

Chapter 1

Introduction

1.1 What is classical mechanics?

“Classical mechanics” can mean quite different things in different contexts. It is often used to refer to Newtonian mechanics, including electrodynamics, as opposed to the “modern physics” developed in the 20th century, ie relativity and quantum mechanics. It is also, perhaps more commonly, used as an antonym to quantum mechanics: “classical” is everything that is not “quantum”. In this sense of the word, general relativity is also considered to be “classical mechanics”, or at least classical physics.

Although the methods introduced in this course can easily be extended to both special and general relativity, we will not discuss that here: the focus will remain on Newtonian mechanics. However, the use of the term “Classical mechanics” as opposed to merely “Mechanics” denotes a shift in emphasis, from Newtonian forces and acceleration to the more general and abstract formulations that were developed in the late 18th and the 19th century, associated with names like Euler, Lagrange, Hamilton and Jacobi. Therefore, this course is *not* more of the stuff you have already studied in modules like MP110, MP112 and MP205, but instead represents a completely new way of looking at mechanics, and one which forms the foundation of nearly all modern mathematical physics.

The focus in this course is on methods and formulations rather than on answers or numbers. In part, this is because the key to solving complicated problems is very often to formulate them properly and to select appropriate methods. However, there are other reasons for this shift in focus:

- Often, we are not that interested in numerical solutions, but more in the qualitative features of a system, and we can find out a lot about this without doing any numerical calculations.
- We will see that wildly different physical systems can look identical from a mathematical point of view, so solving one can immediately give us the solution to the other. Starting with numerical calculations can obscure this.

- *Symmetries* will play an extremely important role, and we will learn to identify and exploit symmetries to simplify and understand mechanical systems. Putting in numbers at the start will often hide the symmetries.

The Lagrange–Hamilton formalism and the symmetry principles which we will become acquainted with here, are use all throughout modern physics:

- quantum mechanics;
- statistical mechanics;
- condensed matter theory (quantum statistical mechanics)
- classical field theory (electromagnetism, general relativity)
- particle physics (quantum field theory and symmetry groups)
- chaos theory
- etc

1.2 Overview

The module will cover the following topics:

- The principle of least action (Hamilton’s principle), the lagrangian and the Euler–Lagrange equations.
- Generalised coordinates (how to formulate a mechanical problem in the most sensible way given symmetries and constraints).
- Canonical momenta and conservation laws; energy conservation.
- Hamilton’s equations of motion.
- Poisson brackets.
- Central force motion, angular momentum conservation.
- Planetary motion, Kepler’s laws.
- Rotations and rotation matrices.
- Inertia tensor, principal axes of inertia.
- Euler’s equations of (rotational) motion.

Learning outcomes

At the end of this course, you should be able to:

- formulate the basic principles of the Lagrange–Hamilton formalism;
- use these principles to derive equations of motion for dynamical systems;
- explain the relation between symmetries and conservation laws;
- apply conservation laws to analyse the motion of dynamical systems; and
- describe the mathematical properties of rotations and systems with rotational symmetry.

Chapter 2

Lagrangian mechanics

2.1 The principle of least action

The starting point for the reformulation of classical mechanics is the *principle of least action*, which may be somewhat flippantly paraphrased as “The world is lazy”, or in the more flowery words of Pierre Louis Maupertuis (1744): *Nature is thrifty in all its actions*:

The laws of movement and of rest deduced from this principle being precisely the same as those observed in nature, we can admire the application of it to all phenomena. The movement of animals, the vegetative growth of plants . . . are only its consequences; and the spectacle of the universe becomes so much the grander, so much more beautiful, the worthier of its Author, when one knows that a small number of laws, most wisely established, suffice for all movements.

This very general formulation does not in itself have any predictive power, but the idea that nature’s “thrift” could be used to derive laws of motion had already been successfully applied in optics for a long time:

Fermat’s principle

The path take between two points by a ray of light is the path that can be traversed in the least time.

This principle was first formulated by Ibn al-Haytham (aka Alhacen) in his *Book of Optics* from 1021, which formed one of the main foundations of geometric optics and the scientific method in general. He proved that it led to the law of reflection. It was restated by Pierre de Fermat in 1662, who also derived Snell’s law of refraction from this principle.

2.1.1 Hamilton's principle

In *mechanics* the proper mathematical formulation of Maupertuis' principle is due to William Rowan Hamilton¹, building on earlier work by Joseph Louis Lagrange.

We will denote the kinetic and potential energy of a particle, or of a mechanical system in general, as

$$\begin{aligned} T &= \text{kinetic energy} \\ V &= \text{potential energy} \end{aligned}$$

T usually depends on the *velocities* $v_i = \frac{dx_i}{dt} \equiv \dot{x}_i$ $T = T(\dot{x}_i)$
 V usually depends on the *positions* x_i $V = V(x_i)$
 but may also depend explicitly on time $V = V(x_i, t)$
 (for example with time-varying external forces).

x_i and \dot{x}_i here denote *all* the coordinates and their time derivatives. So for example we have

$$\begin{aligned} x_i &\rightarrow x && \text{for a single particle in one dimension} \\ x_i &\rightarrow \{x, y, z\} && \text{for a particle in three dimensions} \\ x_i &\rightarrow \{x_1, y_1, z_1, x_2, y_2, z_2, \dots, x_N, y_N, z_N\} && \text{for } N \text{ particles in three dimensions} \end{aligned}$$

We now define the **lagrangian** L as the difference between kinetic and potential energy,

$$L(x_i, \dot{x}_i, t) = T - V. \tag{2.1}$$

Note that L will be a function of the coordinates x_i , the velocities \dot{x}_i , and the time t , although in many cases there is no explicit time dependence; ie, if we know the positions and velocities of all the particles we know the lagrangian.

A particular **path** is given by specifying the coordinates x_i as a function of time, $x_i = x_i(t)$. (Note that if $x_i(t)$ is known, its derivative $\dot{x}_i(t)$ is also known.) For a given path, the **action** S is defined as

$$S[x] \equiv \int_{t_1}^{t_2} L(x(t), \dot{x}(t), t) dt. \tag{2.2}$$

We are now in a position to formulate Hamilton's principle of least action.

¹On a General Method in Dynamics, Phil. Trans. Roy. Soc. (1834) 247; (1835) 95.

The principle of least action:

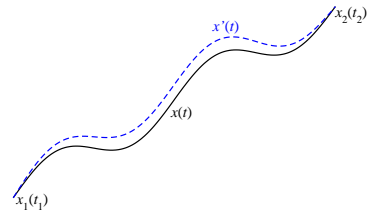
The physical path a system will take between two points in a certain time interval is the one that gives the smallest action S .

Comments:

1. The potential energy V is defined only for conservative forces, so the action as it is written here is defined only for conservative forces. It is possible to generalise this to certain non-conservative forces and obtain the correct equations of motion (we will see examples of this later). However, *all* microscopic (fundamental) forces are conservative.
2. The action S is a “function of a function” since it depends on the function(s) $x_i(t)$. We call this a *functional*, and denote it by putting the function argument in square brackets, $S = S[x]$.

2.2 The Euler–Lagrange equations

What does ‘the path that gives the smallest action’ actually mean, and how can we find it? To work this out, let us consider a path $x(t)$ and another path $x'(t) = x(t) + \alpha h(t)$, where $h(t)$ is some arbitrary smooth function of t , and α is a parameter that we will vary.



Since we are looking for the path the system will take between two specific points in a specific time interval, the endpoints of the two paths must be the same. We therefore have

$$x(t_1) = x'(t_1) = x_1; \quad x(t_2) = x'(t_2) = x_2 \quad \iff \quad h(t_1) = h(t_2) = 0. \quad (2.3)$$

We can now write $S[x + \alpha h] = S(\alpha)$, and treat it as a function of the parameter α . For a given $h(t)$, the minimum of S will occur when $\frac{dS}{d\alpha} = 0$.

