}(t))$. Since $\underline{x}^{(1)}(t), \dots, \underline{x}^{(N)}(t)$ describes N particles in real D dimensional space (see Fig. 5.1), the system carries a total of

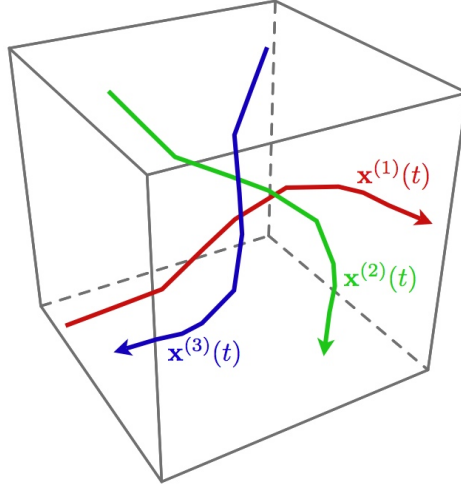


Figure 5.1: N particle paths in real D dimensional space.

DN degrees of freedom: one for each component of each particle. Therefore it is often much more convenient to introduce a single path in DN dimensional *configuration space* (see Fig. 5.2), described by

$$\underline{X}(t) = (\underline{x}^{(1)}(t), \dots, \underline{x}^{(N)}(t))$$

Each $\underline{x}^{(k)}(t)$ contributes D components to $\underline{X}(t)$, and there are N of these, so $\underline{X}(t)$ has DN components. With this in mind, it is convenient to write $\underline{X}(t)$ in components as $X_A(t)$ where $A = 1, \dots, DN$. In configuration space we have a single path with DN components, and therefore DN degrees of freedom describing the system. The number of degrees of freedom is the same whether we work in real or configuration space, as of course it should be since the two descriptions are equivalent. The Lagrangian can now be written in terms of the single path, $X_A(t)$,

$$\mathcal{L}(t, X_A(t), \dot{X}_A(t))$$

It is now clear that the Euler-Lagrange equations must take the form

$$\frac{\partial \mathcal{L}}{\partial X_A} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{X}_A} \right) = 0, \quad A = 1, \dots, DN$$

which is equivalent to

$$\frac{\partial \mathcal{L}}{\partial x_i^{(k)}} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}_i^{(k)}} \right) = 0, \quad i = 1, \dots, D, \quad k = 1, \dots, N$$

Example: Suppose that the k th particle has mass m_k , then the total kinetic energy of the system is

$$T = \sum_{k=1}^N \frac{1}{2} m_k |\dot{\underline{x}}^{(k)}|^2$$

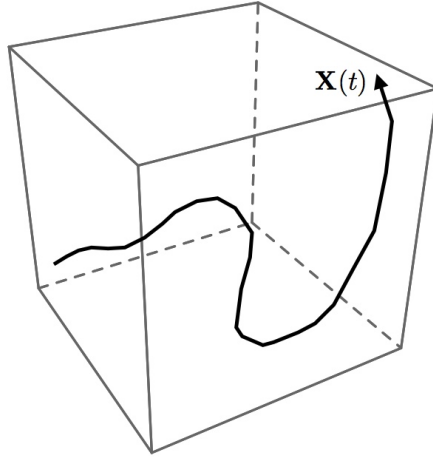


Figure 5.2: A single path in DN dimensional configuration space.

If the potential $V = V(\underline{x}^{(1)}, \dots, \underline{x}^{(N)})$, then given the Lagrangian $\mathcal{L} = T - V$, the Euler-Lagrange equations yield

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x_i^{(k)}} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}_i^{(k)}} \right) &= 0 \\ \Rightarrow -\frac{\partial V}{\partial x_i^{(k)}} &= \frac{d}{dt} \left(m_k \dot{x}_i^{(k)} \right) \end{aligned} \quad (5.3)$$

Now $F_i^{(k)} = -\frac{\partial V}{\partial x_i^{(k)}}$ is the conservative force acting on the k th particle, and $p_i^{(k)} = m_k \dot{x}_i^{(k)}$ is its momentum. Therefore Eq. 5.3 is nothing more than N2L for a many particle system

$$F_i^{(k)} = \frac{d}{dt} p_i^{(k)}$$

Going back to the single path in configuration space, $X_A(t)$, it is clear that we can readily switch to generalised coordinates

$$q_A(t) = q_A(X_A(t), t)$$

Returning to real space, this is equivalent to defining generalised coordinates that can, in principle, mix up the paths, ie

$$q_i^{(k)}(t) = q_i^{(k)}(\underline{x}^{(1)}(t), \dots, \underline{x}^{(N)}(t), t)$$

This is often a very convenient thing to do, as we shall illustrate shortly when we take a look at the “two body problem”. Before that, however, let us conclude this section with a

formal statement of Hamilton's principle.

Hamilton's principle

For a system described by generalised coordinates, q_A , the correct path of motion $q_A(t)$ between the initial state $q_A(t_0)$ at time t_0 and the final state $q_A(t_1)$ at time t_1 corresponds to a stationary path of the action

$$S = \int_{t_0}^{t_1} \mathcal{L} dt$$

where $\mathcal{L} = \mathcal{L}(t, q_A, \dot{q}_A)$ is the Lagrangian describing the system.

The stationary path can be found by solving the Euler-Lagrange equations,

$$\frac{\partial \mathcal{L}}{\partial q_A} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_A} \right) = 0.$$

This works just as well for many particle systems as it does for single particles. The key thing to note is that each degree of freedom in the system corresponds to one component of the generalised coordinates describing the system.

6 The two body problem

Earlier on in the course we studied the orbit of a single particle around a fixed source. We are now in a position to consider what happens when we allow the source to move, as of course it would in the case of the Sun. This represents a particularly nice example of using generalised coordinates to simplify the dynamics.

Consider two heavy particles of mass m_1 and m_2 with positions $\underline{x}^{(1)}(t)$ and $\underline{x}^{(2)}(t)$ respectively (see Fig. 6.1). To calculate the Lagrangian we need to work out the kinetic and potential energy of the system.

$$\text{Kinetic energy: } T = \frac{1}{2}m_1|\dot{\underline{x}}^{(1)}|^2 + \frac{1}{2}m_2|\dot{\underline{x}}^{(2)}|^2 \quad (6.1)$$

$$\text{Potential energy: } V = -\frac{G_N m_1 m_2}{|\underline{x}^{(1)} - \underline{x}^{(2)}|} \quad (6.2)$$

$$\Rightarrow \mathcal{L} = T - V = \frac{1}{2}m_1|\dot{\underline{x}}^{(1)}|^2 + \frac{1}{2}m_2|\dot{\underline{x}}^{(2)}|^2 + \frac{G_N m_1 m_2}{|\underline{x}^{(1)} - \underline{x}^{(2)}|} \quad (6.3)$$

In principle we could work directly with the coordinate system $\underline{x}^{(1)} = (x_1^{(1)}, x_2^{(1)}, x_3^{(1)})$, $\underline{x}^{(2)} = (x_1^{(2)}, x_2^{(2)}, x_3^{(2)})$, and calculate the Euler-Lagrange equations directly. However, this will be a

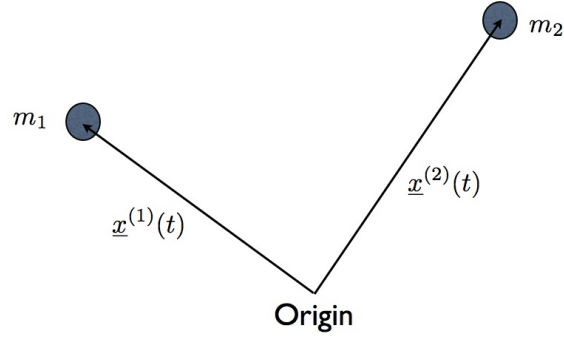


Figure 6.1: The two body problem.

messy business. It is much simpler to change to generalised coordinates defined by

$$\text{Centre of mass position: } \underline{q}^{(1)}(t) = \frac{m_1 \underline{x}^{(1)} + m_2 \underline{x}^{(2)}}{m_1 + m_2} \quad (6.4)$$

$$\text{Relative position: } \underline{q}^{(2)}(t) = \underline{x}^{(1)} - \underline{x}^{(2)} \quad (6.5)$$

It follows that

$$\begin{aligned} \underline{x}^{(1)} &= \underline{q}^{(1)} + \frac{m_2}{m_1 + m_2} \underline{q}^{(2)}, & \underline{x}^{(2)} &= \underline{q}^{(1)} - \frac{m_1}{m_1 + m_2} \underline{q}^{(2)} \\ \Rightarrow |\dot{\underline{x}}^{(1)}|^2 &= |\dot{\underline{q}}^{(1)}|^2 + \frac{2m_2}{m_1 + m_2} \dot{\underline{q}}^{(1)} \cdot \dot{\underline{q}}^{(2)} + \left(\frac{m_2}{m_1 + m_2} \right)^2 |\dot{\underline{q}}^{(2)}|^2 \end{aligned} \quad (6.6)$$

A similar expression to Eq. 6.6 holds for $|\dot{\underline{x}}^{(2)}|^2$. Plugging these expressions into Eq. 6.3, we write the Lagrangian as

$$\mathcal{L} = \frac{1}{2}(m_1 + m_2)|\dot{\underline{q}}^{(1)}|^2 + \frac{1}{2} \left(\frac{m_1 m_2}{m_1 + m_2} \right) |\dot{\underline{q}}^{(2)}|^2 + \frac{G_N m_1 m_2}{|\underline{q}^{(2)}|}$$

The nice thing about this choice of generalised coordinates is that the Lagrangian no longer depends explicitly on $\underline{q}^{(1)}(t)$, ie $\frac{\partial \mathcal{L}}{\partial \underline{q}^{(1)}} = 0$. From the Euler-Lagrange equation for $\underline{q}^{(1)}$, it follows that

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{\underline{q}}^{(1)}} \right) = 0$$

This means the relevant (conjugate) momentum

$$p_i^{(1)} = \frac{\partial \mathcal{L}}{\partial \dot{q}_i^{(1)}} = (m_1 + m_2) \dot{q}_i^{(1)} = \text{constant}$$

where the $p_i^{(1)}$ are the components of momentum $\underline{p}^{(1)}$. The centre of mass position $\underline{q}^{(1)}$ is an example of an *ignorable coordinate* (more later). It follows that the centre of mass moves with a constant velocity \underline{v} . Now consider the Euler-Lagrange equation for the other coordinate, $\underline{q}^{(2)}$. This gives

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial q_i^{(2)}} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i^{(2)}} \right) &= 0 \\ \Rightarrow G_N m_1 m_2 \frac{\partial}{\partial q_i^{(2)}} \left(\frac{1}{|\underline{q}^{(2)}|} \right) - \left(\frac{m_1 m_2}{m_1 + m_2} \right) \ddot{q}_i^{(2)} &= 0 \end{aligned} \quad (6.7)$$

Now lets perform a little careful calculus,

$$\frac{\partial}{\partial q_i^{(2)}} \left(\frac{1}{|\underline{q}^{(2)}|} \right) = -\frac{1}{|\underline{q}^{(2)}|^2} \frac{\partial |\underline{q}^{(2)}|}{\partial q_i^{(2)}} = -\frac{1}{2|\underline{q}^{(2)}|^3} \frac{\partial |\underline{q}^{(2)}|^2}{\partial q_i^{(2)}} = -\frac{q_i^{(2)}}{|\underline{q}^{(2)}|^3} = -\frac{\hat{q}_i^{(2)}}{|\underline{q}^{(2)}|^2}$$

where $\hat{\underline{q}}^{(2)} = \frac{\underline{q}^{(2)}}{|\underline{q}^{(2)}|}$ is the unit vector parallel to $\underline{q}^{(2)}$. The resulting equations of motion are

$$\begin{aligned} \left(\frac{m_1 m_2}{m_1 + m_2} \right) \ddot{\underline{q}}^{(2)} &= -\frac{G_N m_1 m_2}{|\underline{q}^{(2)}|^2} \hat{\underline{q}}^{(2)} \\ \Rightarrow \ddot{\underline{q}}^{(2)} &= -\frac{G_N M}{|\underline{q}^{(2)}|^2} \hat{\underline{q}}^{(2)}, \quad M = m_1 + m_2 \end{aligned} \quad (6.8)$$

We immediately see that the relative motion is equivalent to the motion of a single particle around a fixed source of total mass $M = m_1 + m_2$. What we have done here is describe the motion in terms of the centre of mass and the relative motion of the two particles. The centre of mass moves with constant velocity whereas the relative motion is equivalent to the orbital motions around a fixed source of mass M . In particular there will exist bound orbits in which m_2 orbits m_1 along an ellipse, with m_1 at a focus (and vice versa). Kepler's Laws clearly still apply, although some of the constants of proportionality will have been modified.

Exercise: Show that m_1 also behaves like a particle orbiting a fixed source at the centre of mass position, with mass m_2^3/M^2 . How does m_2 behave relative to the centre of mass.

Hint: The position of m_1 relative to the centre of mass is given by $\underline{s}^{(1)} = \underline{x}^{(1)} - \underline{q}^{(1)}$.

7 Rigid bodies

In our study of planetary dynamics, we were able to treat the planets as point like particles. This approximation works because the scale of the orbit is much larger than the characteristic

scale of the planet. Indeed, for the earth, the orbital scale is around 150 million km, whereas the planetary scale is around 6400 km (the radius of the earth).

However, there are many problems in physics where the size, shape and internal structure of an extended object become important. Examples include the spinning top, or the compound pendulum. This necessitates a study of *rigid bodies* – extended objects held together by internal forces.

We can model a rigid body as a system of particles, where their positions relative to one another are fixed. For example (see Fig. 7.1)

- heavy particles joined together by light rods
- solid bodies, like the earth or a brick.

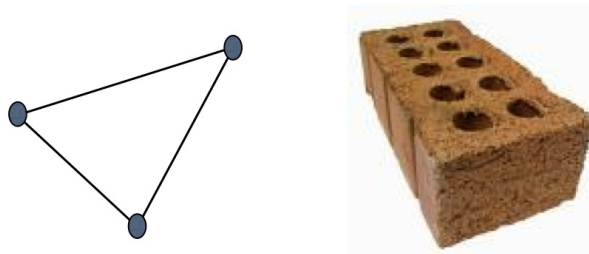


Figure 7.1: Examples of rigid bodies: (a) heavy particles joined together by light rods, and (b) a solid brick.

For a solid body $\sum_{\text{particles}} \rightarrow \int_{\text{volume}}$, eg $\sum_k m_k \rightarrow \int \rho(\underline{x})dV$ where $\rho(\underline{x})$ is the mass density at position \underline{x} .

7.1 Euler's theorem

The general displacement of a rigid body with one fixed point O is a rotation about some axis through O (see Fig. 7.2)

7.2 Angular velocity about a fixed point

Choose any point in the rigid body with position $\underline{r}(t)$, relative to the fixed point O at time t . At a time $t + \delta t$ the point has moved to a position $\underline{r}(t + \delta t)$. By Euler's theorem this

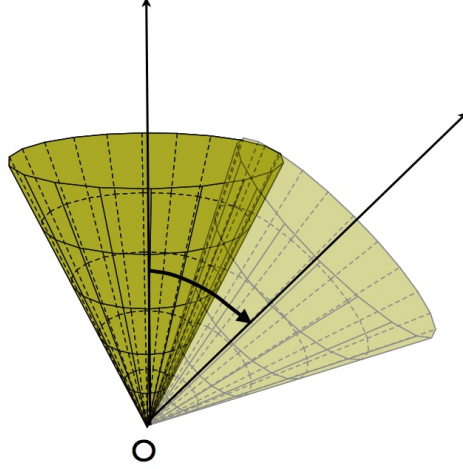


Figure 7.2: Euler's theorem.

displacement must correspond to a rotation through an angle $\delta\phi$ about an axis through O. Let \hat{n} denote the unit vector along that axis, and assume that \underline{r} makes an angle θ with \hat{n} (see Fig. 7.3) The point gets displaced by a vector $\delta\underline{r} = \underline{r}(t + \delta t) - \underline{r}(t)$. Note that $\delta\underline{r}$ is perpendicular to both \hat{n} and \underline{r} and is therefore parallel to $\hat{n} \times \underline{r}$. Furthermore, the displacement vector has magnitude $|\delta\underline{r}| = |\underline{r}| \sin\theta \delta\phi$. This implies

$$\delta\underline{r} = (\hat{n} \times \underline{r})\delta\phi$$

Now since $\delta\underline{r} \approx \dot{\underline{r}}\delta t$ and $\delta\phi \approx \dot{\phi}\delta t$ for small δt , we take $\delta t \rightarrow 0$ and find

$$\dot{\underline{r}} = (\hat{n} \times \underline{r})\dot{\phi} = \underline{w} \times \underline{r}$$

where $\underline{w} = \dot{\phi}\hat{n}$ is defined as the *angular velocity vector*. The angular velocity vector has magnitude, $\dot{\phi}$, corresponding to the rate of rotation, and direction, \hat{n} , along the axis of rotation.

7.3 Angular momentum about a fixed point

A rigid body can be modelled as a system of particles of mass m_k located at position $\underline{x}^{(k)}(t)$, relative to the fixed point. Such a particle has (linear) momentum

$$\underline{p}^{(k)} = m_k \dot{\underline{x}}^{(k)} = m_k (\underline{w} \times \underline{x}^{(k)})$$

and *angular momentum*

$$\underline{h}^{(k)} = \underline{x}^{(k)} \times \underline{p}^{(k)} = m_k \underline{x}^{(k)} \times (\underline{w} \times \underline{x}^{(k)}) \quad (7.1)$$

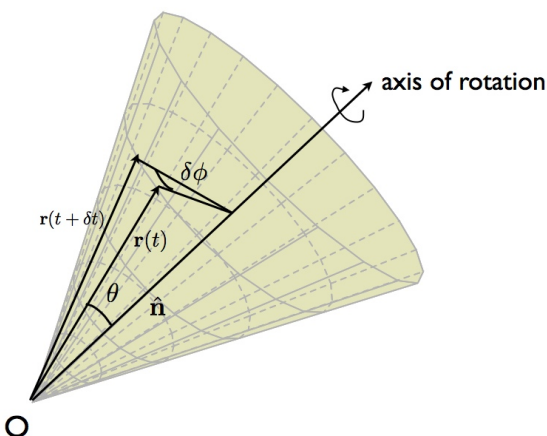


Figure 7.3: The rigid body rotates with angular speed $\dot{\phi}$ about an axis parallel to \hat{n} .

