Chapter 1

Least action

Newton, through his three laws of dynamics, developed an extremely successful description of the motion of objects. These laws can, for example, describe the elliptical orbits of planets to remarkable precision. There is though an alternative presentation of these successes, the principle of least action, which we will explore here. It is a formalism that grew out of optics and will allow us to study an area of mathematics called 'calculus of variation'. Of course it must turn out to be the same as Newton's laws. This alternative formalism makes some dynamics problems easier to solve but, more importantly, it will give us new insights into conservation laws. It is important to master these methods since as one moves to the forefront of modern quantum theories the least action principle becomes the only way to define theories such as that of the strong nuclear force.

1.1 Optics

Our starting point will be to think about the path that light travels by. In these enlightened times we might start from Maxwell's equations and derive a wave equation with light waves as solutions to determine how the light propagates. Before this technology though Fermat proposed:

Fermat's principle of least time: *light propagates between two points so as to minimize its travel time.*

Thus for example in a uniform medium where the speed of light c is a constant the minimum time of travel

$$t = \frac{d}{c} \tag{1.1}$$

is given by the path of shortest distance d, i.e. a straight line. This is still a perfectly good (if limited) description of light.

We can obtain more interesting results by thinking about media where the speed of light changes.

doi:10.1088/978-1-6817-4696-8ch1

1.1.1 Snell's law

Consider two neighbouring regions of space in which light travels at different speeds v_1, v_2—for example a glass–air interface. We will be interested in the light that travels from the point (x_1, y_1) in the first medium to the point (x_2, y_2) in the second (figure 1.1).

In any one medium light travels in a straight line but in this case we have some choice in where the light crosses between the media. Let's consider the arbitrary crossing point $(x, y = 0)$. The time of travel is

$$\begin{aligned} T[x] &= \frac{d_1}{v_1} + \frac{d_2}{v_2} \\ &= \frac{\sqrt{(x - x_1)^2 + y_1^2}}{v_1} + \frac{\sqrt{(x - x_2)^2 + y_2^2}}{v_2} \end{aligned} \tag{1.2}$$

We now want to find the path (i.e. the value of x through which it passes) which minimizes the time taken. Thus

$$\frac{dT}{dx} = \frac{(x - x_1)}{v_1\sqrt{(x - x_1)^2 + y_1^2}} - \frac{(x_2 - x)}{v_2\sqrt{(x_2 - x)^2 + y_2^2}} = 0 \tag{1.3}$$

This equation though is just

$$\boxed{\frac{\sin\theta_1}{v_1} = \frac{\sin\theta_2}{v_2}} \tag{1.4}$$

which is *Snell's law*.

In terms of index of refraction which is defined, relative to the vacuum, as

$$n_1 = \frac{c}{v_1} \tag{1.5}$$

$$n_1 \sin\theta_1 = n_2 \sin\theta_2 \tag{1.6}$$

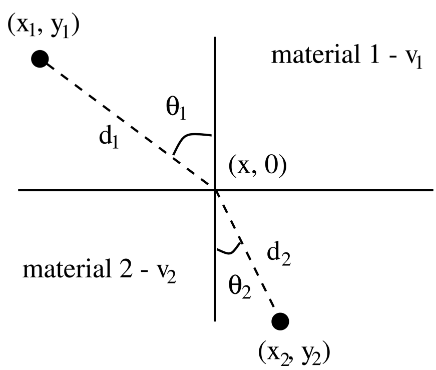

Figure 1.1. Possible paths that light might follow transiting across the interface between two materials.

1.1.2 Complicated problems

We can imagine more complicated problems than that above where the index of refraction is an arbitrary function of position. For example consider light moving in a plane where the speed of the light is $v(x, y)$ (figure 1.2).

Different paths are described by different functions $y(x)$. The time to travel along an arbitrary little piece of path is

$$dT = \frac{\text{distance}}{\text{velocity}} = \frac{\sqrt{dx^2 + dy^2}}{v(x, y)} \tag{1.7}$$

Summing such contributions up along a path gives the total time of travel

$$T[y(x)] = \int_{x_a}^{x_b} \frac{1}{v(x, y)} \sqrt{1 + \left(\frac{dy}{dx}\right)^2} \; dx \tag{1.8}$$

To rewrite this in a more standard form we have found that the time taken to traverse a path is

$$T[y] = \int_{x_a}^{x_b} L(y, \dot{y}, x) \; dx \tag{1.9}$$

where

$$L(y, \dot{y}, x) = \frac{1}{v(x, y)} \sqrt{1 + \dot{y}^2} \tag{1.10}$$

Now we want to find the path $y(x)$ that gives the minimum time.