This allows us to restate the principle of least action:

For any smooth $h_i(t)$ with $h_i(t_1) = h_i(t_2) = 0$, the physical path $x_i(t)$ is such that

$$\frac{d}{d\alpha} S[x + \alpha h] = \frac{d}{d\alpha} \int_{t_1}^{t_2} L(x_i + \alpha h_i, \dot{x}_i + \alpha \dot{h}_i, t) dt = 0. \quad (2.4)$$

We often use the shorthands $\alpha h = \delta x$, $S[x + \delta x] - S[x] = \delta S =$ the *variation* of S , and call $\frac{\delta S}{\delta x}$ the *functional derivative* of S . The principle of least action is then often written as

$$\delta S = 0 \quad \text{or} \quad \frac{\delta S}{\delta x} = 0 \quad \iff \quad \frac{d}{d\alpha} S[x + \alpha h] = 0 \quad \text{for any } h(t). \quad (2.5)$$

Let us now calculate the variation δS . For a single particle in one dimension, we have

$$\frac{d}{d\alpha} S[x + \alpha h] = \frac{d}{d\alpha} \int_{t_1}^{t_2} L(x + \alpha h, \dot{x} + \alpha \dot{h}, t) \quad (2.6)$$

$$= \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial x} h + \frac{\partial L}{\partial \dot{x}} \dot{h} \right) dt \quad (2.7)$$

$$= \int_{t_1}^{t_2} \frac{\partial L}{\partial x} h dt + \left[\frac{\partial L}{\partial \dot{x}} h \right]_{t=t_1}^{t=t_2} - \int_{t_1}^{t_2} \left(\frac{d}{dt} \frac{\partial L}{\partial \dot{x}} \right) h dt \quad (2.8)$$

$$= \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial x} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}} \right) h(t) dt. \quad (2.9)$$

In the first step we used that L is a function of the three variables x, \dot{x}, t , but t does not depend on α . We can then use the chain rule for a function of two variables,

$$\frac{d}{d\alpha} f(x, y) = \frac{\partial f}{\partial x} \frac{dx}{d\alpha} + \frac{\partial f}{\partial y} \frac{dy}{d\alpha}.$$

In the second step we used integration by parts,

$$\int u v dt = uv - \int \dot{u} v dt \quad \text{with} \quad u = \frac{\partial L}{\partial \dot{x}}, \quad v = \dot{h}.$$

In the final step the boundary term vanishes since $h(t_1) = h(t_2) = 0$.

But $h(t)$ is a completely arbitrary smooth function, and we must have $\delta S = 0$ for *any* $h(t)$. This is only possible if the term within the brackets in (2.9) is 0 for all t , ie

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = 0 \quad \textbf{The Euler-Lagrange equation} \quad (2.10)$$

If we have N coordinates x_i , the derivation proceeds following the same steps. Using the chain rule for a function of $2N$ variables, we find

$$\begin{aligned} & \frac{d}{d\alpha} L(x_1 + \alpha h_1, x_2 + \alpha h_2, \dots, x_N + \alpha h_N, \dot{x}_1 + \alpha \dot{h}_1, \dot{x}_2 + \alpha \dot{h}_2, \dots, \dot{x}_N + \alpha \dot{h}_N) \\ &= \frac{\partial L}{\partial x_1} h_1 + \frac{\partial L}{\partial x_2} h_2 + \dots + \frac{\partial L}{\partial x_N} h_N + \frac{\partial L}{\partial \dot{x}_1} \dot{h}_1 + \frac{\partial L}{\partial \dot{x}_2} \dot{h}_2 + \dots + \frac{\partial L}{\partial \dot{x}_N} \dot{h}_N \\ &= \sum_{i=1}^N \left(\frac{\partial L}{\partial x_i} h_i + \frac{\partial L}{\partial \dot{x}_i} \dot{h}_i \right). \end{aligned} \quad (2.11)$$

Using integration by parts on the second term (for each i) gives us

$$\frac{d}{d\alpha} S[x + \alpha h] = \sum_{i=1}^N \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial x_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}_i} \right) h_i(t) dt = 0. \quad (2.12)$$

Since *all* the h_i are *independent*, arbitrary functions, the expression within the brackets must vanish for *each* i :

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{x}_i} - \frac{\partial L}{\partial x_i} = 0 \quad \text{for all } i = 1, \dots, N. \quad (2.13)$$

Example 2.1 Particle in a potential

The kinetic energy of a single particle is

$$T = \frac{1}{2}mv^2 = \frac{1}{2}m(v_x^2 + v_y^2 + v_z^2) = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2). \quad (2.14)$$

We take an arbitrary potential energy $V = V(x, y, z, t)$.

The Euler–Lagrange equations are

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = 0; \quad \frac{d}{dt} \frac{\partial L}{\partial \dot{y}} - \frac{\partial L}{\partial y} = 0; \quad \frac{d}{dt} \frac{\partial L}{\partial \dot{z}} - \frac{\partial L}{\partial z} = 0. \quad (2.15)$$

We find

$$\frac{\partial L}{\partial \dot{x}} = m\dot{x}; \quad \frac{\partial L}{\partial x} = -\frac{\partial V}{\partial x} \quad (2.16)$$

$$\implies \frac{d}{dt} \frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = m\ddot{x} + \frac{\partial V}{\partial x} = 0 \quad \iff \quad m\ddot{x} = -\frac{\partial V}{\partial x}. \quad (2.17)$$

Likewise, we get

$$m\ddot{y} = -\frac{\partial V}{\partial y}, \quad m\ddot{z} = -\frac{\partial V}{\partial z} \quad \text{or}$$

$$\underline{m\ddot{\vec{r}} = m\vec{a} = -\nabla V = \vec{F}} \quad = \text{Newton's 2}^{\text{nd}} \text{ law!}$$

So the EL equations are exactly equivalent to Newton's laws.

Example 2.2 The shortest path between two points

In deriving the Euler–Lagrange equations we did not make any use of the definition of L : it could be *any* function of x, \dot{x} and t . Therefore, the EL equations give the stationary points for *any* functional of the path between two points — for example, the length of the path!

Consider a curve $y = y(x)$ between two points (x_1, y_1) and (x_2, y_2) . The length ds of an infinitesimal segment (dx, dy) of this curve is given by Pythagoras:

$$ds^2 = dx^2 + dy^2 = dx^2 + (y'(x)dx)^2 = (1 + y'(x)^2)dx^2 \quad (2.18)$$

$$\implies ds = \sqrt{1 + y'(x)^2}dx. \quad (2.19)$$

If, to make life simpler for ourselves, we assume that x is monotonically increasing

along the curve, we find that the total length of the curve is

$$S = \int_{x_1}^{x_2} \sqrt{1 + y'(x)^2} dx = \int_{x_1}^{x_2} L(y(x), y'(x), x) dx \quad \text{with} \quad L = \sqrt{1 + y'^2}. \quad (2.20)$$

This looks like what we had before, but with $t \rightarrow x$; $x(t) \rightarrow y(x)$; $\dot{x}(t) \rightarrow y'(x)$.