Making use of the identity $\underline{a} \times (\underline{b} \times \underline{c}) = \underline{b}(\underline{a} \cdot \underline{c}) - \underline{c}(\underline{a} \cdot \underline{b})$, we find that

$$\underline{h}^{(k)} = m_k [|\underline{x}^{(k)}|^2 \underline{w} - (\underline{x}^{(k)} \cdot \underline{w}) \underline{x}^{(k)}] \quad (7.2)$$

To get the total angular momentum of the rigid body we simply sum up the angular momenta of the constituent particles,

$$\underline{h} = \sum_k \underline{h}^{(k)} = \sum_k m_k [|\underline{x}^{(k)}|^2 \underline{w} - (\underline{x}^{(k)} \cdot \underline{w}) \underline{x}^{(k)}]$$

For a solid body like the earth or a brick, the sum becomes an integral. To see this, note that an infinitesimal volume δV located at position $\underline{x}(t)$ inside the body has mass $\rho(\underline{x})\delta V$ where $\rho(\underline{x})$ is the local mass density (see Fig. 7.4). Therefore, the volume δV has angular momentum

$$\delta \underline{h} = \rho(\underline{x})\delta V [|\underline{x}|^2 \underline{w} - (\underline{x} \cdot \underline{w}) \underline{x}]$$

Integrating over the entire rigid body we obtain

$$\underline{h} = \int_{\text{body}} \rho(\underline{x}) [|\underline{x}|^2 \underline{w} - (\underline{x} \cdot \underline{w}) \underline{x}] dV$$

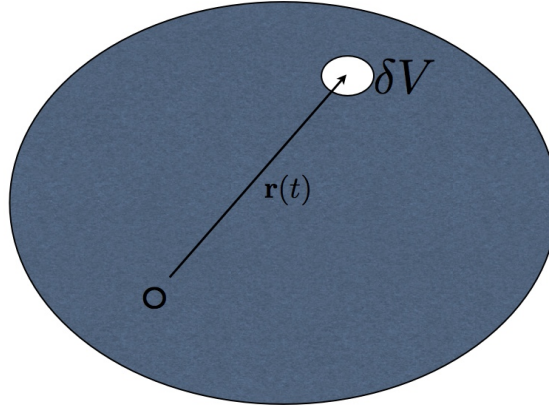


Figure 7.4: Integrating over rigid body.

In terms of components we have

$$\begin{aligned}
 h_i &= \int \rho [|\underline{x}|^2 w_i - (\underline{x} \cdot \underline{w}) x_i] dV \\
 &= \sum_j \int \rho [|\underline{x}|^2 w_i - (x_j w_j) x_i] dV \\
 &= \sum_j \left\{ \int \rho [|\underline{x}|^2 \delta_{ij} - x_i x_j] dV \right\} w_j \quad \text{where } \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \\
 &= \sum_j I_{ij} w_j
 \end{aligned} \tag{7.3}$$

We can write this last relation in vector notation, $\underline{h} = I\underline{w}$, where I is the *moment of inertia tensor*, with matrix components

$$I_{ij} = \int \rho [|\underline{x}|^2 \delta_{ij} - x_i x_j] dV \tag{7.4}$$

In rotational motion the inertia tensor plays the analogous role to inertial mass in linear motion – that is it measures the resistance to changes in the motion. Just as linear momentum is mass times velocity, so angular momentum is moment of inertia times angular velocity.

7.4 The moment of inertia tensor

Let's take a closer look at the moment of inertia tensor. We're probably more familiar with the notion of moment of inertia about a particular axis. This information is already encoded

in the inertia tensor. Indeed, if an axis lies along the direction of a unit vector \hat{n} , then the moment of inertia about that axis is given by $I_{nn} = \sum_{i,j} I_{ij} \hat{n}^i \hat{n}^j$. For example, working with Cartesian coordinates in 3 dimensions, the inertia tensor is given by

$$I = \begin{pmatrix} \int \rho(y^2 + z^2) dV & -\int \rho xy dV & -\int \rho xz dV \\ -\int \rho xy dV & \int \rho(x^2 + z^2) dV & -\int \rho yz dV \\ -\int \rho xz dV & -\int \rho yz dV & \int \rho(x^2 + y^2) dV \end{pmatrix}$$

where $dV = dx dy dz$. The moment of inertia about the x -axis would be given by $I_{xx} = \int \rho(y^2 + z^2) dV$

Principal moments of inertia

The inertia tensor is real and symmetric. This means we can always choose our axes so that I is diagonal.

To demonstrate this, let us consider a simple example in 2 dimensions. Consider a uniform square lamina of mass, M , and side L . We put the corner at the origin, O, and choose the x and y axis so that they coincide with adjacent sides (see Fig. 7.5) Since the

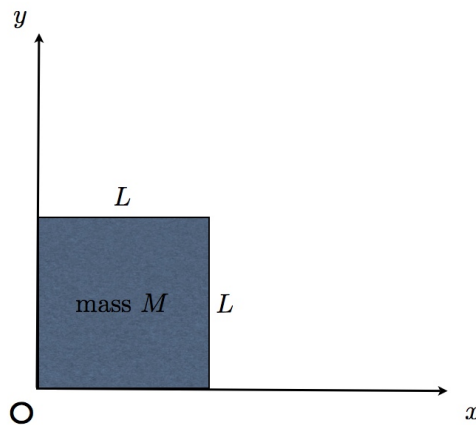


Figure 7.5: A square lamina with mass M and side L . The axes coincide with adjacent sides.

square is uniform it has constant density $\rho = M/L^2$. The inertia tensor is given by

$$I = \begin{pmatrix} \int \rho y^2 dV & -\int \rho xy dV \\ -\int \rho xy dV & \int \rho x^2 dV \end{pmatrix} = ML^2 \begin{pmatrix} \frac{1}{3} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{1}{3} \end{pmatrix}$$

Clearly this is not diagonal. Now define some new axes, x' and y' , corresponding to a rotation of $\frac{\pi}{4}$ about O (see Fig. 7.6). It is easy to see that

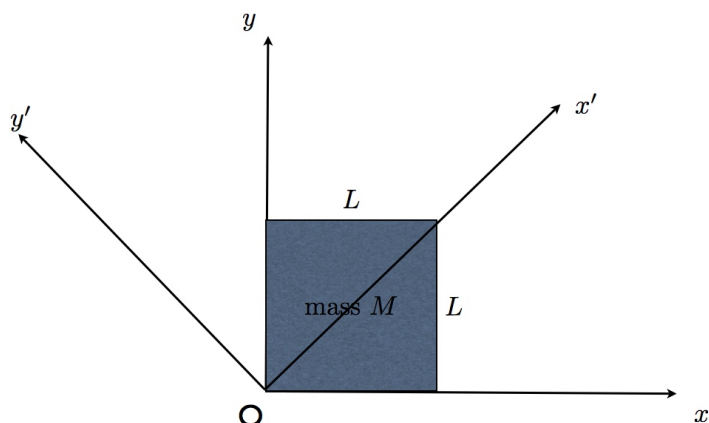


Figure 7.6: A square lamina with mass M and side L . The axes have been rotated through $\frac{\pi}{4}$ about O .

$$x' = \frac{1}{\sqrt{2}}(x + y), \quad y' = \frac{1}{\sqrt{2}}(y - x)$$

In the rotated frame the inertia tensor is

$$I' = \begin{pmatrix} \int \rho y'^2 dV & -\int \rho x' y' dV \\ -\int \rho x' y' dV & \int \rho x'^2 dV \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \int \rho (y - x)^2 dV & -\int \rho (y^2 - x^2) dV \\ -\int \rho (y^2 - x^2) dV & \int \rho (x + y)^2 dV \end{pmatrix} = ML^2 \begin{pmatrix} \frac{1}{12} & 0 \\ 0 & \frac{7}{12} \end{pmatrix}$$

which is diagonal. This demonstrates the fact that we can always rotate our axes to ensure that the inertia tensor is diagonal. In this example, we say that the x' and y' axes are the *principal axes*, with *principle moments of inertia*, $\frac{1}{12}ML^2$ and $\frac{7}{12}ML^2$ respectively.

Exercise: Show that the inertia tensor defined with respect to the original axes, x and y , has eigenvalues corresponding to the principal moments of inertia and eigenvectors parallel to the principal axes.

This is not a fluke. It is always the case. The principal moments of inertia and principal axes can always be found by working out the eigenvalues and eigenvectors of the inertia tensor. In practice, however, it often pays to spend a few seconds looking at the symmetries of the body relative to the fixed point. This will typically lead you into choosing the principal axes to begin with.

Perpendicular axis theorem

The moment of inertia of a lamina about an axis perpendicular to the lamina is the sum of the moments of inertia about two perpendicular axes in the plane of the lamina (see Fig.

7.7).

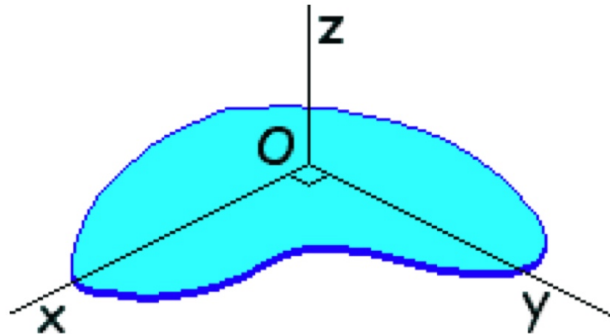


Figure 7.7: The Perpendicular axis theorem states that $I_{zz} = I_{xx} + I_{yy}$.

Proof: Note that $z = 0$ for the lamina, so from Eq. 7.4

$$I_{xx} = \int \rho y^2 dV, \quad I_{yy} = \int \rho x^2 dV$$

However, again by Eq. 7.4,

$$I_{zz} = \int \rho(x^2 + y^2) dV = I_{xx} + I_{yy}$$

.

Parallel axis theorem

Let I_{CM} denote the moment of inertia of a rigid body of mass M about an axis through its centre of mass. The moment of inertia about a parallel axis is given by

$$I = I_{CM} + Md^2 \tag{7.5}$$

where d is the perpendicular distance between the axes. (see Fig. 7.8)

Proof: Define the origin to be the centre of mass position, and the x axis to be the axis through the centre of mass. Then

$$I_{CM} = I_{xx} = \int \rho(y^2 + z^2) dV$$

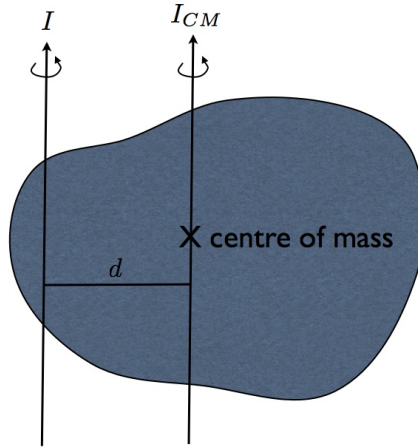


Figure 7.8: The parallel axis theorem states that $I = I_{CM} + Md^2$, where M is the mass of the rigid body.

WLOG we take the parallel axis to pass through the y axis at $y = d$, $z = 0$. We call this axis the x' axis, and further define $y' = y - d$, $z' = z$. The moment of inertia through the parallel axis is therefore

$$\begin{aligned} I &= I_{x'x'} = \int \rho(y'^2 + z'^2)dV = \int \rho((y - d)^2 + z^2)dV = \int \rho(y^2 + z^2 - 2yd + d^2)dV \\ &= I_{CM} - 2d \int \rho y dV + d^2 \int \rho dV = I_{CM} + Md^2 - 2d \int \rho y dV \end{aligned}$$

Now $\int \rho y dV$ defines the y component of the centre of mass, which is at the origin, so this vanishes, and we are left with Eq. 7.5

7.5 Kinetic energy of a rigid body about a fixed point

Since our ultimate goal is to write down Lagrangians for the motion of a rigid body, it is imperative we know how to calculate its kinetic energy. To do so we follow the same reasoning as we did before when working out the angular momentum of the rigid body. In other words, we sum up the individual contributions of the constituent particles. This gives

$$T = \sum_k \frac{1}{2} m_k |\dot{\underline{x}}^{(k)}|^2$$

But $\dot{\underline{x}}^{(k)} = \underline{w} \times \underline{x}^{(k)}$, and so

$$|\dot{\underline{x}}^{(k)}|^2 = (\underline{w} \times \underline{x}^{(k)}) \cdot (\underline{w} \times \underline{x}^{(k)}) = \underline{w} \cdot [\underline{x}^{(k)} \times (\underline{w} \times \underline{x}^{(k)})] = \frac{\underline{w} \cdot \underline{h}^{(k)}}{m_k}$$

where we have made use of the identity $\underline{a} \cdot (\underline{b} \times \underline{c}) = \underline{b} \cdot (\underline{c} \times \underline{a}) = \underline{c} \cdot (\underline{a} \times \underline{b})$. Note that we have also used the expression for angular momentum given by Eq. 7.1. It immediately follows that

$$T = \sum_k \frac{1}{2} \underline{w} \cdot \underline{h}^{(k)} = \frac{1}{2} \underline{w} \cdot \underline{h}$$

where we have used the fact that $\sum_k \underline{h}^{(k)} = \underline{h}$. Clearly this formula also holds for a solid body. Given that $\underline{h} = I\underline{w}$, we see that the rotational kinetic energy is given by

$$T = \frac{1}{2} \underline{w} \cdot I\underline{w} = \sum_{i,j} \frac{1}{2} I_{ij} w_i w_j \quad (7.6)$$

Compare these expressions with the corresponding formulae in linear mechanics

$$T = \frac{1}{2} m |\underline{v}|^2 = \frac{1}{2} \underline{v} \cdot m\underline{v} = \frac{1}{2} \underline{v} \cdot \underline{p}$$

The analogy is again evident, with the inertia tensor playing the role of inertial mass in rotational systems

7.6 Gravitational potential energy of a rigid body

Of course, to write down the full Lagrangian we also need to know about the potential energy. For the most part we will be interested in rigid bodies moving under the influence of the earth's gravitational field. Often we can work in the Galilean approximation and treat the gravitational potential energy of a particle of mass, m to be $V = -m\underline{g} \cdot \underline{x}$, where \underline{x} is the position of the particle above relative to some reference point, and \underline{g} is the constant acceleration due to gravity.

Again, we model our rigid body as a system of particles. The k th particle, of mass m_k and position $\underline{x}^{(k)}(t)$ has an external gravitational potential energy given by

$$V_k = -m_k \underline{g} \cdot \underline{x}^{(k)}$$

The particle also feels internal forces due to the other particles in the system. These are the forces that hold the system together. By N3L, if one particle exerts a force \underline{F} on another particle, then the other particle exerts an equal and opposite force $-\underline{F}$ on the first particle. When we sum up over all particles the net contribution of internal forces vanishes. As a result, for a rigid body, the total potential energy is just given by the net potential energy due to the external field. In this case that is the gravitational field, so the total potential energy is given by

$$V = \sum_k V_k = - \sum_k m_k \underline{g} \cdot \underline{x}^{(k)} = -\underline{g} \cdot \sum_k m_k \underline{x}^{(k)} = -\underline{g} \cdot (M \underline{x}_{CM}) = -M \underline{g} \cdot \underline{x}_{CM}$$

where $\underline{x}_{CM} = \frac{1}{M} \sum_k m_k \underline{x}^{(k)}$ is the centre of mass position, and M is the total mass of the system. In conclusion, the potential energy due to gravity of a rigid body is equivalent to

the potential energy of a particle with the same total mass as the rigid body, located at its centre of mass.

This approximation works well as long as we can treat the acceleration due to gravity as a constant. This is typically the case for rigid bodies that are relatively small (compared to the earth), and are not too far away from the surface of the earth.

8 The compound pendulum

As an application of what we have learnt so far, let us consider the compound pendulum. For simplicity we will restrict attention to the case where the compound pendulum is a lamina of mass, m , attached at a fixed point O to a well oiled pivot. The pivot enables the pendulum to swing freely in the plane of the lamina (see Fig. 8.1).

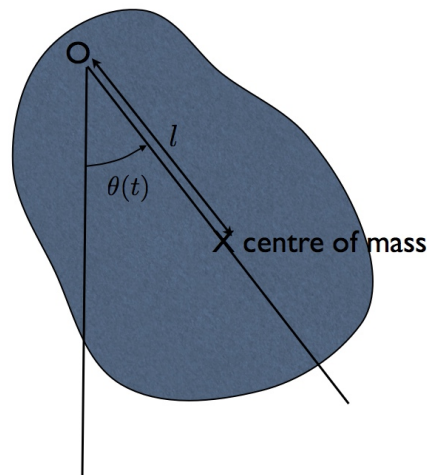


Figure 8.1: A lamina shaped compound pendulum of mass m pivoted at a fixed point O . The centre of mass is a distance l from the pivot. The length l makes an angle $\theta(t)$ with the downward vertical

The moment of inertia of the pendulum about a perpendicular axis through O is given by I , whilst the centre of mass of the pendulum lies at a fixed distance l from O . At time t we say that the line joining the pivot, O , to the centre of mass makes an angle $\theta(t)$ with the downward vertical. Now let's calculate the Lagrangian describing the motion of the pendulum.

We begin with the kinetic energy. The pendulum has angular velocity $\dot{\theta}(t)$ about the pivot. This means it has rotational kinetic energy

$$T = \frac{1}{2} I \dot{\theta}^2$$

The potential energy of the rigid body is equivalent to the potential energy of a particle of mass m at the centre of mass position. Treating the pivot, O, as our reference point, this means the pendulum has potential energy

$$V = -mgl \cos \theta$$

where $g = |g|$ is the (magnitude of) the acceleration due to gravity. The Lagrangian describing the system is therefore given by

$$\mathcal{L} = T - V = \frac{1}{2}I\dot{\theta}^2 + mgl \cos \theta$$

The Euler-Lagrange equation now gives the equation of motion for the system,

$$\frac{\partial \mathcal{L}}{\partial \theta} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{\theta}} \right) = 0 \quad \Rightarrow \quad \ddot{\theta} = -\frac{mgl}{I} \sin \theta$$

The equilibrium points correspond to the solutions $\theta(t) \equiv 0$ and $\theta(t) \equiv \pi$, where the pendulum hangs directly below or above the pivot respectively. Let us take a closer look at the stable equilibrium point at $\theta(t) \equiv 0$ by considering small oscillations about that point. This means we can take $\sin \theta \approx \theta$, and so

$$\ddot{\theta} \approx -\frac{mgl}{I} \theta$$

This corresponds to simple harmonic motion with frequency of oscillation $\omega = \sqrt{\frac{mgl}{I}}$. The solutions take the form $\theta(t) \approx A \cos \omega t + B \sin \omega t$, indicating that this equilibrium point is indeed stable.

Now consider the unstable equilibrium point at $\theta(t) \equiv \pi$. Again, we displace the pendulum slightly away from equilibrium by taking $\theta(t) = \pi + \chi(t)$, where χ is assumed to be small. The equation of motion now give

$$\ddot{\chi} \approx \frac{mgl}{I} \chi$$

which admit exponential solutions of the form $\chi \approx Ae^{\omega t} + Be^{-\omega t}$. The exponentially growing piece indicates that this equilibrium point is indeed unstable, as we suspected.

Exercise: Consider the square lamina of the previous section, and assume it is hinged at one corner, and allowed to swing freely in the plane of the lamina. Show that the frequency of oscillations about the stable equilibrium point is given by

$$\omega = \sqrt{\frac{3g}{2\sqrt{2}L}}$$

Hint: you may assume knowledge of the moments of inertia worked out in the previous section.

9 More on rigid bodies

Up until now we have studied the motion of rigid bodies about a fixed point, such as the pivot in the case of the compound pendulum. What if no fixed point is present? What then? In that case we appeal to a corollary of Euler's theorem due to Chasles. In applications of this, the centre of mass position will play a crucial role.

9.1 Chasles' theorem

The most general displacement of a rigid body is a translation plus a rotation (see Fig. 9.1).

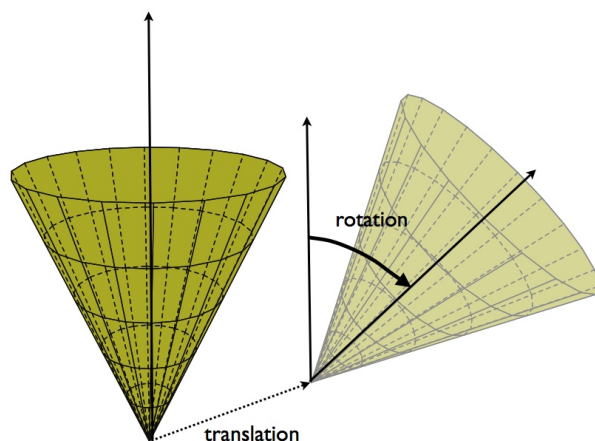


Figure 9.1: Chasles' theorem.