The Euler–Lagrange equation becomes

$$\frac{d}{dt} \frac{\partial L}{\partial y'} - \frac{\partial L}{\partial y} = 0. \quad (2.21)$$

We see immediately that $\partial L / \partial y = 0$. To find $\partial L / \partial y'$ we use the chain rule,

$$\begin{aligned} \frac{\partial L}{\partial y'} &= \frac{dL}{du} \frac{du}{dy'} \quad \text{with} \quad u = 1 + y'^2; \quad L = \sqrt{u} \\ \implies \frac{\partial L}{\partial y'} &= \frac{1}{2\sqrt{1 + y'^2}} \cdot 2y' = \frac{y'}{\sqrt{1 + y'^2}}. \end{aligned}$$

To find $\frac{d}{dx} \frac{\partial L}{\partial y'}$ we use the product rule and the chain rule:

$$\begin{aligned} \frac{\partial L}{\partial y'} &= vw \quad \text{with} \quad v = y', \quad w = \frac{1}{\sqrt{1 + y'^2}} = u^{-1/2} \\ \implies \frac{d}{dx} \frac{\partial L}{\partial y'} &= \frac{dv}{dx} w + v \frac{dw}{du} \frac{du}{dx} = \frac{dy'}{dx} w + y' \frac{dw}{du} \frac{du}{dy'} \frac{dy'}{dx} \\ &= y'' \frac{1}{\sqrt{1 + y'^2}} + y' \cdot \left(-\frac{1}{2} u^{-3/2}\right) \cdot 2y' \cdot y'' \\ &= y'' \left(\frac{1}{\sqrt{1 + y'^2}} - \frac{y'^2}{(1 + y'^2)^{3/2}} \right) = \frac{y''}{\sqrt{1 + y'^2}} \left(1 - \frac{y'^2}{1 + y'^2} \right) \\ &= \frac{y''}{\sqrt{1 + y'^2}} \frac{1}{1 + y'^2}. \end{aligned} \quad (2.23)$$

$$\text{So} \quad \frac{d}{dx} \frac{\partial L}{\partial y'} = 0 \quad \implies \quad y''(x) = 0 \quad \implies \quad y(x) = Ax + B. \quad (2.24)$$

This describes a straight line, so we have shown that the shortest path between two points is a straight line!

So what is the point?

1. The equations are often *easier*: We get rid of complicated vectors and forces, and derive everything from scalars (energy).
2. It is easier to generalise to systems with constraints.

3. We can choose whichever coordinates we want.
4. The lagrangian formalism can be generalised to quantum mechanics (in the Feynman formulation: all paths are possible, but weighted by the action) and field theory (with infinitely many degrees of freedom).

We will look at points 2 and 3 next.

2.3 Generalised coordinates

It is often advantageous to change variables from the cartesian coordinates $\{x_i, y_i, z_i\}$ for each particle $i = 1, \dots, N$ to some other variables $\{q_j\}, j = 1, \dots, n$. These are called *generalised coordinates*.

Consider for example a system of N particles. We need $3N$ independent coordinates to describe the system completely: we say that there are $3N$ *degrees of freedom*.

Now, imagine that there is a *constraint* relating the $3N$ coordinates, for example:

1. Two particles are tied together with a rod of length L , so that

$$(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 = L^2. \quad (2.25)$$

2. The N particles are all moving on the surface of a sphere, ie

$$x_i^2 + y_i^2 + z_i^2 = R^2 \quad \forall i = 1, \dots, N. \quad (2.26)$$

3. A ball in a squash court, $x \geq 0, z \geq 0$.

The first two of these can be described by M equations of the form

$$f_j(\vec{x}_1, \dots, \vec{x}_N, t) = 0, \quad j = 1, \dots, M. \quad (2.27)$$

Such constraints are called *holonomic* (or integrable) constraints, and we will focus only on such constraints in the following. With such constraint equations, the coordinates x_i, y_i, z_i are no longer independent. Instead we have

$$M \text{ relations} \quad \implies \quad n = 3N - M \text{ real degrees of freedom.}$$

By choosing n suitable generalised coordinates to describe these degrees of freedom, we achieve two things:

- We eliminate the forces of constraints which are required in the newtonian formulation. No net work is done by these forces, so they can safely be eliminated.

- The Euler–Lagrange equations look exactly the same in the new coordinates, so the problem is no more difficult (and probably easier) than the original one.

In the first example above, the constraint (2.25) reduces the number of degrees of freedom from 6 to 5. The 5 coordinates can for example be chosen to be the centre of mass coordinates X, Y, Z for the two particles, and two angles θ, ϕ that describe the orientation of the rod.²

In the second example, each particle is described by 2 instead of 3 coordinates. These can be chosen to be the ‘latitude’ θ and ‘longitude’ ϕ of each particle (corresponding to spherical coordinates, see Sec. 2.3.1).

Example 2.3 Simple pendulum

Consider a simple pendulum with length ℓ , mass m in a constant gravitational field g (see Fig. 2.1).

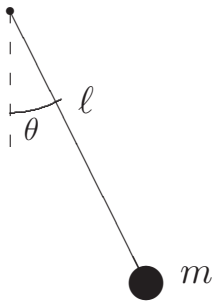


Figure 2.1: A simple pendulum

Here it is convenient to choose the angle θ as our coordinate. The x (horizontal) and z (vertical) coordinates and their time derivatives can be written in terms of θ as

$$x = \ell \sin \theta \qquad \dot{x} = \ell \dot{\theta} \cos \theta, \qquad (2.28)$$

$$z = -\ell \cos \theta \qquad \dot{z} = \ell \dot{\theta} \sin \theta. \qquad (2.29)$$

The kinetic energy is

$$\begin{aligned} T &= \frac{1}{2} m \vec{v}^2 = \frac{1}{2} m (\dot{x}^2 + \dot{z}^2) \\ &= \frac{1}{2} m \ell^2 \dot{\theta}^2 (\cos^2 \theta + \sin^2 \theta) = \frac{1}{2} m \ell^2 \dot{\theta}^2. \end{aligned} \qquad (2.30)$$

The potential energy is

$$V = mgz = -mg\ell \cos \theta. \qquad (2.31)$$

The lagrangian therefore becomes

$$L = T - V = \frac{1}{2} m \ell^2 \dot{\theta}^2 + mg\ell \cos \theta. \qquad (2.32)$$

The Euler–Lagrange equation is

$$\frac{\partial L}{\partial \theta} = \frac{d}{dt} \frac{\partial L}{\partial \dot{\theta}} \implies -mg\ell \sin \theta = \frac{d}{dt} (m \ell^2 \dot{\theta}) \qquad (2.33)$$

$$\implies \ddot{\theta} = -\frac{g}{\ell} \sin \theta. \qquad (2.34)$$

This is the equation of motion for the pendulum.

²In Chapter 5 we will look more at how these angles can be chosen.

Once we have found the equation of motion for θ , and the solution to this equation, we can go back and calculate x and z as functions of time. However, in the example of the simple pendulum, we are not usually interested in this.

We note that the mass m does not appear in the equation of motion. We could have seen this already by inspecting the lagrangian: the EL equations are unchanged if the lagrangian is multiplied by an overall constant α , $L \rightarrow \alpha L$. In this case, since the mass just enters as an overall factor in the lagrangian, the EL equation will not depend on the mass.

Solutions to the equations of motion?

Now we have found the equation of motion for the simple pendulum, and we may want to know the solutions to this equation, ie what the actual motion of the pendulum is for different initial conditions. It is actually possible to integrate the equation (2.34) and write down a solution, but this involves elliptic integrals and lots of other complicated maths, and will not help us to understand the physical system. It will be more useful to find numerical solutions, and in Computational Physics MP354 we will learn how this can be done.

What we can do to understand the system better, is

- look at the general types of solutions we may have. We will do this when we discuss conservation of energy;
- consider limiting cases such as *small oscillations*. This is what we will do now.

If θ is small, we may approximate $\sin \theta$ with the first term in its power expansion (Taylor expansion),

$$\sin \theta = \theta - \frac{1}{3!}\theta^3 + \frac{1}{5!}\theta^5 + \dots \approx \theta. \quad (2.35)$$

In that case (2.34) simplifies to

$$\ddot{\theta} = -\frac{g}{\ell}\theta. \quad (2.36)$$

We recognise this as the equation for a simple harmonic oscillator, $\ddot{x} + \omega^2 x = 0$, with $x \rightarrow \theta, \omega^2 \rightarrow g/\ell$. We therefore see that for small oscillations, the simple pendulum behaves as a simple harmonic oscillator with frequency $\omega_s = \sqrt{g/\ell}$, ie the frequency is inversely proportional to the square root of the length of the pendulum (and independent of the mass).

Example 2.4 Double Atwood machine

Consider the double Atwood machine in Fig. 2.2. We assume that

1. the pulleys are light, so we can ignore their kinetic energy; and
2. the ropes do not slip (or they slide without friction).

Here we have *two* independent degrees of freedom, which we can choose to be x and y . In terms of these, the positions of the three blocks are

$$\begin{aligned}x_1 &= -x, \\x_2 &= -(\ell_1 - x + y), \\x_3 &= -(\ell_1 - x + \ell_2 - y).\end{aligned}$$

The kinetic and potential energy of the three blocks are

$$\begin{aligned}T_1 &= \frac{1}{2}m_1\dot{x}_1^2 \\T_2 &= \frac{1}{2}m_2\left[\frac{d}{dt}(\ell_1 - x - y)\right]^2 = \frac{1}{2}m_2(\dot{y} - \dot{x})^2 \\T_3 &= \frac{1}{2}m_3(\dot{x} + \dot{y})^2 \\V_1 &= -mgx \\V_2 &= -m_2g(\ell_1 - x + y) \\V_3 &= -m_2g(\ell_1 + \ell_2 - x - y)\end{aligned}$$

The lagrangian becomes

$$\begin{aligned}L &= \frac{1}{2}(m_1 + m_2 + m_3)\dot{x}^2 + \frac{1}{2}(m_1 + m_2)\dot{y}^2 + (m_3 - m_2)\dot{x}\dot{y} \\&\quad + (m_1 - m_2 - m_3)gx + (m_2 - m_3)gy + m_2g\ell_2 + m_3g(\ell_1 + \ell_2).\end{aligned}$$

Note that the last two terms are constants which do not play any role in the equations of motion. We get two equations of motion:

$$\begin{aligned}\frac{d}{dt}\frac{\partial L}{\partial \dot{x}} &= (m_1 + m_2 + m_3)\ddot{x} + (m_3 - m_2)\ddot{y} = \frac{\partial L}{\partial x} = (m_1 - m_2 - m_3)g \\ \frac{d}{dt}\frac{\partial L}{\partial \dot{y}} &= (m_2 + m_3)\ddot{y} + (m_3 - m_2)\ddot{x} = \frac{\partial L}{\partial y} = (m_2 - m_3)g\end{aligned}$$

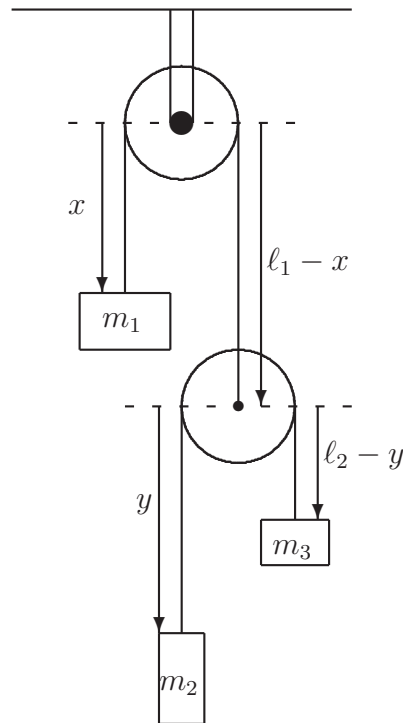


Figure 2.2: Double Atwood machine

Example 2.5 Pendulum with rotating support

Consider a pendulum mounted on the edge of a disc with radius a , rotating with constant angular velocity ω (see Fig. 2.3). If the support point is in the horizontal position at $t = 0$, the angular position of the support point at time t is $\phi = \omega t$, and the cartesian coordinates of the bob at time t are

$$\begin{aligned}x &= a \cos \omega t + \ell \sin \theta \\z &= a \sin \omega t - \ell \cos \theta\end{aligned}$$

giving the velocities

$$\begin{aligned}\dot{x} &= -a\omega \sin \omega t + \ell \dot{\theta} \cos \theta \\ \dot{z} &= a\omega \cos \omega t + \ell \dot{\theta} \sin \theta\end{aligned}$$

This gives us the lagrangian

$$\begin{aligned}L = T - V &= \frac{1}{2}m(\dot{x}^2 + \dot{z}^2) - mgz \\ &= \frac{m}{2}(a^2\omega^2 + \ell^2\dot{\theta}^2 + 2a\omega\ell\dot{\theta} \sin(\theta - \omega t)) - mg(a \sin \omega t - \ell \cos \theta)\end{aligned}$$

This system has only one degree of freedom θ , but the lagrangian depends explicitly on time because of the rotation of the support point. The Euler–Lagrange equation is

$$\begin{aligned}\frac{d}{dt} \frac{\partial L}{\partial \dot{\theta}} &= \frac{d}{dt}(m\ell^2\dot{\theta} + ma\omega\ell \sin(\theta - \omega t)) = m\ell^2\ddot{\theta} + ma\omega\ell(\dot{\theta} - \omega) \cos(\theta - \omega t) \\ &= \frac{\partial L}{\partial \theta} = ma\omega\ell\dot{\theta} \cos(\theta - \omega t) + mg\ell \sin \theta \\ \implies \ell\ddot{\theta} - a\omega^2 \cos(\theta - \omega t) &= g \sin \theta \implies \ddot{\theta} = \frac{a\omega^2}{\ell} \cos(\theta - \omega t) - \frac{g}{\ell} \sin \theta\end{aligned}$$

Finding the equation of motion for this system becomes a bit complicated, but it is still far simpler than it would have been to compute the forces at each point and use Newton’s second law.

We can check that our result is sensible by seeing what happens if there is *no* rotation, ie $\omega = 0$. In this case the system reduces to the simple pendulum, and the equation of motion should be the same. We can immediately see that this is the case.

It is worth noting that the potential energy contains a time-dependent term $mga \sin \omega t$, which one naively would think should contribute to the dynamics of the system — however, it plays no role since it does not contain the coordinate. There is also a constant term $ma^2\omega^2/2$ in the kinetic energy which plays no role.

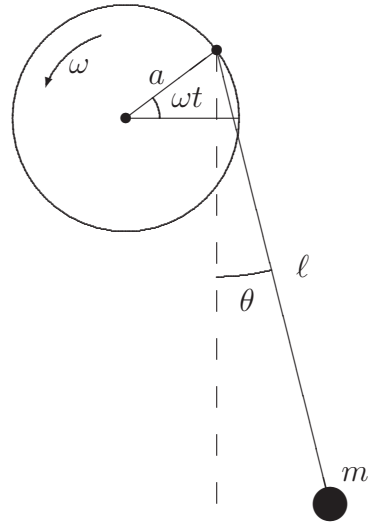


Figure 2.3: Pendulum with rotating support.

2.3.1 Polar and spherical coordinates

When we have rotational motion, or a system with rotational (or spherical) symmetry, it is very often most convenient to use polar coordinates (in 2 dimensions) or spherical coordinates (in 3 dimensions). The definition of these coordinates are given in Fig. 2.4. Since we will be using them often, we need to know what the kinetic energy of a particle is in terms of these coordinates.