An important application of this is to treat the displacement of the rigid body as a translation of the centre of mass, plus a rotation about the centre of mass. Since the centre of mass is always fixed inside the rigid body, the motion of the body relative to the centre of mass can only be a rotation, by Euler's theorem.

9.2 The importance of centre of mass

We have already seen how, in many situations, the gravitational potential energy of a rigid body is equivalent to the potential energy of a particle with the same total mass as the rigid body, located at the centre of mass.

Similarly, the centre of mass often plays an important role in calculating the kinetic energy of a rigid body. The reason is that we can always split the kinetic energy into two

parts: the kinetic energy of the centre of mass, and the kinetic energy of rotation relative to the centre of mass.

To see this, take our system of N particles with the k th particle of mass m_k located at $\underline{x}^{(k)}(t)$. The centre of mass position is given by

$$\underline{x}_{CM} = \frac{1}{M} \sum_k m_k \underline{x}^{(k)}$$

where $M = \sum_k m_k$ is the total mass of the system. Now, each particle has position $\underline{s}^{(k)} = \underline{x}^{(k)} - \underline{x}_{CM}$ relative to the centre of mass.

Exercise: Prove that $\sum_k m_k \underline{s}^{(k)} = 0$.

The total kinetic energy of the system is given by

$$\begin{aligned} T &= \sum_k \frac{1}{2} m_k |\dot{\underline{x}}^{(k)}|^2 \\ &= \sum_k \frac{1}{2} m_k (\dot{\underline{s}}^{(k)} + \dot{\underline{x}}_{CM}) \cdot (\dot{\underline{s}}^{(k)} + \dot{\underline{x}}_{CM}) \\ &= \sum_k \frac{1}{2} m_k |\dot{\underline{s}}^{(k)}|^2 + \dot{\underline{x}}_{CM} \cdot \sum_k m_k \dot{\underline{s}}^{(k)} + \frac{1}{2} |\dot{\underline{x}}_{CM}|^2 \sum_k m_k \end{aligned}$$

But $\sum_k m_k \underline{s}^{(k)} = 0 \Rightarrow \sum_k m_k \dot{\underline{s}}^{(k)} = 0$, and $M = \sum_k m_k$, so

$$T = \frac{1}{2} M |\dot{\underline{x}}_{CM}|^2 + \sum_k \frac{1}{2} m_k |\dot{\underline{s}}^{(k)}|^2$$

This is precisely the statement that:

$$\text{total kinetic energy} = \begin{array}{c} \text{kinetic energy of total mass} \\ \text{at centre of mass} \end{array} + \begin{array}{c} \text{total kinetic energy of particles} \\ \text{relative to centre of mass} \end{array}$$

This fact is of crucial importance in studying the dynamics of a number of systems, not least the double pendulum which we will discuss in the next section.

10 The double pendulum

Consider two uniform rods of mass m_1 and m_2 , with lengths l_1 and l_2 respectively. The rod of mass m_1 is attached at one end to a well oiled fixed pivot, O. The rod of mass m_2 is attached at one end to a well oiled pivot at the other end of the first rod. The rods are allowed to swing freely in a plane (see Fig. 10.1). At time t we say that the rod l_1 makes an angle $\theta(t)$ with the downward vertical, whereas l_2 makes an angle $\phi(t)$. Now let's calculate the Lagrangian. The total kinetic energy is given by $T = T_1 + T_2$, where T_1 is the kinetic

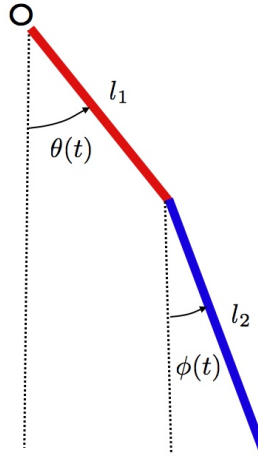


Figure 10.1: The double pendulum. The rod have mass m_1 and m_2 , with lengths l_1 and l_2 respectively

energy of l_1 and T_2 is the kinetic energy of l_2 . Now l_1 contains a fixed point at O.

Exercise: Show that the moment of inertia of l_1 about the fixed point O is given by $I_O = \frac{1}{3}m_1l_1^2$.

Given that it has angular velocity $\dot{\theta}$, its kinetic energy is given by

$$T_1 = \frac{1}{2}I_O\dot{\theta}^2 = \frac{1}{6}m_1l_1^2\dot{\theta}^2$$

The second rod is a little trickier since it is not attached to a fixed point in the system. To calculate T_2 we need to split it into two parts

$$T_2 = T_2^{CM} + T_2^{rel}$$

where T_2^{CM} is the kinetic energy of the total mass at the centre of mass, and T_2^{rel} is the kinetic energy of rotation about the centre of mass. We start with the centre of mass motion. If the centre of mass has speed v , it follows that

$$T_2^{CM} = \frac{1}{2}m_2v^2$$

To calculate v , it is convenient to introduce xy axes, with the origin at O. In relation to Fig. 10.1, x represents horizontal displacement (to the right), and y represents (downward) vertical displacement. Now the centre of mass of l_2 lies at

$$x = l_1 \sin \theta + \frac{1}{2}l_2 \sin \phi, \quad y = l_1 \cos \theta + \frac{1}{2}l_2 \cos \phi$$

The speed v is, of course, given by

$$\begin{aligned} v^2 = \dot{x}^2 + \dot{y}^2 &= (l_1\dot{\theta} \cos \theta + \frac{1}{2}l_2\dot{\phi} \cos \phi)^2 + (-l_1\dot{\theta} \sin \theta - \frac{1}{2}l_2\dot{\phi} \sin \phi)^2 \\ &= l_1^2\dot{\theta}^2 + \frac{1}{4}l_2^2\dot{\phi}^2 + l_1l_2\dot{\theta}\dot{\phi} \cos(\theta - \phi) \end{aligned} \quad (10.1)$$

So the kinetic energy of the total mass at the centre of mass would be

$$T_2^{CM} = \frac{1}{2}m_2 \left(l_1^2\dot{\theta}^2 + \frac{1}{4}l_2^2\dot{\phi}^2 + l_1l_2\dot{\theta}\dot{\phi} \cos(\theta - \phi) \right)$$

Now lets consider the rotational motion relative to the centre of mass. The angular velocity is given by $\dot{\phi}$, so the rotational kinetic energy is

$$T_2^{rel} = \frac{1}{2}I_{CM}\dot{\phi}^2$$

where I_{CM} is the moment of inertia of l_2 about its centre of mass.

Exercise: Show that $I_{CM} = \frac{1}{12}m_2l_2^2$.

It follows that

$$T_2^{rel} = \frac{1}{24}m_2l_2^2\dot{\phi}^2$$

and so the total kinetic energy of l_2 is given by

$$T_2 = T_2^{CM} + T_2^{rel} = \frac{1}{2}m_2 \left(l_1^2\dot{\theta}^2 + \frac{1}{3}l_2^2\dot{\phi}^2 + l_1l_2\dot{\theta}\dot{\phi} \cos(\theta - \phi) \right)$$

We now consider the system as a whole, with both rods. To total kinetic energy is given by

$$T = T_1 + T_2 = \frac{1}{6}m_1l_1^2\dot{\theta}^2 + \frac{1}{2}m_2 \left(l_1^2\dot{\theta}^2 + \frac{1}{3}l_2^2\dot{\phi}^2 + l_1l_2\dot{\theta}\dot{\phi} \cos(\theta - \phi) \right) \quad (10.2)$$

We now turn our attention to the potential energy of the system due to gravity, $V = V_1 + V_2$, where V_1, V_2 are the potential energies of l_1 and l_2 respectively. Again, we can find the potential energy of each rod by assuming that it is equivalent to the potential energy of a particle of the same mass at the centre of mass position. It is straight forward enough to see that

$$\begin{aligned} V_1 &= -\frac{1}{2}m_1gl_1 \cos \theta, & V_2 &= -m_2g(l_1 \cos \theta + \frac{1}{2}l_2 \cos \phi) \\ \Rightarrow V &= -\frac{1}{2}(m_1 + 2m_2)gl_1 \cos \theta - \frac{1}{2}m_2gl_2 \cos \phi \end{aligned} \quad (10.3)$$

We are now ready to write down the Lagrangian using Eqs. 10.2 and 10.3.

$$\begin{aligned} \mathcal{L} = T - V &= \frac{1}{6}(m_1 + 3m_2)l_1^2\dot{\theta}^2 + \frac{1}{6}m_2l_2^2\dot{\phi}^2 + \frac{1}{2}m_2l_1l_2\dot{\theta}\dot{\phi} \cos(\theta - \phi) \\ &\quad + \frac{1}{2}(m_1 + 2m_2)gl_1 \cos \theta + \frac{1}{2}m_2gl_2 \cos \phi \end{aligned} \quad (10.4)$$

Now we can, in principle, derive the equations of motion exactly using the Euler-Lagrange equations.

$$\frac{\partial \mathcal{L}}{\partial \theta} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{\theta}} \right) = 0, \quad \frac{\partial \mathcal{L}}{\partial \phi} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{\phi}} \right) = 0 \quad (10.5)$$

However, the resulting equations are rather messy and not particularly illuminating. The enthusiastic student may wish to derive them as an exercise. In this example, our interest lies in the motion about the stable equilibrium point. Clearly this is given by $\theta(t) \equiv 0$, $\phi(t) \equiv 0$. To study the motion near this point we simply assume that θ and ϕ are small. Note that in the example of the compound pendulum we imposed this condition at the level of the equations of motion. One can just as happily do it at the level of the Lagrangian, and derive the approximate equations of motion from that. In fact, this is usually a much easier way to do it.

To this end we note that for small θ , ϕ ,

$$\cos \theta \approx 1 - \frac{1}{2}\theta^2, \quad \cos \phi \approx 1 - \frac{1}{2}\phi^2, \quad \cos(\theta - \phi) \approx 1 - \frac{1}{2}(\theta - \phi)^2$$

Neglecting all terms beyond quadratic order in the Lagrangian, we find that it is given by

$$\begin{aligned} \mathcal{L} \approx & \frac{1}{6}(m_1 + 3m_2)l_1^2\dot{\theta}^2 + \frac{1}{6}m_2l_2^2\dot{\phi}^2 + \frac{1}{2}m_2l_1l_2\dot{\theta}\dot{\phi} - \frac{1}{4}(m_1 + 2m_2)gl_1\theta^2 - \frac{1}{4}m_2gl_2\phi^2 \\ & + \frac{1}{2}(m_1 + 2m_2)gl_1 + \frac{1}{2}m_2gl_2 \end{aligned} \quad (10.6)$$

The last two terms in the Lagrangian are constant and therefore do not contribute to the dynamics. The Euler-Lagrange Eqs. 10.5 are now very easy to derive. They give

$$\frac{1}{3}(m_1 + 3m_2)l_1^2\ddot{\theta} + \frac{1}{2}m_2l_1l_2\ddot{\phi} + \frac{1}{2}(m_1 + 2m_2)gl_1\theta = 0 \quad (10.7)$$

$$\frac{1}{2}m_2l_1l_2\ddot{\theta} + \frac{1}{3}m_2l_2^2\ddot{\phi} + \frac{1}{2}m_2gl_2\phi = 0 \quad (10.8)$$

It is instructive to write these equations in matrix form

$$\begin{bmatrix} \frac{1}{3}(m_1 + 3m_2)l_1^2 & \frac{1}{2}m_2l_1l_2 \\ \frac{1}{2}m_2l_1l_2 & \frac{1}{3}m_2l_2^2 \end{bmatrix} \begin{pmatrix} \ddot{\theta} \\ \ddot{\phi} \end{pmatrix} + \begin{bmatrix} \frac{1}{2}(m_1 + 2m_2)gl_1 & 0 \\ 0 & \frac{1}{2}m_2gl_2 \end{bmatrix} \begin{pmatrix} \theta \\ \phi \end{pmatrix} = 0 \quad (10.9)$$

This clearly has the form

$$K\underline{\ddot{\psi}} + M\underline{\psi} = 0 \quad (10.10)$$

where $\underline{\psi} = \begin{pmatrix} \theta \\ \phi \end{pmatrix}$ and

$$K = \begin{bmatrix} \frac{1}{3}(m_1 + 3m_2)l_1^2 & \frac{1}{2}m_2l_1l_2 \\ \frac{1}{2}m_2l_1l_2 & \frac{1}{3}m_2l_2^2 \end{bmatrix}, \quad M = \begin{bmatrix} \frac{1}{2}(m_1 + 2m_2)gl_1 & 0 \\ 0 & \frac{1}{2}m_2gl_2 \end{bmatrix}$$

10.1 Normal modes and normal frequencies

As it stands this system of coupled differential is difficult to proceed, we look for the *normal modes* of the system. These are combinations of θ and ϕ that oscillate with a well defined frequency, known as the *normal frequencies*.

How do we find the normal modes and normal frequencies? Well, we begin by taking our matrix equation 10.10, and writing it as

$$\ddot{\underline{\psi}} + A\underline{\psi} = 0 \quad (10.11)$$

where $A = K^{-1}M$. We can certainly do this since $\det K \neq 0$, and so the inverse of K is well defined.

Now let \underline{u}_1 and \underline{u}_2 denote two linearly independent eigenvectors of A with eigenvalues λ_1 and λ_2 respectively, so that $A\underline{u}_1 = \lambda_1\underline{u}_1$ and $A\underline{u}_2 = \lambda_2\underline{u}_2$. The vector $\underline{\psi}$ can be written in terms of these eigenvectors as

$$\underline{\psi} = \Theta(t)\underline{u}_1 + \Phi(t)\underline{u}_2 \quad (10.12)$$

where $\Theta(t)$ and $\Phi(t)$ are linear combinations of $\theta(t)$ and $\phi(t)$.

Exercise: Show that

$$\Theta(t) = \frac{(|\underline{u}_2|^2\underline{u}_1 - (\underline{u}_1 \cdot \underline{u}_2)\underline{u}_2) \cdot \underline{\psi}}{|\underline{u}_1|^2|\underline{u}_2|^2 - (\underline{u}_1 \cdot \underline{u}_2)^2}, \quad \Phi(t) = \frac{(|\underline{u}_1|^2\underline{u}_2 - (\underline{u}_1 \cdot \underline{u}_2)\underline{u}_1) \cdot \underline{\psi}}{|\underline{u}_1|^2|\underline{u}_2|^2 - (\underline{u}_1 \cdot \underline{u}_2)^2},$$

Hint: dot both sides of Eq. 10.12 with \underline{u}_1 and \underline{u}_2 and solve for $\Theta(t)$ and $\Phi(t)$.

Now for the nice bit. We plug our expression for $\underline{\psi}$ (Eq. 10.12) into the matrix equation 10.11. Since $A\underline{u}_1 = \lambda_1\underline{u}_1$ and $A\underline{u}_2 = \lambda_2\underline{u}_2$, the result is

$$(\ddot{\Theta} + \lambda_1\Theta)\underline{u}_1 + (\ddot{\Phi} + \lambda_2\Phi)\underline{u}_2 = 0$$

Since \underline{u}_1 and \underline{u}_2 are linearly independent, it follows that

$$\ddot{\Theta} = -\lambda_1\Theta, \quad \ddot{\Phi} = -\lambda_2\Phi$$

Therefore Θ satisfies simple harmonic motion with frequency $\omega_\Theta = \sqrt{\lambda_1}$, whereas Φ satisfies simple harmonic motion with frequency $\omega_\Phi = \sqrt{\lambda_2}$. We say that $\Theta(t)$ and $\Phi(t)$ are the *normal modes* with corresponding *normal frequencies* ω_Θ and ω_Φ .

In a given problem then, the task is to find the eigenvalues and eigenvectors of the matrix A . This is straight forward. If I is the identity matrix⁴, the eigenvalues satisfy the equation

$$\begin{aligned} \det(A - \lambda I) = 0 & \Rightarrow \det(K^{-1}M - \lambda I) = 0 & \Rightarrow \det(K^{-1}(M - \lambda K)) = 0 \\ & \Rightarrow \det(K^{-1}) \det(M - \lambda K) = 0 & \Rightarrow \det(M - \lambda K) = 0 \end{aligned}$$

⁴Not to be confused with inertia tensor.

since $\det(K^{-1}) \neq 0$. Once we have the eigenvalues, λ , we find the corresponding eigenvectors, \underline{u} , by solving

$$A\underline{u} = \lambda\underline{u}, \quad \Rightarrow \quad M\underline{u} = \lambda K\underline{u}$$

Exercise: Consider the case where $m_1 = 6m$, $m_2 = m$ and $l_1 = l$, $l_2 = 3l$. Show that A has eigenvalues

$$\lambda_1 = \frac{2g}{l}, \quad \lambda_2 = \frac{4g}{9l}$$

with corresponding eigenvectors

$$\underline{u}_1 = \begin{pmatrix} -3 \\ 2 \end{pmatrix}, \quad \underline{u}_2 = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$$

In the example given in the exercise, we can use that fact that $|\underline{u}_1|^2 = 13$, $|\underline{u}_2|^2 = 7$ and $\underline{u}_1 \cdot \underline{u}_2 = 5$ to show that

$$\Theta(t) = \frac{(|\underline{u}_2|^2 \underline{u}_1 - (\underline{u}_1 \cdot \underline{u}_2) \underline{u}_2) \cdot \underline{\psi}}{|\underline{u}_1|^2 |\underline{u}_2|^2 - (\underline{u}_1 \cdot \underline{u}_2)^2} = \frac{1}{14} \begin{pmatrix} -4 \\ 1 \end{pmatrix} \cdot \underline{\psi} = \frac{1}{14} (\phi(t) - 4\theta(t))$$

In conclusion we see that the normal mode $\Theta \propto \phi - 4\theta$ has normal frequency $\omega_\Theta = \sqrt{\frac{2g}{l}}$.

Exercise: Calculate the normal mode Φ and its normal frequency of oscillation.

In summary a system with two or more oscillating degrees of freedom, like the double pendulum, will typically contain special modes that oscillate with a well defined frequency. These are the normal modes and the normal frequencies of oscillation.

11 Constraints

When we analysed the motion of planets around the Sun, we treated the heavy objects as particles, and implicitly assumed that they were, in principle at least, able to roam freely in space under the action of the conservative gravitational force. However, many dynamical systems involve constraints, and corresponding *constraint forces*. For example, a simple pendulum consisting of a mass, m on the end of a light rope of fixed length, l is constrained by the length of the rope. The constraint force corresponds to the tension of the rope, which is initially unknown. We can, of course, cope with these constraints by imposing them at the level of Newton's laws and introducing the necessary constraint forces as unknowns to be determined. But it is a messy business, especially when there are a number of nasty, complicated constraints affecting the system. As we shall see in this section, for the class of constraints known as *holonomic constraints*, the Lagrangian formulation makes life an awful lot easier.

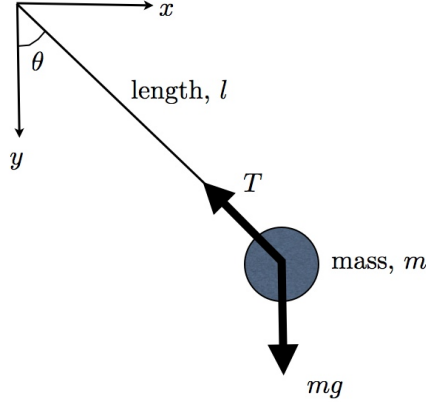


Figure 11.1: The simple pendulum: a mass m is attached to a rope of fixed length l .

Let us begin with a tame example: the simple pendulum described in the previous paragraph (see Fig. 11.1). Since the rope has fixed length, l , we need to introduce the tension of the rope, T , as a constraint force. Working in Cartesian coordinates, as shown in Fig. 11.1, we note that the mass has position $(x(t), y(t))$, where $x^2 + y^2 = l^2$. Applying N2L along each direction we obtain

$$m\ddot{x} = -Tx/l, \quad m\ddot{y} = mg - Ty/l \quad (11.1)$$

We can impose the constraint by setting $x = l \sin \theta$, $y = l \cos \theta$, thereby obtaining

$$T = ml\dot{\theta}^2 + mg \cos \theta, \quad \ddot{\theta} = -\frac{g}{l} \sin \theta \quad (11.2)$$

By multiplying the latter of these equations by $\dot{\theta}$, we can show that

$$\dot{\theta}^2 = \frac{2g}{l} \cos \theta + c$$

where c is constant. It follows that the tension satisfies

$$T = 3mg \cos \theta + mlc$$

The simple pendulum might seem fairly straightforward to solve using Newtonian methods, and indeed it is. However, it is worth noting that a chunk of the analysis went into deriving the tension. This kind of thing can be really problematic in more complicated constrained systems. If we don't care about the value of the constraint forces then it is far easier to make use of Lagrangian mechanics, as we shall soon see. Of course, sometimes we do care about

the values of constraint forces, as illustrated in the exercise below.

Exercise: Suppose that the mass initially hangs vertically downwards, but is given a horizontal velocity v . Show that the rope becomes slack if $2gl < v^2 < 5gl$.

Hint: this gives us the initial condition $\theta(0) = 0$, $\dot{\theta}(0) = v/l$. Note that the rope becomes slack when $T = 0$. You must also worry about whether or not the mass reaches a maximum and turns around.

11.1 Holonomic constraints

In the previous example, we can think of the system as being described by coordinates $(x(t), y(t))$, subject to the constraint $x^2 + y^2 - l^2 = 0$. This is an example of an *holonomic constraint*.

More generally, consider a dynamical system described by $x_A(t)$ where $A = 1, \dots, N$. An holonomic constraint is one that can be written in the form

$$F(x_A, t) = 0$$

Suppose there are $N - n$ such constraints

$$F_\alpha(x_A, t) = 0, \quad \alpha = n + 1, \dots, N$$

How do we deal with these constraints in the Lagrangian formulation? Well, the trick is to introduce $N - n$ new variables, λ_α , called *Lagrange multipliers*. Given the unconstrained Lagrangian $\mathcal{L}(x_A, \dot{x}_A)$ we introduce a new Lagrangian

$$\mathcal{L}' = \mathcal{L}(x_A, \dot{x}_A) + \sum_{\alpha=n+1}^N \lambda_\alpha F_\alpha(x_A, t) \quad (11.3)$$

If we treat the λ_α like a new coordinate, then it is clear that the Euler-Lagrange equations for these new coordinates are given by

$$\frac{\partial \mathcal{L}'}{\partial \lambda_\alpha} = 0 \quad \Rightarrow \quad F_\alpha(x_A, t) = 0$$

In other words, we recover the constraints. On the other hand, the Euler-Lagrange equations for the x_A yield

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}_A} \right) - \frac{\partial \mathcal{L}}{\partial x_A} = \sum_{\alpha=n+1}^N \lambda_\alpha \frac{\partial F_\alpha}{\partial x_A}$$

The left hand side corresponds to the equations of motion for the unconstrained system, whereas the right hand side encodes the role played by the constraint forces in the system.

Indeed, in the previous example, where we had the constraint, $F = x^2 + y^2 - l^2 = 0$, we can describe the system using the modified Lagrangian

$$\mathcal{L}' = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + mgy + \lambda(x^2 + y^2 - l^2)$$

We obtain the following equations of motion for x , y

$$m\ddot{x} = 2\lambda x, \quad m\ddot{y} = mg + 2\lambda y$$

and of course the constraint equation $x^2 + y^2 = l^2$. By comparing with Eq. 11.1, we see that the Lagrange multiplier is proportional to the tension, $\lambda = -T/2l$. The cute thing here is that we didn't really need to worry about the nature of the constraint forces—their properties just fall out of the Euler-Lagrange equations automatically.

Avoiding constraints in constrained systems!

Things are sometimes even cuter, especially if one really doesn't care about the details of the constraint forces. Indeed, there is a particular choice of generalised coordinates that enables us to avoid the constraints altogether. How can this be? Well, the constraints haven't really gone away. It's just that a clever choice of generalised coordinates is able to encode information about those constraints. The point is that the number of degrees of freedom has been reduced from N down to n by imposing $N - n$ constraints. So you only really need n coordinates to describe the system in full. In the pendulum example, the appropriate generalised coordinate would be θ , where we identify $x = l \sin \theta$, $y = l \cos \theta$.

To see how this works in the general case, we introduce generalised coordinates q_A where

$$\begin{aligned} q_i &= q_i(x_A, t), & i &= 1, \dots, n, \\ q_\alpha &= F_\alpha(x_A, t), & \alpha &= n + 1, \dots, N \end{aligned}$$

In other words, we have identified $N - n$ coordinates with the $N - n$ constraints, leaving a total of n non-vanishing coordinates, q_i . Clearly then, the system can be solved in terms of the n non-vanishing generalised coordinates q_i ,

$$x_A = x_A(q_1, \dots, q_n)$$

Indeed, given the modified Lagrangian

$$\mathcal{L}' = L(q_A, \dot{q}_A, t) + \sum_{\alpha=n+1}^N \lambda_\alpha F_\alpha$$

we see that the q_i equation of motion is

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) - \frac{\partial \mathcal{L}}{\partial q_i} = \sum_{\alpha=n+1}^N \lambda_\alpha \frac{\partial F_\alpha}{\partial q_i}$$

But $\frac{\partial F_\alpha}{\partial q_i} = 0$ by definition and so the dynamics of the q_i only depends on the unconstrained Lagrangian—it doesn't care about the Lagrange multipliers!

To see this in action we return to the simple pendulum. Given that we start off with two degrees of freedom, x and y , and one constraint $x^2 + y^2 = l^2$, we only need to introduce

one generalised coordinate, θ , in such a way as to respect the constraint. We therefore set $x = l \sin \theta$, $y = l \cos \theta$. Working with this generalised coordinate, there is no longer be any reason to worry about the constraint. To see this, note that the unconstrained Lagrangian is given by

$$\mathcal{L} = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + mgy = \frac{1}{2}ml^2\dot{\theta}^2 + mgl \cos \theta$$

which yields the correct equation of motion for θ ,

$$\ddot{\theta} = -\frac{g}{l} \sin \theta$$

Of course, we haven't calculated the tension here, but if that is not a concern, we see how this method involves far less work.

In summary, holonomic constraints can be dealt with very easily using the Lagrangian formulation. They can be included explicitly using the method of Lagrange multipliers, or ignored altogether if we are clever enough to identify the key degrees of freedom that describe the system. However, in the interests of health and safety, a general rule of thumb would be: if in doubt, use a Lagrange multiplier.

11.2 Non-holonomic constraints

As stated earlier, a holonomic constraint is one which can be written as $F(x_A, t) = 0$. Any other kind of constraint is called a non-holonomic constraint. Examples include:

Inequalities: a ball rolling off the surface of a solid sphere of radius, R , cannot penetrate the sphere, and is therefore constrained to move along a path $(x(t), y(t), z(t))$ where $x^2 + y^2 + z^2 \geq R^2$.

Velocity dependent constraints: a coin rolling down a slope without slipping is described using 4 coordinates, $(x(t), y(t), \theta(t), \phi(t))$, where x , y denote the cartesian coordinates of the point of contact on the slope; θ denotes the angle the coin makes with the path of steepest descent; and ϕ denotes the angle a fixed point on the coin makes with the vertical (see Fig. 11.2). If the coin has radius R then the velocity of the rim is $R\dot{\phi}$. It follows that we have the velocity dependent constraint

$$\dot{x} = R\dot{\phi} \cos \theta, \quad \dot{y} = R\dot{\phi} \sin \theta$$

This *cannot* be integrated to given constraints of the form $F(x, y, \theta, \phi) = 0$

Note that velocity dependent constraints can be absorbed into the Lagrangian formulation using Lagrange multipliers, whereas the inequalities cannot.

11.3 Other tricks involving Lagrangians (*not examinable*)

Higher derivative Lagrangians

So far we have restricted attention to Lagrangian that only depend on position and velocity. However, occasionally one comes across Lagrangian's with more derivatives, for example

$$\mathcal{L}(x_i, \dot{x}_i, \ddot{x}_i) \tag{11.4}$$

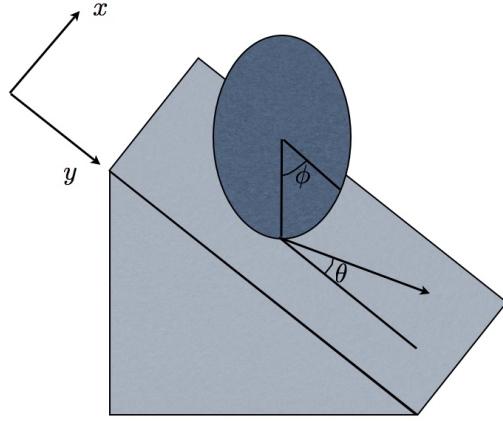


Figure 11.2: Nonholonomic constraints: a coin rolling down a slope without slipping.

If there are lots of higher derivatives, with a complicated structure this can make it quite difficult to extract the equations of motion. However, we can get round this by making use of Lagrange multipliers and introducing an enlarged system of generalised coordinates. In the example given we introduce generalised coordinates, $q_A = (x_i, y_i)$, where

$$y_i = \dot{x}_i$$

Now we can write our dynamical system in terms of the Lagrangian containing only first derivatives

$$\mathcal{L}' = \mathcal{L}(x_i, y_i, \dot{y}_i) + \sum_i \lambda_i (\dot{x}_i - y_i) \quad (11.5)$$

The λ_i equation gives the constraint $y_i = \dot{x}_i$, whereas the x_i and y_i equations give

$$\begin{aligned} \frac{\partial \mathcal{L}'}{\partial x_i} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}'}{\partial \dot{x}_i} \right) = 0 &\quad \Rightarrow \quad \frac{\partial \mathcal{L}}{\partial x_i} = \dot{\lambda}_i \\ \frac{\partial \mathcal{L}'}{\partial y_i} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}'}{\partial \dot{y}_i} \right) = 0 &\quad \Rightarrow \quad \frac{\partial \mathcal{L}}{\partial y_i} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{y}_i} \right) = \lambda_i \end{aligned}$$

Combining these equations yields the higher derivative version of the Euler-Lagrange equations found previously,

$$\frac{\partial \mathcal{L}}{\partial x_i} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}_i} \right) + \frac{d^2}{dt^2} \left(\frac{\partial \mathcal{L}}{\partial \ddot{x}_i} \right) = 0$$

The method generalises to systems with even higher numbers of derivatives. In the example given, the acceleration dependent Lagrangian is equivalent to a Lagrangian depending only on position and velocity, but with twice as many degrees of freedom. In this way, higher order derivatives can be thought of as corresponding to extra degrees of freedom in the system.

Higher/fractional power Lagrangians

Occasionally we come across a Lagrangian containing a higher power term, or indeed a fractional power term of the form

$$\mathcal{L} = [W(t, x_i(t), \dot{x}_i(t))]^n + \dots \quad (11.6)$$

These can be awkward to vary, especially in theories of gravity where there may be an underlying summation over tensor indices (we have not shown these here for simplicity). However we can adopt a neat trick to simplify our analysis in this case. We introduce an *auxiliary field*, $A(t)$, and write down the modified Lagrangian

$$\mathcal{L}' = nA^{n-1}W(t, x_i(t), \dot{x}_i(t)) - (n-1)A^n + \dots \quad (11.7)$$

which is equivalent to the original Lagrangian. The nice thing about this Lagrangian is that it is linear in W which can often make the analysis much simpler. To see that the two Lagrangians are equivalent, note that the A equation yields $A = W$. Substituting this back into Eq. 11.7 recovers the original Lagrangian Eq. 11.6.

A nice example of this trick is in rewriting the action for a point particle in General Relativity. A point particle of mass m has action

$$S[x^\mu] = -m \int d\tau \sqrt{-\sum_{\mu,\nu} g_{\mu\nu} \dot{x}^\mu \dot{x}^\nu} \quad (11.8)$$

where τ is proper time, $g_{\mu\nu}(x(\tau))$ is the spacetime metric with spacetime indices μ and ν , evaluated on the particle's trajectory $x^\mu(\tau)$. We can introduce an auxiliary field $A(\tau)$, and write down the equivalent action

$$S'[x^\mu] = -\frac{m}{2} \int d\tau \frac{1}{\sqrt{A(\tau)}} \left(-\sum_{\mu,\nu} g_{\mu\nu} \dot{x}^\mu \dot{x}^\nu \right) + \sqrt{A(\tau)} \quad (11.9)$$

If we redefine our auxiliary field by letting $A(\tau) = m^2 \eta^2(\tau)$, the new action can be written in the form

$$S'[x^\mu] = \frac{1}{2} \int d\tau \frac{1}{\eta(\tau)} \left(\sum_{\mu,\nu} g_{\mu\nu} \dot{x}^\mu \dot{x}^\nu \right) - m^2 \eta(\tau) \quad (11.10)$$

The action is now completely void of square roots and very easy to work with. In string theory, a similar procedure takes us from the the Nambu-Goto action for the fundamental string to the much more manageable Polyakov action.

12 Hamilton's principle –again

We now turn our attention to the fascinating area of Hamiltonians as a complementary means to understand the symmetries of systems and to derive their equations of motion. As this part of the course develops we will see that in some situations it is preferable to analyse the system through a Lagrangian whereas in other situations the Hamiltonian is the physicists preferred tool of choice.

In chapter 5 Tony introduced you to the formal statement of Hamilton's Principle:

Hamilton's principle or The Principle of Least Action

For a system described by generalised coordinates, q_A , the correct path of motion $q_A(t)$ between the initial state $q_A(t_0)$ at time t_0 and the final state $q_A(t_1)$ at time t_1 corresponds to a stationary path of the action

$$S = \int_{t_0}^{t_1} \mathcal{L} dt$$

where $\mathcal{L} = \mathcal{L}(t, q_A, \dot{q}_A)$ is the Lagrangian describing the system.

The stationary path is found by solving the Euler-Lagrange equations,

$$\frac{\partial \mathcal{L}}{\partial q_A} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_A} \right) = 0. \quad (12.1)$$

In the symmetries course you will come across some beautiful theorems associated with the action principle, known as Noether's Theorem after the German mathematician Emmy Noether. The theorems connect symmetries of a system with physically conserved quantities. Here we consider just two types of constants of the motion:

1. Conserved momentum like quantities
2. Conserved psuedo-energy

12.1 Constant Generalised Momentum

If $\mathcal{L} = \mathcal{L}(t, q_A, \dot{q}_A)$ is such that for a particular generalised coordinate, say q_j the Lagrangian does not depend on q_j , i.e. if

$$\frac{\partial \mathcal{L}}{\partial q_j} = 0 \quad \text{some } j \quad (12.2)$$

then if we define the **generalised momentum**

$$p_j \equiv \frac{\partial \mathcal{L}}{\partial \dot{q}_j}, \quad (12.3)$$

then from the Euler Lagrange equation for the j th coordinate we have

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_j} \right) = 0 \quad (12.4)$$

hence

$$p_j = \text{const} \quad (12.5)$$

for this j . The coordinate q_j is then called an ignorable or cyclic coordinate. In particular if q_j is a linear coordinate it follows that p_j is a linear momentum, where as if q_j is an angular coordinate then p_j corresponds to an angular momentum.

Let us look at an example to see the usefulness of ignorable coordinates. Fig 12.1 shows two bodies of mass M and m . The larger body, mass M is moving in the positive x direction and located at position x , the smaller body is forced to oscillate with frequency ω , and any given moment t is a distance y from the the centre of the block M given by

$$y(t) = a \cos(\omega t), \text{ a const.}$$

We want to solve for $x(t)$ using the Euler-Lagrange equations. The Lagrangian is obtained

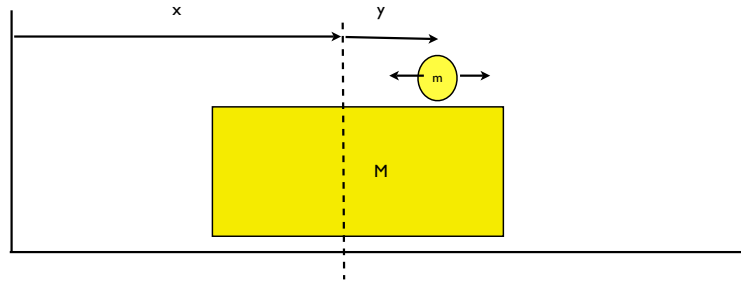


Figure 12.1: Example of Generalised Coordinate.

from the kinetic energy and potential energy as usual. In this case we have

$$\text{Kinetic Energy} : T = \frac{1}{2}M\dot{x}^2 + \frac{1}{2}M(\dot{x} + \dot{y})^2 \quad (12.6)$$

$$\text{Potential Energy} : V = 0 \quad (12.7)$$

Using $L = T - V$, expanding the Kinetic Energy and substituting in for $y(t)$ and $\dot{y} = -a\omega \sin(\omega t)$ we have

$$L = \frac{1}{2}(M + m)\dot{x}^2 - ma\omega\dot{x} \sin(\omega t) + \frac{1}{2}ma^2\omega^2 \sin^2(\omega t)$$

Looking closely at L we see that it is independent of the coordinate x which implies that x is an ignorable constant. Therefore the associated generalised momentum p_x is constant, i.e.

$$p_x = \frac{\partial L}{\partial \dot{x}} = (M + m)\dot{x} - ma\omega \sin(\omega t) = C = \text{const.} \quad (12.8)$$

The constant C can be determined by the initial conditions. For example, imagine that at $t = 0$, the mass M is moving with speed $\dot{x} = v$, then (12.8) gives

$$C = (M + m)v$$

hence the same equation for general t gives

$$\dot{x} = v + \frac{am\omega}{M + m} \sin \omega t$$

hence upon integration we obtained the general solution

$$x(t) = vt - \frac{am}{M + m} \cos \omega t + \text{const.}$$

The first term on the RHS describes the steady motion of the block and the second term describes the added wobble due to the mass m . Note that for a very small mass $m \ll M$, the amplitude of the wobble is small, whereas for the case where the mass is large, $m \gg M$ we see the amplitude is a .