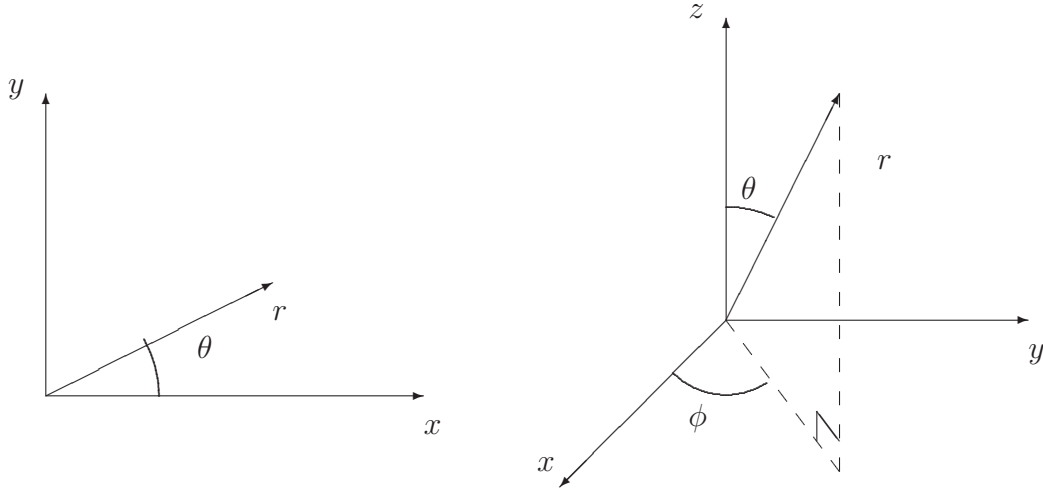


Figure 2.4: Plane polar coordinates (r, θ) (left) and spherical coordinates (r, θ, ϕ) (right).

Polar coordinates

The relation between cartesian and polar coordinates is given by

$$x = r \cos \theta \quad \Longrightarrow \quad \dot{x} = \dot{r} \cos \theta - r \dot{\theta} \sin \theta \quad (2.37)$$

$$y = r \sin \theta \quad \Longrightarrow \quad \dot{y} = \dot{r} \sin \theta + r \dot{\theta} \cos \theta \quad (2.38)$$

This gives for the kinetic energy,

$$\begin{aligned} T &= \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) \\ &= \frac{m}{2}(\dot{r}^2 \cos^2 \theta + r^2 \dot{\theta}^2 \sin^2 \theta - 2r\dot{r}\dot{\theta} \cos \theta \sin \theta + \dot{r}^2 \sin^2 \theta + r^2 \dot{\theta}^2 \cos^2 \theta + 2r\dot{r}\dot{\theta} \cos \theta \sin \theta) \\ &= \frac{m}{2}(\dot{r}^2 + r^2 \dot{\theta}^2). \end{aligned} \quad (2.39)$$

Spherical coordinates

The relation between cartesian and spherical coordinates is given by

$$x = r \sin \theta \cos \phi \implies \dot{x} = \dot{r} \sin \theta \cos \phi + r \dot{\theta} \cos \theta \cos \phi - r \dot{\phi} \sin \theta \sin \phi \quad (2.40)$$

$$y = r \sin \theta \sin \phi \implies \dot{y} = \dot{r} \sin \theta \sin \phi + r \dot{\theta} \cos \theta \sin \phi + r \dot{\phi} \sin \theta \cos \phi \quad (2.41)$$

$$z = r \cos \theta \implies \dot{z} = \dot{r} \cos \theta - r \dot{\theta} \sin \theta \quad (2.42)$$

Using this we find that

$$T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\dot{\phi}^2 \sin^2 \theta). \quad (2.43)$$

The complete derivation is left as an exercise.

2.4 Lagrange multipliers [Not exam material]

Using the constraint equations to reduce the number of coordinates is usually the most straightforward way of handling constraints. But it is not always practical:

- It may not be straightforward to solve the constraint equations.
- The constraint equations may involve velocities.
- The constraint equations may be expressed as differential rather than algebraic equations.
- We may want to know the forces of constraint (for example, to find out when they become too large or too small).

2.4.1 Velocity constraints and differential constraints

A constraint relation of the form

$$g_\alpha(x, \dot{x}, t) = 0, \quad x = \{x_1, \dots, x_N\} \quad (2.44)$$

involving velocities \dot{x}_i , can in general not be integrated to yield relations between coordinates. However, consider the equation

$$\sum_i A_i(x, t) \dot{x}_i + B(x, t) = 0, \quad (2.45)$$

If we can write

$$A_i = \frac{\partial f}{\partial x_i} \quad B = \frac{\partial f}{\partial t} \quad \text{with} \quad f = f(x, t), \quad (2.46)$$

then (2.45) is equivalent to an algebraic relation among the coordinates x_i :

$$\sum_i \frac{\partial f}{\partial x_i} \frac{dx_i}{dt} + \frac{\partial f}{\partial t} = \frac{df}{dt} = 0 \quad (2.47)$$

$$\iff f(x, t) - C = 0. \quad (2.48)$$

This is an example of a *differential constraint*

$$\sum_i \frac{\partial f_j}{\partial q_i} dq_i + \frac{\partial f_j}{\partial t} dt = 0. \quad (2.49)$$

2.4.2 Lagrange undetermined multipliers

Assume now that we are keeping our original coordinates q_i (whatever they are) and write the constraint equations on the differential form (2.49). We can now try to rederive the Euler–Lagrange equations.

Consider a variation $q_i^\alpha(t) = q_i(t) + \alpha h_i(t)$. We find, as before,

$$\frac{dS}{d\alpha} = \frac{d}{d\alpha} \int_{t_1}^{t_2} L(q_i^\alpha, \dot{q}_i^\alpha, t) dt = \int_{t_1}^{t_2} \sum_{i=1}^N \left(\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right) h_i dt = 0. \quad (2.50)$$

However, the functions h_i are no longer independent, since the q_i^α must satisfy the equations (2.49). In particular, we must have

$$\sum_i \frac{\partial f_j}{\partial q_i} \frac{dq_i^\alpha}{d\alpha} \sum_i \frac{\partial f_j}{\partial q_i} h_i = 0. \quad (2.51)$$

Note that the second term in (2.49) does not contribute, since $dt/d\alpha = 0$ as t is not varied.

If we have m constraint equations, we can use them to eliminate m of the functions h_i :

$$\begin{aligned} \frac{\partial f_1}{\partial q_N} h_N(t) &= - \sum_{i=1}^{N-1} \frac{\partial f_1}{\partial q_i} h_i(t) \iff h_N(t) = - \sum_{i=1}^{N-1} \frac{\partial f_1 / \partial q_i}{\partial f_1 / \partial q_N} h_i(t) \quad (2.52) \\ \frac{\partial f_2}{\partial q_{N-1}} h_{N-1}(t) &= - \sum_{i=1}^{N-2} \frac{\partial f_2}{\partial q_i} h_i(t) - \frac{\partial f_2}{\partial q_N} h_N(t) \\ &= - \sum_{i=1}^{N-2} \frac{\partial f_2}{\partial q_i} h_i(t) - \frac{\partial f_2 / \partial q_N}{\partial f_1 / \partial q_N} \sum_{i=1}^{N-2} \frac{\partial f_1}{\partial q_i} h_i(t) - \frac{\partial f_2 / \partial q_N}{\partial f_1 / \partial q_N} \frac{\partial f_1}{\partial q_{N-1}} h_{N-1}(t) \quad (2.53) \end{aligned}$$

... etc.