12.2 Constant Pseudo-energy – known as the Hamiltonian

If the Lagrangian does not depend explicitly on time t , i.e. $L = L(q_i, \dot{q}_i)$ only such that $\frac{\partial L}{\partial t} = 0$, then we can define a quantity E such that

$$E \equiv \sum_{i=1}^N \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} - L(q_i, \dot{q}_i) = \text{const}$$

Proof

For a general Lagrangian $L(t, q_i, \dot{q}_i)$ we have applying the chain rule

$$\frac{dL}{dt} = \frac{\partial L}{\partial t} + \sum_i \frac{\partial L}{\partial q_i} \dot{q}_i + \sum_i \frac{\partial L}{\partial \dot{q}_i} \ddot{q}_i. \quad (12.9)$$

Now if we consider the function

$$h(t, (q_i, \dot{q}_i)) = \sum_{i=1}^N \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} - L(t, q_i, \dot{q}_i) \quad (12.10)$$

then

$$\frac{dh}{dt} = \sum_i \frac{\partial L}{\partial \dot{q}_i} \ddot{q}_i + \sum_i \dot{q}_i \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) - \frac{dL}{dt}.$$

Substituting in for $\frac{dL}{dt}$ from (12.9) we obtain

$$\frac{dh}{dt} = \sum_i \frac{\partial L}{\partial \dot{q}_i} \ddot{q}_i + \sum_i \dot{q}_i \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) - \left(\frac{\partial L}{\partial t} + \sum_i \frac{\partial L}{\partial q_i} \dot{q}_i + \sum_i \frac{\partial L}{\partial \dot{q}_i} \ddot{q}_i \right). \quad (12.11)$$

The first and last terms on the RHS of (12.11) cancel as do the second and next to last terms once we have made use of the Euler-Lagrange equation. This leaves us with the simple yet elegant result

$$\frac{dh}{dt} = -\frac{\partial L}{\partial t} \quad (12.12)$$

implying that if $\frac{\partial L}{\partial t} = 0$, then $h = \text{const.}$ This is known as the Hamiltonian and we will see a lot more of it as the course proceeds. In particular we will see how it can be related to conservation of energy. Before doing that we will consider an example where the Hamiltonian is conserved. The example consists of a trolley of mass M moving in the positive x direction with a pendulum of length l and bob mass m hanging freely from it, beautifully illustrated in Fig. 12.2

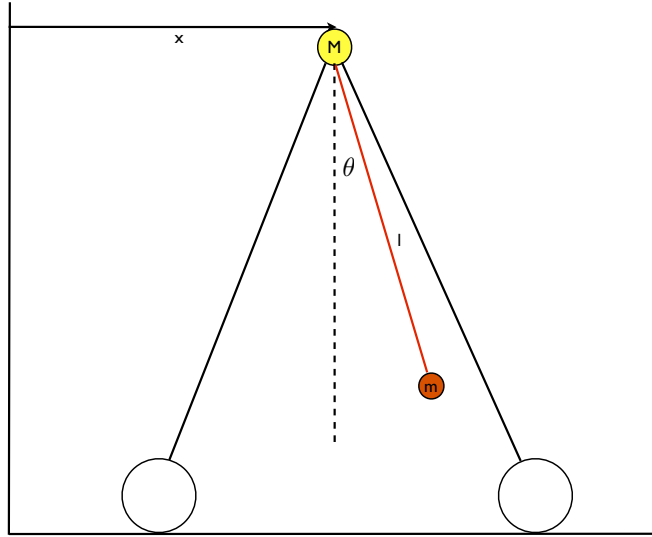


Figure 12.2: Trolley moving in x direction with pendulum hanging freely from it.

At a given instant of time the pendulum is at an angle θ to the vertical. Considering the kinetic energy of the system we consider the trolley and pendulum separately. The trolley is straightforward and has kinetic energy $\frac{1}{2}M\dot{x}^2$. We have to be more careful with the pendulum bob and take into account the fact that it is moving in an arc in the $x - y$ plane. A little bit of thought should tell you that that x and y velocity components for the bob are given by

$$v_x = \dot{x} + l\dot{\theta} \cos \theta \quad (12.13)$$

$$v_y = l\dot{\theta} \sin \theta, \quad (12.14)$$

therefore the kinetic energy of the bob is given by $\frac{1}{2}m(v_x^2 + v_y^2)$. The potential energy of the bob is given by $V = -mgl \cos \theta$ (where we have assumed $V = 0$ corresponds to the centre of mass of the trolley). The Lagrangian for the system now follows

$$L = T - V = \frac{1}{2}(M + m)\dot{x}^2 + ml\dot{x}\dot{\theta} \cos \theta + \frac{1}{2}ml^2\dot{\theta}^2 + mgl \cos \theta. \quad (12.15)$$

Now we have the Lagrangian we immediately start looking for ignorable coordinates because we know there are then associated constants of the motion with this system. Eqn. (12.15) is manifestly independent of the coordinate x , i.e. $\frac{\partial L}{\partial x} = 0$ implying that x is an ignorable coordinate. The associated constant generalised momentum p_x is given from (12.15) by

$$p_x = \frac{\partial L}{\partial \dot{x}} = (M + m)\dot{x} + ml\dot{\theta} \cos \theta = \text{const} \quad (12.16)$$

For completeness we also see that

$$p_\theta = \frac{\partial L}{\partial \dot{\theta}} = ml\dot{x} \cos \theta + ml^2\dot{\theta} \quad (12.17)$$

Further analysis of (12.15) shows that there is no explicit t dependence, i.e. $\frac{\partial L}{\partial t} = 0$. As we have just seen in (12.12), this implies that the quantity which we now suggestively call E is obtained from (12.10) as

$$E = \dot{x} \frac{\partial L}{\partial \dot{x}} + \dot{\theta} \frac{\partial L}{\partial \dot{\theta}} - L = \text{const}. \quad (12.18)$$

Specifically, using (12.15)-(12.17) we obtain

$$E = \frac{1}{2}(M + m)\dot{x}^2 + ml\dot{x}\dot{\theta} \cos \theta + \frac{1}{2}ml^2\dot{\theta}^2 - mgl \cos \theta = \text{const}. \quad (12.19)$$

From (12.19) we see that $E = T + V$, the sum of the kinetic and potential energy terms. In other words it is the energy of the system. This need not always be the case, so we will digress for a moment to establish what it is that has happened here which has meant that we can energy conservation when the Lagrangian is independent of time. Basically it has happened because the kinetic energy T is quadratic in $\dot{\theta}$, \dot{x} where as the potential V is independent of them. Let us think about this in a bit more detail for a more general case. We can always write the Lagrangian $L(t, q_i, \dot{q}_i)$ as

$$L(t, q_i, \dot{q}_i) = L_0(\dot{q}_i) + L_1(\dot{q}_i) + L_2(\dot{q}_i)$$

where L_2 is quadratic in \dot{q}_i so that $\dot{q}_i \frac{\partial L_2}{\partial \dot{q}_i} = 2L_2$, L_1 is linear in \dot{q}_i so that $\dot{q}_i \frac{\partial L_1}{\partial \dot{q}_i} = L_1$ and L_0 is independent of \dot{q}_i so that $\dot{q}_i \frac{\partial L_0}{\partial \dot{q}_i} = 0$. Substituting into (12.10) we find

$$\begin{aligned} h(t, q_i, \dot{q}_i) &= \sum_{i=1}^N \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} - L(t, q_i, \dot{q}_i) \\ &= 2L_2 + L_1 - (L_0 + L_1 + L_2) \\ &= L_2 - L_0 \end{aligned} \quad (12.20)$$

Now we often have (but not always) that the Lagrangian is given by $L = T - V$ where the kinetic energy T is quadratic in \dot{q}_i and the potential V is independent of \dot{q}_i . In that case (which of course is the one for the trolley above) we then have

$$h = T + V$$

and so the Hamiltonian corresponds to the total energy $E = T + V$.

returning to the case of our supermarket trolley/pendulum system (you should try it out next time you are in Sainsburys!) let us imagine the initial condition to be that at $t = 0$ we have $\theta = 0$, $\dot{x} = 0$ but $l\dot{\theta} = v$. In other words initially the trolley is static but the pendulum is moving with angular velocity v . *How far up does the pendulum swing?* From (12.16) we have, recalling $p_x = mv$ at $t = 0$ when $\theta = 0$, $\dot{x} = 0$,

$$p_x = mv = (M + m)\dot{x} + ml\dot{\theta} \cos \theta = \text{const}$$

hence

$$\dot{x} = \frac{m(v - l\dot{\theta} \cos \theta)}{M + m}. \quad (12.21)$$

Similarly from (12.19) we have (using the initial conditions)

$$E = \frac{1}{2}mv^2 - mgl = \text{const}, \quad (12.22)$$

Substituting for the energy and velocity \dot{x} into (12.19) we obtain the rather complicated looking equation

$$\frac{1}{2}(M+m) \left(\frac{m(v - l\dot{\theta} \cos \theta)}{M + m} \right)^2 + m^2 l \dot{\theta} \cos \theta \left(\frac{(v - l\dot{\theta} \cos \theta)}{M + m} \right) + \frac{1}{2}ml^2\dot{\theta}^2 - mgl \cos \theta = \frac{1}{2}mv^2 - mgl \quad (12.23)$$

Things are not as bad as they may seem. The pendulum stops for an instance when it has reached its maximum height, hence the maximum swing angle $\theta = \theta_0$ corresponds to when $\dot{\theta} = 0$. Substituting in (12.23) and rearranging we obtain the equation for θ_0 as

$$-\frac{1}{2} \frac{Mv^2}{M + m} = gl(\cos \theta_0 - 1).$$

A related question we can now answer is *what is the size of v required so that the pendulum bob reaches the vertical, i.e. so that $\theta_0 = \pi$?* Straightforward substitution gives

$$v^2 = \frac{4gl(M + m)}{M}.$$

Does the result make sense? Consider the two obvious limits: first for $M \gg m$ we have $v^2 \simeq 4gl$ which is the usual result obtained for a fixed pendulum which it effectively is as the trolley would barely be moving. On the other hand for the case $M \ll m$ we see that $v^2 \sim \frac{4m}{M}gl$ indicating that the required velocity becomes large as expected.

12.3 Resonance

The pendulum we have been considering above has been at a fixed pivot point in the y axis, say $(x,y=0)$. Imagine that we drop this restriction and now force the pivot to move such that it satisfies

$$x = F(t), \quad y = G(t)$$

for some arbitrary driving terms $F(t)$ and $G(t)$. This means there is an extra component of velocity for the bob in the y direction. Equations (12.13) and (12.14) now become

$$v_x = \dot{x} + l\dot{\theta} \cos \theta \quad (12.24)$$

$$v_y = \dot{y} + l\dot{\theta} \sin \theta, \quad (12.25)$$

but with $\dot{x} = \dot{F}$ and $\dot{y} = \dot{G}$. The kinetic energy and potential of the bob is now

$$T = \frac{1}{2}m(v_x^2 + v_y^2) = \frac{1}{2}m \left(\dot{x}^2 + \dot{y}^2 + l^2\dot{\theta}^2 + 2l\dot{x}\dot{\theta} \cos \theta + 2l\dot{y}\dot{\theta} \sin \theta \right)$$

and

$$V = mg(y - l \cos \theta),$$

hence the Lagrangian becomes

$$L = \frac{1}{2}m \left(\dot{F}^2 + \dot{G}^2 + l^2\dot{\theta}^2 + 2l\dot{F}\dot{\theta} \cos \theta + 2l\dot{G}\dot{\theta} \sin \theta \right) - mg(G - l \cos \theta) \quad (12.26)$$

The equations of motion follow. In particular considering the θ coordinate, the corresponding Euler-Lagrange equation becomes (after a little bit of algebra)

$$l^2\ddot{\theta} + l\ddot{F} \cos \theta + l\ddot{G} \sin \theta + gl \sin \theta = 0 \quad (12.27)$$

We can make progress by looking at particular cases:

Case (a): the pivot only moves horizontally, i.e. $G = 0$. To simplify things and allow us to obtain a solution, we consider the specific case of small oscillations of the pendulum so that $\theta \ll 1$. This then allows us to make use of the small angle results, $\cos \theta \sim 1 - \frac{\theta^2}{2}$, $\sin \theta \sim \theta$. To lowest order in θ , (12.27) becomes

$$\ddot{\theta} + \frac{g}{l}\theta \simeq -\frac{1}{l}\ddot{F}$$

representing the case of a pendulum moving with simple harmonic motion but under the influence of a forcing term, $F(t)$. The fundamental frequency of the pendulum is $\omega = \sqrt{\frac{g}{l}}$. Now if the forcing term is periodic with period say $\frac{2\pi}{\Omega}$ then we can write $F(t)$ as a Fourier series of the form

$$F(t) = \sum_{n=1}^{\infty} [a_n \cos(n\Omega t) + b_n \sin(n\Omega t)].$$

We obtain *resonance* (i.e. growing amplitude solutions) for all frequencies such that $n\Omega = \omega$. In particular we obtain resonance at the fundamental frequency of F , Ω , and all higher harmonics $n\Omega$ – a feature known as superharmonic resonance.

Case (b): the pivot moves vertically, i.e. $F = 0$. If we once again consider the case of small oscillations then Eqn. (12.27) becomes

$$\ddot{\theta} + \frac{\ddot{G}}{l}\theta + \frac{g}{l}\theta = 0$$

Considering the case of a periodic driving term of the form $G(t) = a \cos(\Omega t)$ with $a \ll 1$ we have

$$\ddot{\theta} + \left(\frac{g}{l} - \frac{\Omega^2 a}{l} \cos \Omega t \right) \theta = 0,$$

which is the famous *Mathieu* equation. It leads to parametric excitation, hence resonance for a particular range of frequencies Ω . In particular if $\Omega = 2\omega$ the system enters *subharmonic resonance*.

Let us take a closer look at resonance, which as we have mentioned is an effect which can arise when an oscillator is subject to an applied external periodic force. First of all a reminder (I hope) of some known solutions. A particle mass m near a position of stable equilibrium under a conservative force (i.e. $F(x)$ only) can always be treated approximately as a Simple Harmonic Oscillator. If however there is energy loss, we must include in the equation of motion a force depending on velocity (i.e. $F(x, \dot{x})$). So long as the displacements are small about the equilibrium position we can treat x and \dot{x} as small quantities and drop terms quadratic in them. This leaves us with the damped harmonic oscillator for which the force is $F = -kx - \lambda\dot{x}$, where k is the usual force constant and $\lambda > 0$ is a constant (note not a Lagrange multiplier here). The equation of motion (Newton's Law $F=ma$) is

$$m\ddot{x} + \lambda\dot{x} + kx = 0 \tag{12.28}$$

The solutions are well known and can be obtained by introducing $x = e^{pt}$ and substituting it into (12.28). This gives

$$mp^2 + \lambda p + k = 0 \tag{12.29}$$

with roots

$$p = -\gamma \pm \sqrt{\gamma^2 - \omega_0^2} \tag{12.30}$$

where

$$\gamma = \frac{\lambda}{2m}, \quad \text{and} \quad \omega_0 = \sqrt{\frac{k}{m}}. \tag{12.31}$$

ω_0 is the frequency of the undamped oscillator. A useful quantity is the *Quality Factor*, Q of the oscillator, defined as the dimensionless number

$$Q = \frac{m\omega_0}{\lambda} = \frac{\omega_0}{2\gamma}. \tag{12.32}$$

Three important limits follow:

1. Large damping - λ large, so that $\gamma > \omega_0$ and $Q < 1$.

The general solution is

$$x = \frac{1}{2}Ae^{-\gamma_+t} + \frac{1}{2}Be^{-\gamma_-t} \quad (12.33)$$

where $\gamma_{\pm} = \gamma \pm \sqrt{\gamma^2 - \omega_0^2}$. The displacement tends exponentially to zero and for large times the dominant term is the one with the smaller coefficient γ_- , with a characteristic decay time of order $1/\gamma_-$.

2. Small damping - λ small, so that $\gamma < \omega_0$ and $Q > 1$.

The general solution is

$$x = \frac{1}{2}Ae^{i\omega t - \gamma t} + \frac{1}{2}Be^{-i\omega t - \gamma t} \quad (12.34)$$

$$= ae^{-\gamma t} \cos(\omega t - \theta) \quad (12.35)$$

where $\omega = \sqrt{\omega_0^2 - \gamma^2}$, $A = ae^{-i\theta}$ and $B = ae^{i\theta}$. The solution represents an oscillation with exponentially decreasing amplitude $ae^{-\gamma t}$, angular frequency $\omega < \omega_0$. The *relaxation time* of the oscillator is $1/\gamma = 2m/\lambda$.

3. Critical damping - limiting case $\gamma = \omega_0$ and $Q = \frac{1}{2}$.

In this case $\omega = 0$ and the two roots of p coincide. The general solution in this case is

$$x = (a + bt)e^{-\gamma t}. \quad (12.36)$$

It provides the shortest possible response time.

In an isolated system the forces are functions of position and velocity but not explicitly of the time. Here we consider the case where there is an applied external force which is given as a function of time $F(t)$. We then deal with the eqn

$$m\ddot{x} + \lambda\dot{x} + kx = F(t) \quad (12.37)$$

(i) If $x_1(t)$ is any solution of (12.37) and $x_0(t)$ is a solution of the corresponding *homogeneous* equation (12.28), it follows that $x_1(t) + x_0(t)$ will also be a solution of (12.37). Therefore all we have to do is find *one* particular solution of (12.37). Its general solution is then found by adding to this solution the general solution of the homogeneous equation (12.28). It will be general because it contains two arbitrary constants. Consider the case where the applied force is periodic in time,

$$F(t) = F_1 \cos \omega_1 t \quad (12.38)$$

where F_1 and ω_1 are real constants. It is convenient to rewrite this in the form $F(t) = \text{Re}(F_1 e^{i\omega_1 t})$ and to solve first the equation with a complex force

$$m\ddot{z} + \lambda\dot{z} + kz = F_1 e^{i\omega_1 t} \quad (12.39)$$

The real part x of this solution will be a solution of (12.37) with the force (12.38). Now a particular solution to (12.37) is

$$z = A_1 e^{i\omega_1 t} = a_1 e^{i(\omega_1 t - \theta_1)} \quad (12.40)$$

where $A_1 = a_1 e^{-i\theta_1}$ is a complex constant. By equating real and imaginary parts of (12.37) it follows that

$$(\omega_0^2 - \omega_1^2)a_1 = \frac{F_1}{m} \cos \theta_1 \quad (12.41)$$

$$2\gamma\omega_1 a_1 = \frac{F_1}{m} \sin \theta_1 \quad (12.42)$$

The amplitude and phase of the forced oscillation can then be read off and are given by

$$a_1 = \frac{\frac{F_1}{m}}{\sqrt{(\omega_0^2 - \omega_1^2)^2 + 4\gamma^2\omega_1^2}} \quad (12.43)$$

$$\tan \theta_1 = \frac{2\gamma\omega_1}{(\omega_0^2 - \omega_1^2)} \quad (12.44)$$

If $F_1 > 0$, the correct choice between the two solutions for θ_1 differing by π is the one lying in the range $0 < \theta_1 < \pi$. By considering the real part of z for the particular solution, for the case where the damping is less than critical ($\gamma < \omega_0$), we obtain for the general solution to (12.37)

$$x = a_1 \cos(\omega_1 t - \theta_1) + a e^{-\gamma t} \cos(\omega t - \theta), \quad (12.45)$$

where a and θ are arbitrary constants to be fixed by the initial conditions. By taking the $t \rightarrow \infty$ limit, we see that at late times, no matter what initial conditions we choose, the oscillations are ultimately governed solely by the external force. It is clear from (12.43) and (12.44) that the amplitude a_1 and phase θ_1 of the forced oscillations are strongly dependent on the angular frequencies ω_0 and ω_1 . If the damping is small, the amplitude can become very large when the frequencies are almost equal. Fixing γ and the forcing frequency ω_1 we see that the amplitude is maximum when $\omega_0 = \omega_1$. In this case the system is in *resonance*. Try to show that the corresponding amplitude is

$$a_1 = \frac{F_1}{2m\gamma\omega_1} = \frac{F_1}{\lambda\omega_1}, \quad (12.46)$$

which can of course be large if λ is small.