This becomes messy and not very illuminating in general, so let us consider just the case with $m = 1, N = 3$. Eq. (2.52) then becomes

$$h_3 = - \frac{\partial f / \partial q_1}{\partial f / \partial q_3} h_1 - \frac{\partial f / \partial q_2}{\partial f / \partial q_3} h_2 \quad (2.54)$$

and (2.50) becomes

$$\int_{t_1}^{t_2} \left\{ \left(\frac{\partial L}{\partial q_1} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_1} \right) h_1 + \left(\frac{\partial L}{\partial q_2} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_2} \right) h_2 + \left(\frac{\partial L}{\partial q_3} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_3} \right) h_3 \right\} dt = 0 \quad (2.55)$$

$$\begin{aligned} \iff \int_{t_1}^{t_2} \left\{ \left[\left(\frac{\partial L}{\partial q_1} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_1} \right) - \frac{\partial f / \partial q_1}{\partial f / \partial q_3} \left(\frac{\partial L}{\partial q_3} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_3} \right) \right] h_1 \right. \\ \left. + \left[\left(\frac{\partial L}{\partial q_2} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_2} \right) - \frac{\partial f / \partial q_2}{\partial f / \partial q_3} \left(\frac{\partial L}{\partial q_3} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_3} \right) \right] h_2 \right\} dt = 0 \end{aligned} \quad (2.56)$$

Since h_1 and h_2 are now both arbitrary functions, the terms inside the square brackets must both be zero. So we have

$$\begin{aligned} \left(\frac{\partial L}{\partial q_1} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_1} \right) \left(\frac{\partial f}{\partial q_1} \right)^{-1} &= \left(\frac{\partial L}{\partial q_3} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_3} \right) \left(\frac{\partial f}{\partial q_3} \right)^{-1} \\ \left(\frac{\partial L}{\partial q_2} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_2} \right) \left(\frac{\partial f}{\partial q_2} \right)^{-1} &= \left(\frac{\partial L}{\partial q_3} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_3} \right) \left(\frac{\partial f}{\partial q_3} \right)^{-1} \end{aligned} \quad \text{at all } t. \quad (2.57)$$

The two sides of the equations can be any function of t ; we call this function $-\lambda(t)$. This gives the equations

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} + \lambda(t) \frac{\partial f}{\partial q_i} = 0 \quad (2.58)$$

The function $\lambda(t)$ is called a *Lagrange undetermined multiplier*.

With m constraint equations, we get m functions $\lambda_k(t)$, ie

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} + \sum_{k=1}^m \frac{\partial f_k}{\partial q_i} \lambda_k(t) = 0 \quad (2.59)$$

The Euler–Lagrange equations with Lagrange multipliers

We now have $N + m$ unknown functions $q_i(t)$, $\lambda_k(t)$, but we also have $N + m$ equations: the N EL equations and the m constraint equations. This will therefore completely determine the system once the initial conditions are given.

If we know the Lagrange multipliers, we can find the (generalised) constraint forces F_i . These are given by

$$F_i = \sum_k \frac{\partial f_k}{\partial q_i} \lambda_k. \quad (2.60)$$

2.5 Canonical momenta and conservation laws

Assume the lagrangian L does not depend explicitly on the coordinate q_i . Such coordinates are called **cyclic**. The Euler–Lagrange equations for the cyclic coordinate q_i

becomes

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} = \frac{\partial L}{\partial q_i} = 0 \implies \boxed{\frac{\partial L}{\partial \dot{q}_i} \equiv p_i = \text{constant}.} \quad (2.61)$$

We call the quantity p_i the *canonical momentum* conjugate to (or corresponding to) q_i .

Why momentum?

Consider the 'usual' case where

1. we use cartesian coordinates $q_i = x_i$;
2. there are no constraints; and
3. the potential depends only on the coordinates, $V = V(x)$.

In this case we have

$$L = T - V = \frac{1}{2}m \sum_j \dot{x}_j^2 - V(x) \implies \frac{\partial L}{\partial \dot{x}_i} = m\dot{x}_i = p_i = \text{ordinary momentum.}$$

So we have found the law of conservation of momentum p_i if the potential V does not depend on the coordinate x_i — ie, if the system is *translationally invariant in the i -direction*. Note that if V does not depend on x_i this implies that there are no net forces in the i -direction.

We may in a similar way demonstrate conservation of total momentum for a system of n particles if the potential energy does not depend on the centre of mass coordinate. But the concept of canonical momenta is much more general and powerful than this, and can be used to derive a whole host of other conservation laws. One of the most important is *angular momentum*, which we will look at next.

2.5.1 Angular momentum

Consider a one-particle rotationally symmetric 2-dimensional system, and let us use polar coordinates (r, θ) to describe the particle. Rotational symmetry then means that the potential energy $V(r, \theta) = V(r)$, independent of the angle θ . The lagrangian is then

$$L = T - V = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) - V(r). \quad (2.62)$$

We see that θ is a cyclic coordinate, and the canonical momentum p_θ is therefore conserved. What is this canonical momentum?

We straightforwardly find

$$p_\theta = \frac{\partial L}{\partial \dot{\theta}} = \frac{\partial}{\partial \dot{\theta}} \left(\frac{1}{2}mr^2\dot{\theta}^2 \right) = mr^2\dot{\theta}. \quad (2.63)$$

But $\dot{\theta}$ is the same as the angular velocity ω , and we know that the velocity v_θ in the angular direction (perpendicular to the radius r) is $v_\theta = r\omega = r\dot{\theta}$, so $p_\theta = r(mv_\theta)$. But this is exactly equal to the angular momentum of the particle,

$$L_z = (\vec{r} \times \vec{p})_z = rv_\theta. \quad (2.64)$$

So the canonical momentum conjugate to the angle θ is the angular momentum, which is conserved if the system is rotationally symmetric, ie the lagrangian does not depend on θ .

Angular momentum in spherical coordinates

In section 2.3.1 we found that the kinetic energy in spherical coordinates (see Fig. 2.4 is

$$T = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2 + r^2 \sin^2 \theta \dot{\phi}^2). \quad (2.43)$$

The angle ϕ corresponds to rotations about the z -axis: if a particle rotates about the z -axis, ϕ changes while r and θ are unchanged. If the potential energy does not depend on ϕ , we have rotational symmetry about the z -axis, and the canonical momentum p_ϕ is conserved. From (2.43) we find

$$p_\phi = mr^2 \sin^2 \theta \dot{\phi} = r(mr \sin \theta \dot{\phi})(\sin \theta). \quad (2.65)$$

We now want to show that this is equal to the z -component of the angular momentum, $L_z = (\vec{r} \times \vec{p})_z$.

We can put a coordinate system $(\hat{r}, \hat{\theta}, \hat{\phi})$ at the point \vec{r} and decompose the velocity in its (r, θ, ϕ) components,

$$\vec{v} = v_r \hat{r} + v_\theta \hat{\theta} + v_\phi \hat{\phi} \quad (2.66)$$

The unit vector \hat{r} denotes the radial direction, ie the direction where r changes, while θ, ϕ are unchanged. Similarly, $\hat{\theta}$ denotes the direction where θ changes while r, ϕ are unchanged, and $\hat{\phi}$ denotes the direction where ϕ changes while r, θ are unchanged. The three vectors are orthogonal, and $\hat{\phi}$ is also orthogonal to \hat{z} , since motion in the ϕ -direction does not change z .

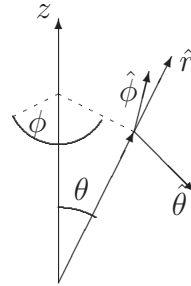


Figure 2.5: Unit vectors in spherical coordinates.

The velocity component v_ϕ is the rotational velocity about the z -axis, which again is equal to the distance from the axis times the angular velocity about the axis. Since ϕ is the rotational angle about the z -axis, the angular velocity is $d\phi/dt = \dot{\phi}$. The distance from the axis is $r \sin \phi$, so

$$v_\phi = (r \sin \phi) \dot{\phi}. \quad (2.67)$$

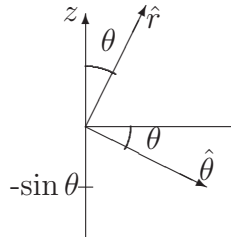
We can now work out the vector product $\vec{r} \times m\vec{v}$ in the spherical coordinate system. Since $\vec{r} = r\hat{r}$, we need the cross product of \hat{r} with each unit vector. Using the right-hand rule we find

$$\hat{r} \times \hat{\theta} = \hat{\phi}, \quad \hat{r} \times \hat{\phi} = -\hat{\theta}, \quad \hat{r} \times \hat{r} = 0. \quad (2.68)$$

The z -component of the angular momentum is therefore

$$L_z = (\vec{r} \times m \vec{v})_z = mr[\hat{r} \times (v_r \hat{r} + v_\theta \hat{\theta} + v_\phi \hat{\phi})]_z = mr[v_\theta \phi - v_\phi \dot{\theta}]_z = -mrv_\phi \hat{\theta} \cdot \hat{z}. \quad (2.69)$$

We now need to work out the scalar product $\hat{\theta} \cdot \hat{z}$. Looking at the figure on the right, we see that since θ is the angle of \vec{r} with the z -axis, and $\hat{\theta}$ is orthogonal to \vec{r} (but still in the $z-r$ plane), the projection of $\hat{\theta}$ onto the z -axis is $\hat{\theta} \cdot \hat{z} = -\sin \theta$.