Now consider fixing γ and the oscillator frequency ω_0 , and vary the forcing frequency ω_1 . Show that the maximum amplitude now occurs for a slightly lower frequency than ω_0 , namely

$$\omega_1 = \sqrt{\omega_0^2 - 2\gamma^2}. \quad (12.47)$$

Of course if γ is very small (small λ) the difference is not very much. It is worth noting that the natural frequency of the oscillator $\omega = \sqrt{\omega_0^2 - \gamma^2}$ given above lies between this resonant frequency and the frequency ω_0 of the undamped oscillator.

13 Charged particle in an electromagnetic field

Up to now we have considered examples where the potential is only a function of the coordinate, i.e. $V(q_i)$. We now turn our attention to the case of velocity dependent potentials by considering the motion of a particle mass m , charge q moving with velocity \underline{v} in an electromagnetic field, where the electric and magnetic fields are given by $\underline{E}(\underline{x}, t)$ and $\underline{B}(\underline{x}, t)$ respectively. The particle feels a Lorentz force given by

$$\underline{F} = q\underline{E} + q\underline{v} \wedge \underline{B},$$

hence Newton's Law ($F = ma$) gives us the equation of motion

$$\frac{d}{dt}(m\dot{\underline{x}}) = q(\underline{E} + \underline{v} \wedge \underline{B}) \quad (13.1)$$

The question we would like to answer is whether it is possible to construct a Lagrangian L such that the resulting equation of motion of this form? Recall the Euler-Lagrange equation is given by

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}_i} \right) - \frac{\partial L}{\partial x_i} = 0$$

with $L = T - V$, then it is easy to see that in order to obtain the LHS of Eqn. (13.1) we simply need to take for the kinetic energy of the particle $T = \frac{1}{2}m\dot{x}_i\dot{x}_i$ (recall the implied sum over $i = 1, 2, 3$). Our aim then is to find an electromagnetic 'potential' $M(\dot{x}_i, x_i, t)$ such that $L = T - M$. It is clear that M needs to be a velocity dependent potential because the RHS of Eqn. (13.1) contains an explicit \dot{x}_i term in it.

Let us take a little aside here on electromagnetic theory. In the absence of dielectric or magnetic media, the electric field \underline{E} and the magnetic field \underline{B} satisfy the four *Maxwell equations*. One pair are the homogeneous equations

$$\nabla \wedge \underline{E} + \frac{\partial \underline{B}}{\partial t} = 0 \quad (13.2)$$

$$\nabla \cdot \underline{B} = 0, \quad (13.3)$$

and the second pair involving the electric charge density ρ and current density \underline{j} are

$$\mu_0^{-1} \nabla \wedge \underline{B} - \epsilon_0 \frac{\partial \underline{E}}{\partial t} = \underline{j} \quad (13.4)$$

$$\epsilon_0 \nabla \cdot \underline{E} = \rho, \quad (13.5)$$

where μ_0 and ϵ_0 are universal constants (actually they give the speed of light in vacuum through $c = \frac{1}{\sqrt{\mu_0\epsilon_0}}$). The key equations here are (13.2) and (13.3). The potentials we are looking for are introduced by using vector identities to rewrite \underline{E} and \underline{B} in terms of potentials. In particular making use of the fact that $\nabla \cdot (\nabla \wedge \underline{A}) = 0$ for any vector \underline{A} it follows that in (13.3) we can write

$$\underline{B} = \nabla \wedge \underline{A} \quad (13.6)$$

where $\underline{A}(\underline{x}, t)$ is the *magnetic vector potential*. Now, substituting into (13.2)) we obtain

$$\nabla \wedge \left(\underline{E} + \frac{\partial \underline{A}}{\partial t} \right) = 0$$

Making use of a second vector identity, namely $\nabla \wedge \nabla \phi = 0$ for any scalar ϕ we can write

$$\underline{E} = -\frac{\partial \underline{A}}{\partial t} - \nabla \phi \quad (13.7)$$

for some *electric scalar potential* $\phi(\underline{x}, t)$. An important point we need to make is that equations (13.7) and (13.6) do not fully specify ϕ and \underline{A} – there remains a degree of arbitrariness in their values which needs to be pinned down. This introduces us to the idea of “choosing a gauge” which is basically deciding how ϕ and \underline{A} are to be linked. Why have they not been fully specified? Consider Eqn. (13.6) for a given \underline{A} . Now because of the vector identity $\nabla \wedge \nabla \psi = 0$ for any scalar ψ , it follows there is another vector potential $\underline{A}' = \underline{A} + \nabla \psi$ which is also a solution to Eqn. (13.6). In other words, the vector potential \underline{A} is not unique. Moreover if we consider Eqn. (13.7) for a given $\phi(\underline{x}, t)$, it follows that we can define a new scalar potential $\tilde{\phi} = \phi + f(t)$ (for an arbitrary $f(t)$) which is also a solution because $\nabla f(t) = 0$. Pinning down the values for these terms is an example of a choice of gauge. one such example is the ‘Lorentz gauge’ and is given in four-vector notation by $\nabla_\mu A^\mu = 0$ where $A^0 = \phi$ and $A^i = -\underline{A}$. In more familiar notation choosing the Lorentz gauge implies

$$\frac{1}{c^2} \frac{\partial \phi}{\partial t} + \nabla \cdot \underline{A} = 0, \quad (13.8)$$

a condition that explicitly links ϕ and \underline{A} .

We now return to the equation of motion (13.1) and consider say the i th component of the RHS. It can be written as

$$q(E_i + \epsilon_{ijk} \dot{x}_j B_k). \quad (13.9)$$

Here $\epsilon_{i,j,k}$ is the Levi-Civita (or Epsilon) tensor and i, j, k can each have values 1, 2, 3. It is defined by

$$\begin{aligned} \epsilon_{i,j,k} &= 0 && \text{if two indices are the same} \\ \epsilon_{1,2,3} &= \epsilon_{2,3,1} = \epsilon_{3,1,2} = 1 \\ \epsilon_{1,3,2} &= \epsilon_{2,1,3} = \epsilon_{3,2,1} = -1 \end{aligned}$$

Thus it is invariant with respect to cyclic permutation of indices and changes its sign if indices are interchanged. As an example, consider the vector product $\mathbf{c} = \mathbf{a} \times \mathbf{b}$. By explicitly evaluating the vector product show that the i th component c_i can be expressed as

$$c_i = \sum_{j,k=1}^3 \epsilon_{i,j,k} a_j b_k$$

It follows from this result that if $\text{curl } \mathbf{A} = \mathbf{B}$ then the i th component B_i is given by

$$B_i = \epsilon_{i,j,k} \frac{\partial A_k}{\partial x_j} \quad (13.10)$$

Also, note that in Eq. (13.9) we are using the implied summation convention i.e. $\epsilon_{ijk}B_k = \epsilon_{ij1}B_1 + \epsilon_{ij2}B_2 + \epsilon_{ij3}B_3$. Now substituting in for the components E_i and B_k from (13.6) and (13.7) we can rewrite (13.9) as

$$q \left(-\frac{\partial\phi}{\partial x_i} - \frac{\partial A_i}{\partial t} + \epsilon_{ijk} \dot{x}_j \epsilon_{klm} \frac{\partial A_m}{\partial x_l} \right). \quad (13.11)$$

A useful identity in three dimensions is $\epsilon_{ijk}\epsilon_{klm} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}$ where the Kronecker delta tensor is defined by

$$\begin{aligned} \delta_{i,j} &= 0 & \text{if } i \neq j \\ \delta_{i,j} &= 1 & \text{if } i = j. \end{aligned} \quad (13.12)$$

This allows us to rewrite (13.11) as

$$\begin{aligned} &= q \left(-\frac{\partial\phi}{\partial x_i} - \frac{\partial A_i}{\partial t} + (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}) \dot{x}_j \frac{\partial A_m}{\partial x_l} \right) \\ &= q \left(-\frac{\partial\phi}{\partial x_i} - \frac{\partial A_i}{\partial t} + \dot{x}_j \frac{\partial A_j}{\partial x_i} - \dot{x}_j \frac{\partial A_i}{\partial x_j} \right) \end{aligned} \quad (13.13)$$

Recall that because we have already accounted for the kinetic part of the Lagrangian, we require (13.13) to be equal to

$$\frac{d}{dt} \left(\frac{\partial M}{\partial \dot{x}_i} \right) - \frac{\partial M}{\partial x_i}.$$

By inspection (or guess work) let us try a solution of the form

$$M = q(\phi - \dot{x}_j A_j) \quad (13.14)$$

(recall assumed summation of indices in $\dot{x}_j A_j$. We now need to substitute this into the Euler-Lagrange equation for M and see whether we recover (13.13). Doing this we have

$$\frac{d}{dt} \left(\frac{\partial M}{\partial \dot{x}_i} \right) - \frac{\partial M}{\partial x_i} = -q \frac{dA_i}{dt} - q \frac{\partial\phi}{\partial x_i} + q \dot{x}_j \frac{\partial A_j}{\partial x_i} \quad (13.15)$$

But since $A_i(\underline{x}, t)$ then by using the Chain rule on the total derivative $\frac{dA_i}{dt}$

$$\frac{dA_i}{dt} = \frac{\partial A_i}{\partial t} + \frac{\partial A_i}{\partial x_j} \dot{x}_j,$$

we find

$$\frac{d}{dt} \left(\frac{\partial M}{\partial \dot{x}_i} \right) - \frac{\partial M}{\partial x_i} = q \left(-\frac{\partial\phi}{\partial x_i} - \frac{\partial A_i}{\partial t} + \dot{x}_j \frac{\partial A_j}{\partial x_i} - \dot{x}_j \frac{\partial A_i}{\partial x_j} \right) \quad (13.16)$$

the same as (13.13) as required. We conclude that a Lagrangian for a charged particle in an electromagnetic field is of the form $L = T - M$ with $T = \frac{1}{2}m\dot{x}_i\dot{x}_i$ and $M = q(\phi - \dot{x}_j A_j) = q(\phi - \underline{v} \cdot \underline{A})$ is a velocity dependent potential. Note that we have done this calculation assuming

a cartesian system of coordinates, but as you saw in Tony's lectures we can conclude that the equations of motion in general coordinates q_i are

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) - \frac{\partial L}{\partial q_i} = 0.$$

Example: Consider a charged particle, mass m , charge q moving in a steady uniform magnetic field which points in the z direction, i.e. $\mathbf{B} = (0, 0, B)$. By considering cylindrical polar coordinates (r, θ, z) , we can show that the vector potential \mathbf{A} can be given by the time independent form $\mathbf{A} = (0, A_\theta(r), 0)$ where

$$A_\theta = \frac{1}{2}Br \tag{13.17}$$

Specifically, we know $\mathbf{B} = \text{curl } \mathbf{A}$, and since in cylindrical polars we have:

$$\text{curl } \mathbf{A} = \left(\left(\frac{1}{r} \frac{\partial A_z}{\partial \theta} - \frac{\partial A_\theta}{\partial z} \right), \left(\frac{\partial A_r}{\partial z} - \frac{\partial A_z}{\partial r} \right), \frac{1}{r} \left(\frac{\partial(rA_\theta)}{\partial r} - \frac{\partial A_r}{\partial \theta} \right) \right),$$

then by comparing the z components of the two expressions for $\text{curl } \mathbf{A}$ we have

$$\frac{1}{r} \left(\frac{\partial(rA_\theta)}{\partial r} - \frac{\partial A_r}{\partial \theta} \right) = B.$$

It follows that a particular solution of this equation is the time independent form

$$\mathbf{A} = (0, A_\theta(r), 0)$$

where

$$B = \frac{1}{r} \frac{\partial(rA_\theta)}{\partial r}$$

hence

$$A_\theta(r) = \frac{1}{2}Br.$$

Now we know in general that the electric field \mathbf{E} is given in terms of the vector and scalar potential as

$$\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t}$$

Here we only have a magnetic field, $\mathbf{E} = 0$ and since $\mathbf{A}(r)$ only (i.e. indep of t), then we have $\nabla\phi = 0$ which has $\phi = 0$ as a consistent solution.

We have $L = T - M$ where $T = \frac{1}{2}mv^2$ and $M = q(\phi - \mathbf{v} \cdot \mathbf{A})$. Now since the instantaneous velocity of the particle at position (r, θ, z) is given by $\mathbf{v} = \dot{\mathbf{x}} = (\dot{r}, r\dot{\theta}, \dot{z})$, it follows that

$$v^2 = \mathbf{v} \cdot \mathbf{v} = \dot{r}^2 + r^2\dot{\theta}^2 + \dot{z}^2$$

and

$$\mathbf{v} \cdot \mathbf{A} = \frac{1}{2}\dot{\theta}Br^2$$

hence, since $\phi = 0$, the Lagrangian for the system is given by

$$L = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2 + \dot{z}^2) + \frac{q}{2}Br^2\dot{\theta} \quad (13.18)$$

Since $\frac{\partial L}{\partial z} = \frac{\partial L}{\partial \theta} = 0$ then z and θ are ignorable co-ordinates. The associated generalised momenta are given by

$$p_z = \frac{\partial L}{\partial \dot{z}} = m\dot{z} = \text{const} \quad (13.19)$$

$$p_\theta = \frac{\partial L}{\partial \dot{\theta}} = mr^2\dot{\theta} + \frac{q}{2}Br^2 = \text{const} \quad (13.20)$$

and represent generalised linear and angular momenta respectively.

Another conserved quantity follows from the fact that $\frac{\partial L}{\partial t} = 0$. The corresponding conserved quantity is given by the Hamiltonian. In terms of the coordinates (r, θ, z) the conserved energy is given by

$$\dot{r}\frac{\partial L}{\partial \dot{r}} + \dot{\theta}\frac{\partial L}{\partial \dot{\theta}} + \dot{z}\frac{\partial L}{\partial \dot{z}} - L = E = \text{const.}$$

From (13.18) we obtain after a little algebra

$$E = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2 + \dot{z}^2) \quad (13.21)$$

So far we have not considered the r -eqn. From (13.18) the Euler-Lagrange eqn $\frac{\partial L}{\partial r} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{r}}\right) = 0$ gives

$$m\ddot{r} - qBr\dot{\theta} - mr\dot{\theta}^2 = 0$$

A solution where $r = R = \text{const}$ therefore implies that

$$\dot{\theta} = -\frac{qB}{m} = \text{const.} \quad (13.22)$$

Substituting into (13.20) we obtain

$$p_\theta = -\frac{qBR^2}{2} = \text{const} \quad (13.23)$$

Such a solution corresponds to a particle orbiting in a circle of radius R about the origin in the $z = 0$ plane, and having a constant ‘gyro-frequency’ $\dot{\theta}$.

14 Hamilton’s Equations

We now start discussing an alternative yet equally powerful approach to solving dynamical systems. It was pioneered by the great Irish physicist, astronomer and mathematician, Sir

William Rowan Hamilton (4 August 1805 – 2 September 1865) and the equations to be used *Hamilton's Equations* bear his name. He suggested that the Lagrangian method that we have been using up to now could be extended by using the generalised momenta p_1, p_2, \dots, p_n in place of the generalised velocities $\dot{q}_1, \dot{q}_2, \dots, \dot{q}_n$. Although more 'abstract' than the Lagrangian formalism it possesses the advantage of getting to the symmetries of a problem quicker because they often correspond to situations where the generalised momenta are constants of the motion. Therefore the approach is well suited to finding conserved quantities, and we now turn our attention to the derivation of Hamilton's Equations.

We start by writing the Lagrangian for the generalised co-ordinate system $q_i(t)$, $i = 1..n$, $L(\dot{q}_1, \dot{q}_2, \dots, \dot{q}_n, q_1, q_2, \dots, q_n) = L(\dot{q}, q)$ for brevity, where q stands now for *all* the generalised coordinates and \dot{q} for *all* their time derivatives. Now recall the Euler Lagrange equation (12.1) and the definition of the generalised momenta in (12.3). Combining them we obtain the result

$$\frac{dp_\alpha}{dt} = \frac{\partial L}{\partial q_\alpha}, \quad \alpha = 1, 2, \dots, n \quad (14.1)$$

where we recall the generalised momenta is defined in (12.3), $p_\alpha = \frac{\partial L}{\partial \dot{q}_\alpha}$.

So far, with the Lagrangian system we have considered $2n$ variables (q, \dot{q}) to uniquely specify every part of the system. An alternative approach is to solve $p_\alpha = \frac{\partial L}{\partial \dot{q}_\alpha}$ for $\dot{q}_\alpha(q, p)$ thereby giving \dot{q}_α as a function of (q, p) . We still have a system of $2n$ variables (q, p) which serves us just as well as before to specify the instantaneous position and velocity of every particle.

Example: consider a particle of mass m moving in a plane and with instantaneous position r, θ in polar coordinates. It has kinetic energy $T = \frac{1}{2}m\dot{r}^2 + \frac{1}{2}mr^2\dot{\theta}^2$ and its potential is $V(r, \theta)$ leading to the Lagrangian

$$L = \frac{1}{2}m\dot{r}^2 + \frac{1}{2}mr^2\dot{\theta}^2 - V(r, \theta). \quad (14.2)$$

The associated generalised momenta are then

$$p_r = \frac{\partial L}{\partial \dot{r}} = m\dot{r} \quad (14.3)$$

$$p_\theta = \frac{\partial L}{\partial \dot{\theta}} = mr^2\dot{\theta} \quad (14.4)$$

which can be inverted to give

$$\dot{r} = \frac{p_r}{m} \quad (14.5)$$

$$\dot{\theta} = \frac{p_\theta}{mr^2}. \quad (14.6)$$

This in turn then implies that the instantaneous position and velocity of the particle can be fixed by the values $(r, \theta, p_r, p_\theta)$.

14.1 Hamiltonian Function – $H(q, p)$

Earlier in section 12.2 we briefly introduced the concept of the Hamiltonian. We now make it more formal: the Hamiltonian Function $H(p, q)$ is defined by

$$H(q, p) = \sum_{\beta=1}^n p_{\beta} \dot{q}_{\beta}(q, p) - L(q, \dot{q}(q, p)) \quad (14.7)$$

To derive Hamilton's two key equations we first differentiate (14.7) wrt p_{α} for the first and wrt q_{α} for the second. For the first case we have

$$\frac{\partial H}{\partial p_{\alpha}} = \dot{q}_{\alpha} + \sum_{\beta=1}^n p_{\beta} \frac{\partial \dot{q}_{\beta}}{\partial p_{\alpha}} - \sum_{\beta=1}^n \frac{\partial L}{\partial \dot{q}_{\beta}} \frac{\partial \dot{q}_{\beta}}{\partial p_{\alpha}}. \quad (14.8)$$

Using Eqn. (12.3) we see that the second and third terms on the RHS of (14.8) cancel leaving us with

$$\frac{\partial H}{\partial p_{\alpha}} = \dot{q}_{\alpha}. \quad (14.9)$$

For the second case obtain

$$\frac{\partial H}{\partial q_{\alpha}} = \sum_{\beta=1}^n p_{\beta} \frac{\partial \dot{q}_{\beta}}{\partial q_{\alpha}} - \sum_{\beta=1}^n \frac{\partial L}{\partial \dot{q}_{\beta}} \frac{\partial \dot{q}_{\beta}}{\partial q_{\alpha}} - \frac{\partial L}{\partial q_{\alpha}} \quad (14.10)$$

Once again this can be simplified. using Eqn. (12.3) we see that the first and second terms on the RHS of (14.10) cancel. Moreover from Eqn. (14.1) we obtain

$$\frac{\partial H}{\partial q_{\alpha}} = -\dot{p}_{\alpha} \quad (14.11)$$

Together Eqns. (14.9) and (14.11) form *Hamilton's Equations*.

We note in passing that the Euler-Langrange equations form a set of n second order equations where as Hamilton's equations form a set of $2n$ first order equations.

Example: Central Conservative force $F(r)$ in a plane: (recall a conservative force $F(x)$ depends on on x satisfying $F(x) = -\frac{dV}{dx}$, and a Central force is one where $V(r)$ only, i.e. independent of θ, ϕ , hence $\underline{F}(r) = \hat{r} \frac{dV(r)}{dr}$). We want to obtain Hamilton's equations for a particle moving in a plane with potential energy function $V(r)$. The Lagrangian (14.2) becomes

$$L = \frac{1}{2} m \dot{r}^2 + \frac{1}{2} m r^2 \dot{\theta}^2 - V(r). \quad (14.12)$$

The Hamiltonian follows from (14.7)

$$H = (p_r \dot{r} + p_{\theta} \dot{\theta}) - L(r, \theta, \dot{r}, \dot{\theta}) \quad (14.13)$$

Substituting for the Lagrangian (14.12) whilst eliminating the velocities using (14.5) and (14.6) we obtain

$$H = \frac{p_r^2}{2m} + \frac{p_\theta^2}{2mr^2} + V(r) \quad (14.14)$$

which is nothing other than the total energy $T + V$. Hamilton's equations (14.9) - (14.11) then become:

$$\dot{r} = \frac{\partial H}{\partial p_r} = \frac{p_r}{m} \quad (14.15)$$

$$\dot{\theta} = \frac{\partial H}{\partial p_\theta} = \frac{p_\theta}{mr^2} \quad (14.16)$$

$$-\dot{p}_r = \frac{\partial H}{\partial r} = -\frac{p_\theta^2}{mr^3} + \frac{dV(r)}{dr} \quad (14.17)$$

$$-\dot{p}_\theta = \frac{\partial H}{\partial \theta} = 0 \quad (14.18)$$

Note that Eqns. (14.15)-(14.16) simply reproduces the earlier results (14.5) - (14.6) which was obtained through generalised momenta Eqn. (12.3), $p_\alpha = \frac{\partial L}{\partial \dot{q}_\alpha}$. Now, from (14.18) we obtain our first conserved quantity. Since $\dot{p}_\theta = 0$ we can integrate to obtain

$$p_\theta = J = \text{const}$$

corresponding to *conservation of angular momentum*. Considering (14.17), making use of (14.15) and our newly obtained angular momentum J we can rewrite it as:

$$\dot{p}_r = m\ddot{r} = \frac{J^2}{mr^3} - \frac{dV}{dr}$$

Multiplying by \dot{r} and integrating wrt t we obtain the *radial energy equation* :

$$\frac{1}{2}m\dot{r}^2 + \frac{J^2}{2mr^2} + V(r) = E = \text{const.}$$

To end this section we return to the question of when is the energy of a system conserved, and how it can be interpreted in terms of the Hamiltonian. A *Natural System* is one whose kinetic energy T contains no explicit time dependence and that is no more than quadratic in velocity (i.e. contains no more than \dot{q}_i^2 terms). This then implies that

$$\sum_{\alpha=1}^n \frac{\partial T}{\partial \dot{q}_\alpha} \dot{q}_\alpha = 2T$$

Now assuming the potential only depends on position (i.e. $V(q_\alpha)$), then from Eqn. (12.3) we have

$$p_\alpha \equiv \frac{\partial L}{\partial \dot{q}_\alpha} = \frac{\partial(T - V)}{\partial \dot{q}_\alpha} = \frac{\partial T}{\partial \dot{q}_\alpha},$$

hence the Hamiltonian (14.7) becomes

$$H(q, p) = \sum_{\beta=1}^n p_{\beta} \dot{q}_{\beta}(q, p) - L(q, \dot{q}(q, p)) = \sum_{\beta=1}^n \frac{\partial T}{\partial \dot{q}_{\beta}} \dot{q}_{\beta}(q, p) - (T - V)$$

which simplifies to give

$$H(q, p) = T + V.$$

We conclude that for a natural system the value of the Hamiltonian function is equal to the total energy of the system. However, this is not always the case. Consider the case where the system is forced and has an explicit time dependence say through the forcing term. In that case the Hamiltonian function does not give the total energy.

Earlier, in Eqn. (12.12) we saw that the condition for constant H was that $\frac{\partial L}{\partial t} = 0$, i.e. that the Lagrangian had no explicit time dependence. What does it imply for the time dependence of H ? We can answer this by allowing H to be time dependent, i.e. by considering $H(p, q, t)$. We then have (applying the chain rule)

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} + \sum_i \frac{\partial H}{\partial q_i} \dot{q}_i + \sum_i \frac{\partial H}{\partial p_i} \dot{p}_i.$$

Substituting in for \dot{q}_i and \dot{p}_i from Hamilton's equations we obtain

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} + \sum_i \left(\frac{\partial H}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial H}{\partial p_i} \frac{\partial H}{\partial q_i} \right).$$

leaving us with the equally elegant result

$$\frac{dH}{dt} = \frac{\partial H}{\partial t}$$

which implies that the Hamiltonian is time dependent *only* because of the explicit time dependence present. Otherwise it would be constant!

in particular a natural conservative system has neither T nor V depending explicitly on time, hence $\frac{\partial H}{\partial t} = 0$ implying $\frac{dH}{dt} = 0$ and leading naturally to the law of conservation of energy

$$H = T + V = E = \text{const.}$$

The Hamiltonian formalism leading to first order equations, makes it well suited to finding conservation laws or constants of the motion. We have just seen that conservation of energy emerges from having t as an ignorable coordinate. Earlier we saw that conservation of momentum meant that the coordinate q_j say was ignorable. In general for a coordinate q_{α} to be ignorable, the Hamiltonian $H(q, p, t)$ is such that

$$-\dot{p}_{\alpha} \equiv \frac{\partial H}{\partial q_{\alpha}} = 0$$

implying that

$$p_\alpha = \text{const.}$$

This is of course perfectly consistent with the condition we obtained earlier in (12.2)-(12.5) when considering directly the Lagrangian $L(q, \dot{q}, t)$. In that case we saw that if $\frac{\partial L}{\partial q_\alpha} = 0$ this implied $\dot{p}_\alpha = \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_\alpha} = 0$ hence $p_\alpha = \text{const.}$

15 The symmetric top – or toy gyroscope

In this section we will apply Hamilton's equations and demonstrate the power of ignorable coordinates in identifying symmetries of a system by considering the dynamics of a toy that five year old children can master without knowing anything about Hamiltonians or Lagrangians – or are we underestimating their innate mathematical ability?

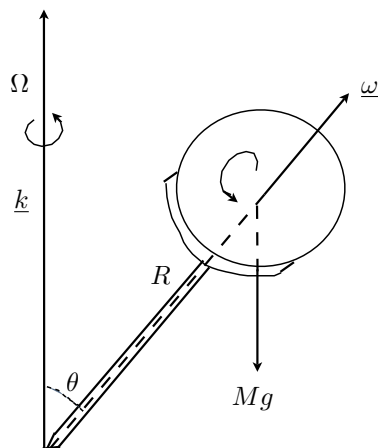


Figure 15.1: Symmetric rigid body pivoted at a point on its axis of symmetry and moving under gravity.

A spinning top can be considered as a symmetric rigid body of mass M pivoted at a point on its axis of symmetry and moving under gravity. Its potential energy can be seen from Fig. (15.1) to be

$$V = MgR \cos \theta \tag{15.1}$$

where R is the distance of centre of mass of the top from the origin and θ is the angle it makes to the vertical. Now when dealing with a *rigid body* we need to make use of some of the machinery which you came across in chapter 7. In particular the moment of inertia and the principal axes. To remind you, the principal axes of inertia are perpendicular axes through any given point which ensure the moment of inertia tensor I_{ij} is diagonal (i.e. I_{11}, I_{22}, I_{33}

only, labeled I_1, I_2, I_3) with $I_{ij} = 0, i \neq j$. We now introduce the set of unit vectors, $\{\hat{e}_1, \hat{e}_2, \hat{e}_3\}$ along these axes. Note that the principal axes $\{\hat{e}_1, \hat{e}_2, \hat{e}_3\}$ are fixed in the body, not in space. (In this sense they do not have to correspond to the fixed cartesian axes we defined in chapter 1). They can rotate with the body in space. Considering a rigid body which is free to rotate about a fixed point, the position of every point of the body is fixed once $\{\hat{e}_1, \hat{e}_2, \hat{e}_3\}$ has been specified. For example the position of a point on the body is given by

$$\underline{r} = r_1\hat{e}_1 + r_2\hat{e}_2 + r_3\hat{e}_3, \quad r_1, r_2, r_3 \text{ const} \quad (15.2)$$

and the velocity of that point is given by

$$\dot{\underline{r}} = r_1\dot{\hat{e}}_1 + r_2\dot{\hat{e}}_2 + r_3\dot{\hat{e}}_3. \quad (15.3)$$

The instantaneous angular velocity is given by

$$\underline{\omega} = \omega_1\hat{e}_1 + \omega_2\hat{e}_2 + \omega_3\hat{e}_3, \quad (15.4)$$

and the instantaneous angular momentum is

$$\underline{J} = I_1\omega_1\hat{e}_1 + I_2\omega_2\hat{e}_2 + I_3\omega_3\hat{e}_3, \quad (15.5)$$

where to remind you I_1, I_2, I_3 are the principal moments of inertia. Recall from (7.6), the *rotational kinetic energy* is given by

$$T = \frac{1}{2}\underline{\omega} \cdot I\underline{\omega} = \sum_{i,j} \frac{1}{2}I_{ij}\omega_i\omega_j,$$

then for the case of diagonal principal moments of inertia this then becomes

$$T = \frac{1}{2}I_1\omega_1^2 + \frac{1}{2}I_2\omega_2^2 + \frac{1}{2}I_3\omega_3^2. \quad (15.6)$$

15.1 Euler angles

In order to uniquely determine the orientation of a rigid body about a fixed point or about its centre of mass we must specify three angles. Although these can be chosen in a number of ways it is traditional to choose Euler's angles to do this. By specifying two of the angles, it is possible to fix the direction of one of the axes, say \hat{e}_3 . These are the conventional polar angles $\theta, 0 \leq \theta \leq \pi$ and $\varphi, 0 \leq \varphi \leq 2\pi$. The third angle ψ specifies the angle through which the rigid body has been rotated from a standard position about this axis.

To achieve an arbitrary orientation we can start with the body in a standard position and make three successive rotations. Let us do it:

1. Initially we set the body up so that $\{\hat{e}_1, \hat{e}_2, \hat{e}_3\}$ of the body coincide with the fixed axes $\{\underline{i}, \underline{j}, \underline{k}\}$.

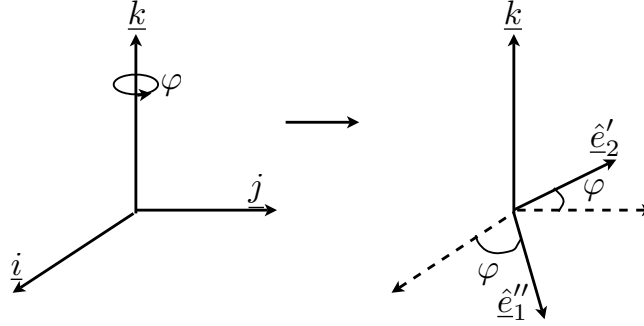


Figure 15.2: Step 1: Rotate by φ around the space-frame axis \underline{k} .

2. The first two rotations bring the axis \hat{e}_3 to its required position given by θ, φ .
3. Specifically, first rotate through an angle φ about the \underline{k} axis. This brings the three axes into positions $\{\hat{e}'_1, \hat{e}'_2, \underline{k}\}$ as seen in Figure. 15.2.
4. Next rotate through an angle θ about the \hat{e}'_2 axis. This brings the three axes into positions $\{\hat{e}'_1, \hat{e}'_2, \hat{e}_3\}$ as seen in Figure. 15.3.
5. Finally rotate through an angle ψ about the \hat{e}_3 axis. This brings the three axes into positions $\{\hat{e}_1, \hat{e}_2, \hat{e}_3\}$ as seen in Figure. 15.4.– and we are done.

15.2 Angular velocity, angular momentum and kinetic energy

Now the three Euler angles φ, θ, ψ fixes the orientations of the three axes and therefore completely specify the orientation of the rigid body. They lead to a simple expression for the *instantaneous angular velocity* $\underline{\omega}$. We can obtain it by considering the motion of a rigid body in an infinitesimal time dt during which the angles undergo an infinitesimal change

$$(\varphi, \theta, \psi) \longrightarrow (\varphi + d\varphi, \theta + d\theta, \psi + d\psi).$$

Now from the definition of the Euler angles we know that the change in $\varphi, d\varphi$ leads to a rotation about the \underline{k} axis leading to an angular velocity in that direction of $\dot{\varphi}\underline{k}$. Similarly the change in $\theta, d\theta$ leads to a rotation about the \hat{e}'_2 axis leading to an angular velocity in that direction of $\dot{\theta}\hat{e}'_2$, and for the change $d\psi$, there is a rotation about the \hat{e}_3 axis and a corresponding angular velocity $\dot{\psi}\hat{e}_3$. If all three angles change, the total angular velocity becomes

$$\underline{\omega} = \dot{\varphi}\underline{k} + \dot{\theta}\hat{e}'_2 + \dot{\psi}\hat{e}_3. \quad (15.7)$$

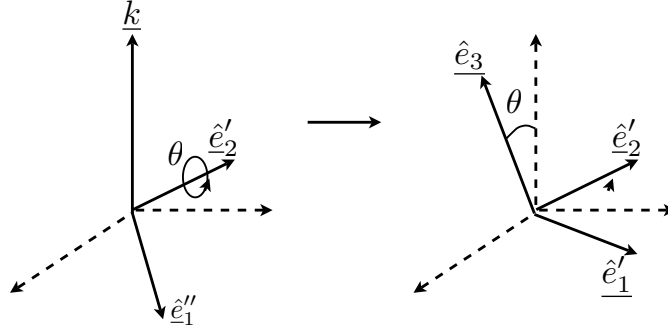


Figure 15.3: Step 2: Rotate by θ around the new axis $\underline{\hat{e}}'_2$.

In order to obtain the angular momentum and kinetic energy of the rotating rigid body we require the components of $\underline{\omega}$ in the direction of the principal axes $\{\underline{\hat{e}}_1, \underline{\hat{e}}_2, \underline{\hat{e}}_3\}$, as given by (15.5) and (15.6). Now rather than deal with the general case we will limit ourselves to the case of a *symmetric body* with symmetry axis $\underline{\hat{e}}_3$ and $I_1 = I_2$. (The definition of a symmetric rigid body is one where two of the principal moments of inertia coincide).

Any orthogonal pair of axes in the plane of $\underline{\hat{e}}_1$ and $\underline{\hat{e}}_2$ will serve as principal axes together with $\underline{\hat{e}}_3$. In particular the pair $\underline{\hat{e}}'_1, \underline{\hat{e}}'_2$ will work just fine. Using these three axes, it means that in (15.7) for the angular velocity $\underline{\omega}$ we just require \underline{k} in terms of $\underline{\hat{e}}'_1, \underline{\hat{e}}'_2, \underline{\hat{e}}_3$. Inspection of the right hand diagram in figure (15.3) shows that

$$\underline{k} = -\sin \theta \underline{\hat{e}}'_1 + \cos \theta \underline{\hat{e}}_3.$$

Substituting for \underline{k} in (15.7), we find

$$\underline{\omega} = -\dot{\varphi} \sin \theta \underline{\hat{e}}'_1 + \dot{\theta} \underline{\hat{e}}'_2 + (\dot{\psi} + \dot{\varphi} \cos \theta) \underline{\hat{e}}_3 \quad (15.8)$$

Then substituting (15.8) into (15.5) and (15.6) we obtain for the angular momentum and kinetic energy of the symmetric rigid body:

$$\begin{aligned} \underline{J} &= I_1 \omega_1 \underline{\hat{e}}'_1 + I_2 \omega_2 \underline{\hat{e}}'_2 + I_3 \omega_3 \underline{\hat{e}}_3, \\ &= -I_1 \dot{\varphi} \sin \theta \underline{\hat{e}}'_1 + I_1 \dot{\theta} \underline{\hat{e}}'_2 + I_3 (\dot{\psi} + \dot{\varphi} \cos \theta) \underline{\hat{e}}_3, \quad \text{using } I_1 = I_2 \end{aligned} \quad (15.9)$$

$$\begin{aligned} T &= \frac{1}{2} I_1 \omega_1^2 + \frac{1}{2} I_2 \omega_2^2 + \frac{1}{2} I_3 \omega_3^2 \\ T &= \frac{1}{2} I_1 \dot{\varphi}^2 \sin^2 \theta + \frac{1}{2} I_1 \dot{\theta}^2 + \frac{1}{2} I_3 (\dot{\psi} + \dot{\varphi} \cos \theta)^2 \end{aligned} \quad (15.10)$$

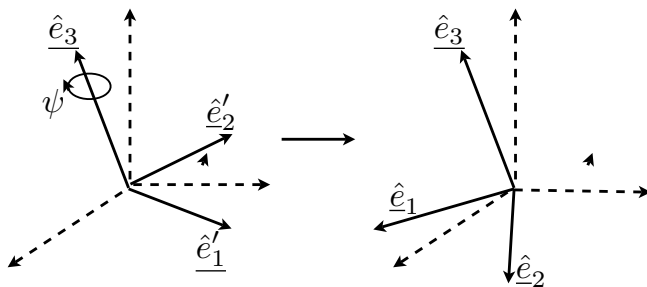


Figure 15.4: Step 3: Rotate by ψ around the latest axis \hat{e}_3 .

15.3 Precession of a symmetric top

Having established that we can use the Euler angles to define the position of the rotating rigid body, we proceed to use them as generalised coordinates to obtain the dynamics of the spinning symmetric top. The kinetic energy is given by (15.10) and the potential is given by (15.5). We will analyse the system both from the Lagrangian standpoint and the Hamiltonian standpoint.

Case 1: Lagrangian approach

The Lagrangian for the symmetric spinning top is given by (15.10) and (15.5):

$$\begin{aligned} L &= T - V \\ &= \frac{1}{2}I_1\dot{\varphi}^2 \sin^2 \theta + \frac{1}{2}I_1\dot{\theta}^2 + \frac{1}{2}I_3(\dot{\psi} + \dot{\varphi} \cos \theta)^2 - MgR \cos \theta \end{aligned} \quad (15.11)$$

Note that $\frac{\partial L}{\partial \varphi} = \frac{\partial L}{\partial \psi} = 0$ implying that both φ and ψ are ignorable coordinates. We will return to that point shortly. Meanwhile on with the Euler-Lagrange equations of motion for θ , φ and ψ . We obtain (recall $\frac{d}{dt} \frac{\partial L}{\partial \dot{\theta}} - \frac{\partial L}{\partial \theta} = 0$ for θ etc...)

$$\theta : \frac{d}{dt}(I_1\dot{\theta}) = I_1\dot{\varphi}^2 \sin \theta \cos \theta - I_3(\dot{\psi} + \dot{\varphi} \cos \theta)\dot{\varphi} \sin \theta + MgR \sin \theta \quad (15.12)$$

$$\varphi : \frac{d}{dt}p_\varphi = \frac{d}{dt}(I_1\dot{\varphi} \sin^2 \theta + I_3(\dot{\psi} + \dot{\varphi} \cos \theta) \cos \theta) = 0 \quad (15.13)$$

$$\psi : \frac{d}{dt}p_\psi = \frac{d}{dt}(I_3(\dot{\psi} + \dot{\varphi} \cos \theta)) = 0 \quad (15.14)$$

From (15.14) we obtain $\omega_3 \equiv \dot{\psi} + \dot{\varphi} \cos \theta = \text{const}$ which implies that the component of the angular velocity ω_3 about the axis of symmetry \hat{e}_3 is constant.