Therefore we find that the z -component of the angular momentum is

$$L_z = (\vec{r} \times m \vec{v})_z = -mrv_\phi(-\sin \theta) = r \cdot (mr \sin \theta \dot{\phi}) \cdot \sin \theta = p_\phi. \quad (2.70)$$

So the canonical momentum p_ϕ is indeed the angular momentum about the z -axis, and it is conserved if we have rotational symmetry about the z -axis.

If we have full spherical symmetry, this means we have rotational symmetry about *all* 3 axes, so by the same argument as above L_x and L_y must also be conserved. Therefore, *for a spherically symmetric system, the angular momentum vector $\vec{L} = \vec{r} \times \vec{p}$ is conserved.*

Naïvely one would think that if we have full rotational symmetry, the angle θ should also be irrelevant, and the canonical momentum p_θ should also be conserved. However, this is not the case: although the potential energy does not depend on θ , the kinetic energy does, through the term $\frac{1}{2}mr^2 \sin^2 \theta \dot{\phi}^2$. This θ -dependence is an artefact of how we have chosen the coordinate system, but it is an unavoidable artefact: no matter how we choose our spherical coordinate angles, these coordinates must break the full spherical symmetry somehow. We realise the full symmetry by noting that we could have chosen the coordinates differently, eg we could have chosen θ to be the angle with the x -axis and ϕ to correspond to rotations about the x -axis — which would have led us to find that L_x is conserved.

2.6 Energy conservation: the hamiltonian

We know that when we have conservative forces, the potential energy depends only on positions, and not on time, and the total energy is conserved. We have derived conservation of linear and angular momentum in lagrangian mechanics, so we may ask ourselves if we can also derive energy conservation within the same framework?

The answer to this is that not only can we do this, but the energy conservation theorem we arrive at is more general than the one we already know!

To see how this works, let us take the (total) time derivative of the lagrangian $L =$

$L(q_i(t), \dot{q}_i(t), t)$. Using the chain rule and the Euler–Lagrange equations we get

$$\begin{aligned} \frac{dL}{dt} &= \sum_i \frac{\partial L}{\partial q_i} \dot{q}_i + \sum_i \frac{\partial L}{\partial \dot{q}_i} \ddot{q}_i + \frac{\partial L}{\partial t} \\ &= \sum_i \left(\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right) \dot{q}_i + \sum_i \frac{\partial L}{\partial \dot{q}_i} \frac{d\dot{q}_i}{dt} + \frac{\partial L}{\partial t} \end{aligned} \quad (2.71)$$

$$\begin{aligned} &= \frac{d}{dt} \sum_i \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i + \frac{\partial L}{\partial t} \\ \iff \frac{d}{dt} \left(\sum_i \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i - L \right) + \frac{\partial L}{\partial t} &\equiv \frac{dH}{dt} + \frac{\partial L}{\partial t} = 0, \end{aligned} \quad (2.72)$$

where we have defined

$$H = \sum_i \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i - L = \sum_i p_i \dot{q}_i - L = \mathbf{the\ hamiltonian} \quad (2.73)$$

So we find that if the lagrangian does not depend explicitly on time, then the *hamiltonian* or *energy function* H is conserved.

To see how this relates to energy conservation as we know it from before, consider a system of particles in cartesian coordinates, described by the lagrangian

$$L = T - V = \frac{1}{2} \sum_i m_i \dot{q}_i^2 - V(q).$$

The hamiltonian for this system is

$$H = \sum_i \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i - L \quad (2.74)$$

2.6.1 When is H (not) conserved?

We found that H is conserved if L does not depend explicitly on time, ie $L(q, \dot{q}, t) = L(q, \dot{q})$. We would like to understand in what circumstances an explicit time dependence could appear in the lagrangian. One possibility would be that the potential energy depends explicitly on time in the first place. But there are also other possibilities. The kinetic energy, written in terms of the original cartesian (or, for that sake, ordinary polar or spherical coordinates) does not have any explicit time dependence. But time dependence could still appear in either the kinetic or the potential energy when we write it in terms of generalised coordinates.

To see how this can happen, let us recall why we introduced generalised coordinates in the first place:

1. **Constraints:** There are fewer actual degrees of freedom in the system because of constraints. We use generalised coordinates to denote the real (relevant) degrees of freedom. An example of this would be the pendulum, where the original x and z coordinates are reduced to the single coordinate θ .
2. **Symmetries:** There are symmetries in the system which mean that using non-cartesian coordinates may give a simpler description. An example of this would be using polar coordinates for a system with rotational symmetry.

Explicit time-dependence can appear in both those types of cases, leaving us with three possibilities for how explicit time-dependence could appear in the lagrangian:

1. The potential energy is explicitly time-dependent, $V = V(x, t)$. Physically, this means that there are external or non-conservative forces, so the energy of the system is not conserved.
2. The constraints are time-dependent. An example of that would be Example 5, the pendulum with rotating support. In such cases, external forces are usually required to maintain the constraint, so the energy of the system is not conserved.
3. We have chosen to use time-dependent transformations $x_i = f_i(q, t)$ between the old coordinates x and the new coordinates q because this may simplify the description of the system. In this case, the hamiltonian may be not conserved even if the total energy is conserved.

2.6.2 When is H (not) equal to the total energy?

We showed the hamiltonian H is equal to the total energy E when

$$L = T - V = \frac{1}{2} \sum_i m_i \dot{q}_i^2 - V(q).$$

More generally, it is the case when

1. V is independent of the velocities \dot{q}_i , $V = V(q, t)$, and
2. T is a *homogeneous quadratic function* of \dot{q} ,

$$T = \sum_{ij} a_{ij} \dot{q}_i \dot{q}_j.$$

Proof

Take

$$L = T - V = \sum_{ij} a_{ij} \dot{q}_i \dot{q}_j - V(q, t) \tag{2.75}$$

We note that we can always arrange it so that $a_{ij} = a_{ji}$, since $\dot{q}_i \dot{q}_j = \dot{q}_j \dot{q}_i$. The canonical momenta are

$$p_k = \frac{\partial L}{\partial \dot{q}_k} = \sum_i a_{ik} \dot{q}_i + \sum_j a_{kj} \dot{q}_j = 2 \sum_j a_{kj} \dot{q}_j. \quad (2.76)$$

The two terms appear because we get a contribution both from the $k = j$ term and from the $k = i$ term in the sum. The hamiltonian is then

$$H = \sum_i p_i \dot{q}_i - L = 2 \sum_{ij} a_{ij} \dot{q}_i \dot{q}_j - \sum_{ij} a_{ij} \dot{q}_i \dot{q}_j + V(q, t) = T + V = E, \quad (2.77)$$

which completes the proof.

So when does T (not) have this form?

Let us start from cartesian coordinates, where

$$T = \frac{1}{2} \sum_i m_i \dot{x}_i^2 \quad i = 1, \dots, 3N. \quad (2.78)$$

We now introduce generalised coordinates q_j , which are related to the x_i via general, time-dependent transformations,

$$x_i = x_i(q_1, \dots, q_m, t) \quad (2.79)$$

Using the chain rule, we find

$$\dot{x}_i = \frac{dx_i}{dt} = \sum_{j=1}^m \frac{\partial x_i}{\partial q_j} \dot{q}_j + \frac{\partial x_i}{\partial t}, \quad (2.80)$$

$$\begin{aligned} \dot{x}_i^2 &= \left(\sum_{j=1}^m \frac{\partial x_i}{\partial q_j} \dot{q}_j + \frac{\partial x_i}{\partial t} \right) \left(\sum_{k=1}^m \frac{\partial x_i}{\partial q_k} \dot{q}_k + \frac{\partial x_i}{\partial t} \right) \\ &= \sum_{j,k=1}^m \frac{\partial x_i}{\partial q_j} \dot{q}_j \frac{\partial x_i}{\partial q_k} \dot{q}_k + 2 \sum_{j=1}^m \frac{\partial x_i}{\partial q_j} \frac{\partial x_i}{\partial t} \dot{q}_j + \left(\frac{\partial x_i}{\partial t} \right)^2. \end{aligned} \quad (2.81)$$

The kinetic energy is therefore

$$\begin{aligned} T &= \frac{1}{2} \sum_{i=1}^{3N} m_i \dot{x}_i^2 \\ &= \frac{1}{2} \sum_{j,k=1}^m \left(\sum_{i=1}^{3N} m_i \frac{\partial x_i}{\partial q_j} \frac{\partial x_i}{\partial q_k} \right) \dot{q}_j \dot{q}_k + \sum_{j=1}^m \left(\sum_{i=1}^{3N} m_i \frac{\partial x_i}{\partial q_j} \frac{\partial x_i}{\partial t} \right) \dot{q}_j + \frac{1}{2} \sum_{i=1}^{3N} m_i \left(\frac{\partial x_i}{\partial t} \right)^2. \end{aligned} \quad (2.82)$$

This can be written on the form

$$T = \sum_{jk} a_{jk} \dot{q}_j \dot{q}_k + \sum_j b_j \dot{q}_j + c, \quad (2.83)$$

with

$$a_{jk} = \frac{1}{2} \sum_{i=1}^{3N} m_i \frac{\partial x_i}{\partial q_j} \dot{q}_j \frac{\partial x_i}{\partial q_k}, \quad b_j = \sum_{i=1}^{3N} m_i \frac{\partial x_i}{\partial q_j} \frac{\partial x_i}{\partial t}, \quad c = \frac{1}{2} \sum_{i=1}^{3N} m_i \left(\frac{\partial x_i}{\partial t} \right)^2. \quad (2.84)$$

If the transformations do *not* depend on time, $x_i = x_i(q_1, \dots, q_m)$, then $\frac{\partial x_i}{\partial t} = 0$, so $b_j = 0, c = 0$ and therefore,

$$T = \sum_{ij} a_{ij} \dot{q}_i \dot{q}_j. \quad (2.85)$$

So the kinetic energy will be a homogeneous quadratic function of the generalised velocities \dot{q}_i whenever the constraints or coordinate transformations do not depend on time. Conversely, we may have $H \neq E$ if

1. the potential V depends on velocities, *or*
2. the constraints of coordinate transformations are time-dependent.

Example 2.6 Spring mounted on moving cart

Consider a body with mass m sitting at the end of a horizontal spring with spring constant k , with the other end attached to a cart moving with a constant velocity v . Since one end of the spring is fixed to the moving cart, the equilibrium point x_0 of the body on the spring is also moving with velocity v . If we say that $x_0 = 0$ when $t = 0$, we therefore have $x_0 = vt$.

The potential energy of the body is given by the displacement $x - x_0$ from equilibrium, $V = \frac{1}{2}k(x - x_0)^2 = \frac{1}{2}k(x - vt)^2$. The kinetic energy is the usual one, so the lagrangian is

$$L = T - V = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}k(x - vt)^2. \quad (2.86)$$

The canonical momentum is $p_x = m\dot{x}$, which gives us the hamiltonian

$$H = p_x \dot{x} - L = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}k(x - vt)^2 = T + V = E. \quad (2.87)$$

Since the lagrangian depends explicitly on time, the hamiltonian (and the total energy) is *not* conserved. We can understand this by noting that the motor driving the cart will have to do work to maintain a constant velocity; in the absence of this the cart will undergo oscillations along with the body attached to the spring.

We can now introduce a new coordinate

$$q = x - vt \quad \implies \quad \dot{q} = \dot{x} - v \quad (2.88)$$

$$\implies \quad L(q, \dot{q}, t) = \frac{1}{2}m(\dot{q} + v)^2 - \frac{1}{2}kq^2 = \frac{1}{2}m\dot{q}^2 + mv\dot{q} - \frac{1}{2}kq^2 + \frac{mv^2}{2}. \quad (2.89)$$

The canonical momentum is $p = m(\dot{q} + v)$, and the hamiltonian is

$$H_q = p\dot{q} - L = m\dot{q}(\dot{q} + v) - \left(\frac{1}{2}m\dot{q}^2 + mv\dot{q} - \frac{1}{2}kq^2 + \frac{mv^2}{2} \right) = \frac{1}{2}m\dot{q}^2 + \frac{1}{2}kq^2 - \frac{mv^2}{2}. \quad (2.90)$$

When written in terms of q , the lagrangian does *not* depend explicitly on time, and therefore the hamiltonian (2.90) *is* conserved! However, it is *not* equal to the total energy.

Example 2.7 Electrodynamics

One case where the distinction between ordinary and canonical momentum is important is electrodynamics. A particle with charge Q moving with velocity \vec{v} in an electric field \vec{E} and a magnetic field \vec{B} is

$$\vec{F} = Q(\vec{E} + \vec{v} \times \vec{B}) \quad (2.91)$$

Using Maxwell's laws, we can introduce the electrostatic and vector ('magnetic') potentials ϕ, \vec{A} :

$$\nabla \cdot \vec{B} = 0 \quad \iff \quad \vec{B} = \nabla \times \vec{A}, \quad (2.92)$$

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \quad \iff \quad \vec{E} = -\nabla\phi - \frac{\partial \vec{A}}{\partial t}. \quad (2.93)$$

It is now possible (although the proof is not straightforward, so we will not present it here) to derive the force (2.91) from a potential

$$U = Q\phi - Q\vec{A} \cdot \vec{v}. \quad (2.94)$$

The lagrangian is then

$$L = T - U = \frac{1}{2}m \sum_{i=1}^3 \dot{x}_i^2 - Q\phi + Q \sum_{i=1}^3 A_i \dot{x}_i. \quad (2.95)$$

The canonical momentum is

$$p_i = \frac{\partial L}{\partial \dot{x}_i} = m\dot{x}_i + QA_i. \quad (2.96)$$

This is *not* the ordinary momentum, a distinction which becomes quite important in quantum mechanics, where it is the canonical momentum that enters into the commutation relations that are used to quantise the system. Note that \vec{A} and/or ϕ must depend on x , otherwise the problem is trivial (there are no forces), so the momentum is not conserved.

The hamiltonian of the particle is

$$\begin{aligned} H &= \sum_{i=1}^3 p_i \dot{x}_i - L = \sum_{i=1}^3 (m\dot{x}_i + QA_i)\dot{x}_i - \frac{1}{2}m \sum_{i=1}^3 \dot{x}_i^2 + Q\phi - Q \sum_{i=1}^3 A_i \dot{x}_i \\ &= \frac{1}{2}m \sum_{i=1}^3 \dot{x}_i^2 + Q\phi. \end{aligned} \quad (2.97)$$

We see that the vector (magnetic) potential does not contribute to the hamiltonian. Physically, this is because no net work is done by the magnetic field.