We now find the conditions under which a steady precession at a *constant* angle θ to the vertical \underline{k} can occur. From (15.13) and (15.14) we see that the condition $\theta = \text{const}$ leads to both $\dot{\varphi}$ and $\dot{\psi}$ being constant. This in turn implies that the axis of the top precesses around the vertical with constant angular velocity $\dot{\varphi} = \Omega$ say as indicated in Figure. (15.1).

What we now want to obtain is the relation between the angular velocities Ω and ω_3 . From (15.12), assuming the axis is not vertical, i.e. $\sin \theta \neq 0$, then for constant θ we see

$$I_1 \Omega^2 \cos \theta - I_3 \omega_3 \Omega + MgR = 0$$

hence

$$\Omega = \frac{I_3 \omega_3 \pm [I_3^2 \omega_3^2 - 4I_1 MgR \cos \theta]^{\frac{1}{2}}}{2I_1 \cos \theta} \quad (15.15)$$

The minimum value of ω_3 for which real roots for Ω exist is therefore

$$\omega_3^{\min} = \frac{[4I_1 MgR \cos \theta]^{\frac{1}{2}}}{I_3}. \quad (15.16)$$

Considering the limits of ω_3 we see that:

1. for $0 \leq \theta \leq \frac{\pi}{2}$ there is a minimum value of the rotational angular velocity ω_3 about the axis.
2. if the top spins with $\omega_3 < \omega_3^{\min}$, it will start to wobble.
3. if the top spins with $\omega_3 > \omega_3^{\min}$, then from (15.15) we see that there are two possible values of Ω .
4. if the top spins with $\omega_3 \gg \omega_3^{\min}$, then from (15.15) the two solutions are

$$\Omega_1 \simeq \frac{I_3 \omega_3}{I_1 \cos \theta}, \quad \Omega_2 \simeq \frac{MgR}{I_3 \omega_3} \ll \omega_3 \quad (15.17)$$

Ω_1 is of order ω_3 and corresponds to a rapid precessional motion in which the gravitational force is negligible. It is the case of free precessional motion of a rigid body.

Let's consider the case of $\theta > \frac{\pi}{2}$. This corresponds to the situation where the top is suspended below its point of support, and we see that *steady* precession is possible for *any* value of ω_3 because in (15.15) the term inside the square root is positive definite (recall $\cos \theta < 0$ for $\frac{\pi}{2} < \theta < \frac{3\pi}{2}$). Remarkably, and somewhat counter intuitively *steady* precession can even occur when $\omega_3 = 0$. From (15.15) we see that in that case

$$\Omega = \sqrt{\frac{MgR}{I_1 |\cos \theta|}}.$$

Case 2: Hamiltonian approach

Starting with the definition of the Lagrangian for the spinning top in (15.11) and recalling the definition of the generalised momenta in (12.3) we can write

$$p_\varphi = I_1 \dot{\varphi} \sin^2 \theta + I_3 (\dot{\psi} + \dot{\varphi} \cos \theta) \cos \theta \quad (15.18)$$

$$p_\theta = I_1 \dot{\theta} \quad (15.19)$$

$$p_\psi = I_3 (\dot{\psi} + \dot{\varphi} \cos \theta) \quad (15.20)$$

Solving these three equations for $(\dot{\theta}, \dot{\varphi}, \dot{\psi})$ we obtain

$$\dot{\varphi} = \frac{(p_\varphi - p_\psi \cos \theta)}{I_1 \sin^2 \theta} \quad (15.21)$$

$$\dot{\theta} = \frac{p_\theta}{I_1} \quad (15.22)$$

$$\dot{\psi} = \frac{p_\psi}{I_3} - \frac{(p_\varphi - p_\psi \cos \theta)}{I_1 \sin^2 \theta} \cos \theta \quad (15.23)$$

Given the definition of the Hamiltonian Function in (14.7),

$$H(q, p) = \sum_{\beta=1}^n p_\beta \dot{q}_\beta(q, p) - L(q, \dot{q}(q, p))$$

and using (15.21)-(15.23) to replace the velocities $(\dot{\theta}, \dot{\varphi}, \dot{\psi})$ in (15.18)-(15.20) and (15.11) we obtain, after quite a bit of algebra,

$$H = \frac{(p_\varphi - p_\psi \cos \theta)^2}{2I_1 \sin^2 \theta} + \frac{p_\theta^2}{2I_1} + \frac{p_\psi^2}{2I_3} + MgR \cos \theta \quad (15.24)$$

This is the same of course as using $H = T + V$. It is also easy to see that the first set of Hamilton's Equations (14.9) reproduce (15.21)-(15.23). From the form of the Hamilton (15.24) it is clear that φ and ψ are *ignorable* implying that

$$p_\varphi = \text{const} \quad p_\psi = \text{const}.$$

This allows us to reduce the system to one of one degree of freedom only, the coordinate θ being that degree of freedom. Doing this allows us to write the Hamiltonian (15.24) in the suggestive form

$$H = \frac{p_\theta^2}{2I_1} + U(\theta) \quad (15.25)$$

where

$$U(\theta) = \frac{(p_\varphi - p_\psi \cos \theta)^2}{2I_1 \sin^2 \theta} + \frac{p_\psi^2}{2I_3} + MgR \cos \theta. \quad (15.26)$$

We now proceed to investigate this one dimensional system.

15.4 General motion of the symmetric top

Hamilton's equations (14.9) and (14.11) for the the coordinates θ and generalised momentum p_θ are

$$\dot{\theta} = \frac{\partial H}{\partial p_\theta} = \frac{p_\theta}{I_1} \quad (15.27)$$

$$-\dot{p}_\theta = \frac{\partial H}{\partial \theta} = \frac{\partial U}{\partial \theta} \quad (15.28)$$

Recalling I_1 is constant then we can differentiate (15.27) to obtain

$$I_1 \ddot{\theta} = \dot{p}_\theta = -\frac{\partial H}{\partial \theta} = -\frac{\partial U(\theta)}{\partial \theta}. \quad (15.29)$$

This is difficult to solve, but we can glean information and obtain the qualitative features of motion from energy conservation. Note that the Hamiltonian (15.25) is independent of t , i.e. $\frac{\partial H}{\partial t} = 0$, hence $H = E = \text{const}$. We begin now to look at various limiting cases in order to understand the general features of the spinning top.

Case a: $\dot{\theta} = 0$ i.e. $p_\theta = 0$

In that case we see from (15.25) that θ is such that $U(\theta) = E$.

Case b: Motion confined to region where $U(\theta) \leq E$.

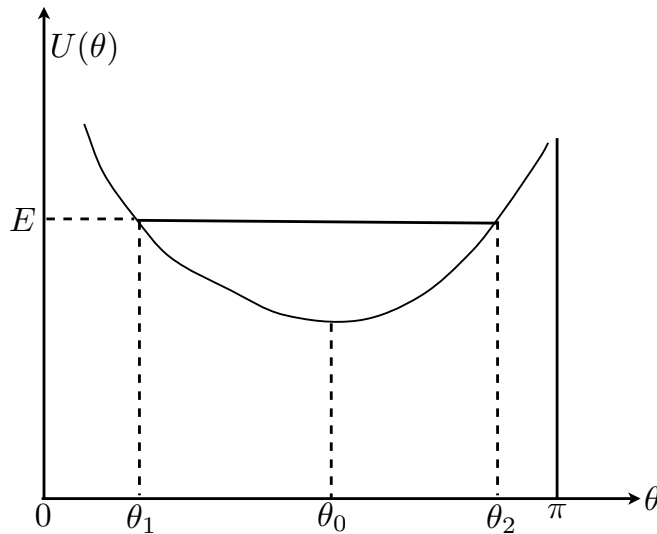


Figure 15.5: Typical shape of $U(\theta)$ with one min at θ_0 and tending to infinity as $\theta \rightarrow 0, \pi$.

Considering the shape of $U(\theta)$, to start with we will assume $p_\varphi \neq \pm p_\psi$, although we will relax this later as the condition $p_\varphi = \pm p_\psi$ is an important special case. From (15.26) it is

clear that $U(\theta) \rightarrow \infty$ as $\theta \rightarrow 0$ and π . It is also possible to show that there is only one minimum $U_{\min}(\theta_0)$ at θ_0 , such that $\frac{dU}{d\theta} = 0$. When the energy $E = U_{\min}(\theta_0)$ there exists an equilibrium situation and θ remains fixed at $\theta = \theta_0$. It corresponds to a *steady precession*. For $E > U_{\min}(\theta_0)$ it is clear from the figure that θ oscillates between a minimum value θ_1 and a maximum value θ_2 .

15.5 Describing the motion of the top

From (15.21) we see that the condition for the angular velocity of the axis about the vertical to vanish (i.e. $\dot{\varphi} = 0$) is

$$\cos \theta_{\text{vanish}} = \frac{p_\varphi}{p_\psi} \quad (15.30)$$

Case 1: If this θ_{vanish} lies outside the range between $\theta_1 \rightarrow \theta_2$ (or if $|\frac{p_\varphi}{p_\psi}| > 1$ it implies that $\dot{\varphi}$ will never vanish and the axis precesses around the vertical in a fixed direction and wobbles up and down between θ_1 and θ_2 as indicated in the beautifully accurate diagram that is Figure. (15.6).

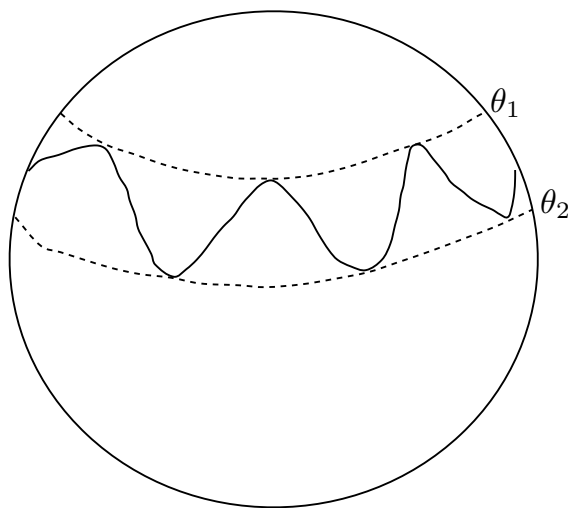


Figure 15.6: Case 1: Track of the end of the spinning top oscillating between between $\theta_1 \rightarrow \theta_2$ as seen on the sphere.

Case 2: $\theta_1 < \cos^{-1}(\frac{p_\varphi}{p_\psi}) < \theta_2$. In this situation θ_{vanish} lies within the range implying $\dot{\varphi} \rightarrow 0$ as $\theta \rightarrow \theta_{\text{vanish}}$. The effect is that the axis of symmetry now moves in loops with the angular velocity having one sign near the top of the loop and the opposite near the bottom, as seen in the even better Fig. (15.7).

Case 3: Limiting case where $\theta_1 = \cos^{-1}(\frac{p_\varphi}{p_\psi}) = \theta_{\text{vanish}}$. in this situation the loop shrinks

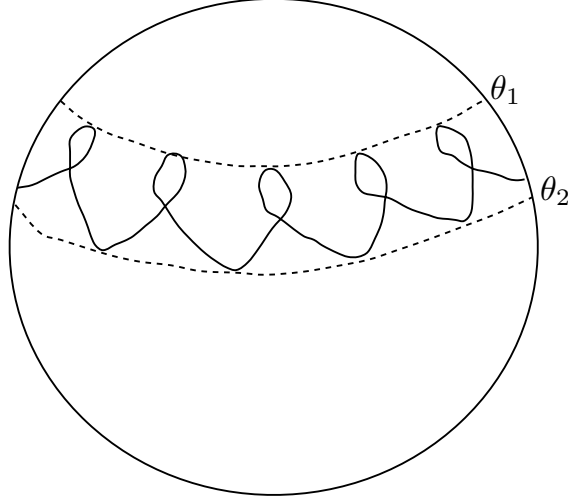


Figure 15.7: Case 2: Track of the end of the spinning top oscillating between $\theta_1 \rightarrow \theta_2$ as seen on the sphere.

to cusps where the axis of the top comes instantaneously to rest at the top of each loop as seen in the Picasso like Figure. (15.8). It occurs if the top is set spinning with its axis initially at rest. Note that there can be no cusped motion with cusps at θ_2 because they correspond to points of the minimum kinetic energy and motion must always be below such points.

15.6 Stability of a vertical top

Imagine the case where we set the top spinning with angular velocity ω_3 with its axis vertical. How will it move? Now if the axis of the top moves through the vertical $\theta = 0$, then it follows that clear $U(0)$ must be finite. Looking at the form of $U(\theta)$ in (15.26), the *only* way that can be the case is if $p_\varphi = p_\psi = I_3 \omega_3$, the final equality coming from (15.23) and the fact $\dot{\psi} = \omega_3$. Note we could have consider motion where the axis passes through the downward vertical $\theta = \pi$ which would have led to $p_\varphi = -p_\psi = -I_3 \omega_3$, but such a downward pointing top is always stable and so not as interesting. Returning to the upwards pointing case we set $p_\varphi = p_\psi = I_3 \omega_3$ in (15.26) to obtain

$$U(\theta) = \frac{I_3^2 \omega_3^2}{2I_1} \tan^2 \frac{\theta}{2} + \frac{1}{2} I_3 \omega_3^2 + MgR \cos \theta \quad (15.31)$$

where we have used the identity $\frac{1-\cos \theta}{\sin \theta} = \frac{2 \sin^2(\theta/2)}{2 \sin(\theta/2) \cos(\theta/2)} = \tan(\theta/2)$.

Consider the case of small θ , i.e. $\theta \ll 1$. We can expand then $\tan^2(\theta/2)$ and $\cos \theta$ in

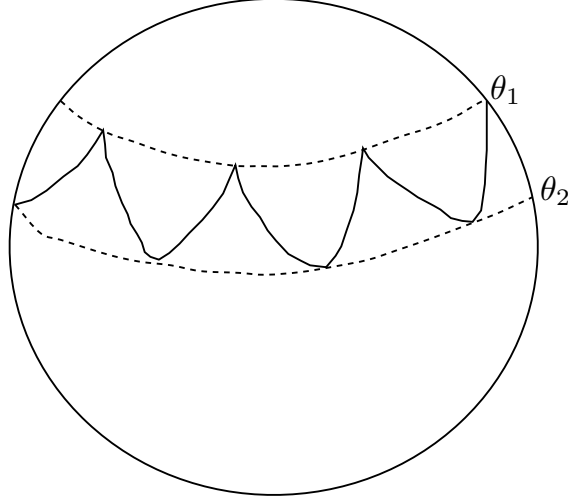


Figure 15.8: Case 3: Track of the end of the spinning top oscillating between between $\theta_1 \rightarrow \theta_2$ as seen on the sphere.

(15.31) to $O(\theta^2)$ to get

$$U(\theta) \simeq \left(\frac{1}{2} I_3 \omega_3^2 + MgR \right) + \frac{1}{2} \left(\frac{I_3^2 \omega_3^2}{4I_1} - MgR \right) \theta^2 \quad (15.32)$$

Given the quadratic nature of (15.32) we see that $\theta = 0$ is a stable equilibrium position, as long as $U(\theta)$ has a minimum at $\theta = 0$. For that to occur the coefficient of the θ^2 term must be positive. If it were negative then $\theta = 0$ would be an unstable equilibrium point. Looking at the θ^2 term we see that there exists a minimum value ω_3 for which the vertical top is stable. That value is given by

$$\omega_3^2 = \frac{4MgRI_1}{I_3^2} = \omega_0^2. \quad (15.33)$$

If the top is set spinning with $\omega_3 > \omega_0$ it will *remain vertical*. When ω_3 falls below ω_0 because of say friction, the top will then begin to wobble.

Lets consider the energy of the vertical top given by (15.25), $E = \frac{p_\theta^2}{2I_1} + U(\theta)$ where $\theta = \dot{\theta} = 0$. in that case from (15.19) and (15.31),

$$E = \frac{1}{2} I_3 \omega_3^2 + MgR \quad (15.34)$$

More generally, from (15.25) it is clear that the angles at which $\dot{\theta} = 0$ are given by solutions of $U(\theta) = E$. Equating (15.31) and (15.34) the condition $U(\theta) = E$ is

$$\frac{I_3^2 \omega_3^2}{2I_1} \tan^2 \frac{\theta}{2} + \frac{1}{2} I_3 \omega_3^2 + MgR \cos \theta = \frac{1}{2} I_3 \omega_3^2 + MgR.$$

Rearranging we have

$$\tan^2 \frac{\theta}{2} \left(\frac{I_3^2 \omega_3^2}{2I_1} - 2MgR \cos^2 \frac{\theta}{2} \right) = 0 \quad (15.35)$$

The two solutions are $\theta = \theta_1 = 0$ or

$$\cos^2 \frac{\theta_2}{2} = \frac{I_3^2 \omega_3^2}{4I_1 MgR} = \frac{\omega_3^2}{\omega_0^2} \quad (15.36)$$

and are shown schematically in Figure. (15.9). Inverting we see that the condition $\dot{\theta} = 0$ leads to $\theta_1 = 0$ and $\theta_2 = 2 \cos^{-1}(\omega_3/\omega_0)$. If the top is set spinning with its axis vertical and

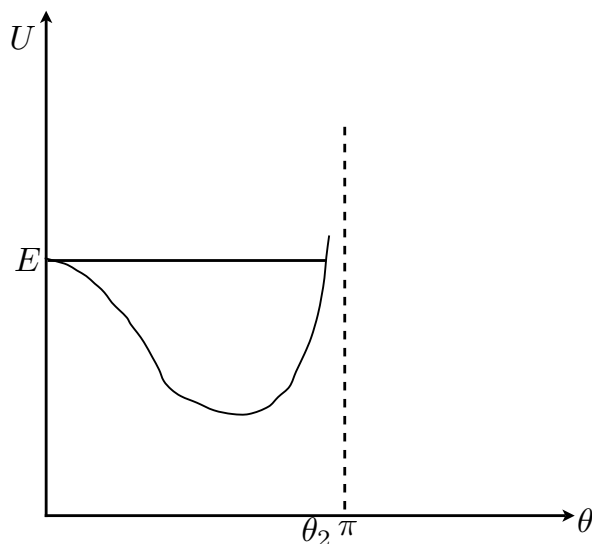


Figure 15.9: Energy of vertical top when $\dot{\theta} = 0$.

almost stationary with $\omega_3 < \omega_0$ it will oscillate in subsequent motion between the vertical and angle θ_2 . Note that θ_2 increases as ω_3 decreases, tending to π as $\omega_3 \rightarrow 0$. When $\omega_3 = 0$, the top behaves like a compound pendulum and swings in a circle through both the upward and downward verticals.