Exercise 1.1: To remind yourself about partial differentiation, define a function by

$$T = a(t) \; b(t)^3 \dot{b}(t) t^{10}$$

where the dot indicates a derivative with respect to t. Give expressions for

$$\frac{\partial T}{\partial t}, \qquad \frac{\partial T}{\partial b}, \qquad \frac{\partial T}{\partial \dot{b}}, \qquad \frac{dT}{dt}$$

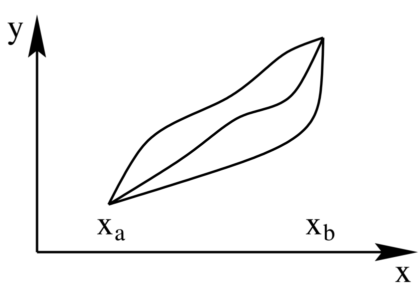

Figure 1.2. Possible paths that light might travel in a plane with varying speed of light $v(x, y)$.

This is the sort of problem that calculus of variation is designed to address, as discussed in appendix A.

As we show in appendix A, the problem of finding the path that minimizes the time, is equivalent to solving a differential equation called the Euler–Lagrange equation,

$$\frac{d}{dx}\left(\frac{\partial L}{\partial \dot{y}}\right) - \frac{\partial L}{\partial y} = 0 \tag{1.11}$$

which corresponds to equation (A.9) with q identified as y and s identified as x.

Let's look at a couple of examples.

1.1.3 Light in vacuum

In vacuum the speed of light is a constant so $v(x, y) = c$.

The Euler–Lagrange equation is

$$\frac{d}{dx}\left(\frac{\dot{y}}{\sqrt{1 + \dot{y}^2}}\right) = 0 \tag{1.12}$$

Integrating this gives

$$\frac{\dot{y}}{\sqrt{1 + \dot{y}^2}} = \text{constant} \tag{1.13}$$

The only solution of this is

$$\dot{y} = \text{constant}, m \tag{1.14}$$

or integrating

$$y = mx + c \tag{1.15}$$

i.e. a straight line. This is our first example of the solution of the Euler–Lagrange equation giving the path that minimizes T. m and c are determined by the initial and final position of the light.

1.1.4 Light in the atmosphere

In the atmosphere the air temperature and density change with height resulting in the speed of light depending on height—$v(h)$. Equivalently we can write the refractive index $n(h)$ with

$$v(h) = \frac{c}{n(h)} \tag{1.16}$$

Our result for the length of time light takes to travel some path $h(x)$ can be written as an *optical path length*

$$cT[h] = \int_{x_1}^{x_2} dxL, \qquad L = n(h)\sqrt{1 + \dot{h}^2} \tag{1.17}$$

We can use the fact that L is independent of x to simplify the Euler–Lagrange equation as follows. Note that

$$\frac{dL}{dx} = \frac{\partial L}{\partial x} + \frac{\partial L}{\partial h}\dot{h} + \frac{\partial L}{\partial \dot{h}}\ddot{h} \tag{1.18}$$

The first term on the right is zero. Now replace $\frac{\partial L}{\partial h}$ using the Euler–Lagrange equation

$$\frac{\partial L}{\partial h} = \frac{d}{dx}\left(\frac{\partial L}{\partial \dot{h}}\right) \tag{1.19}$$

and we find

$$\frac{dL}{dx} = \frac{d}{dx}\left(\frac{\partial L}{\partial \dot{h}}\right)\dot{h} + \frac{\partial L}{\partial \dot{h}}\ddot{h} \tag{1.20}$$

which is just

$$\frac{d}{dx}\left[L - \dot{h}\frac{\partial L}{\partial \dot{h}}\right] = 0 \tag{1.21}$$

which gives us

$$L - \dot{h}\frac{\partial L}{\partial \dot{h}} = \text{constant, } D \tag{1.22}$$

Note that this is only a first order equation rather than the second order Euler–Lagrange equation so is simpler to solve.

In our problem, using the explicit form for L above we have

$$n\sqrt{1 + \dot{h}^2} - \frac{\dot{h}^2 n}{\sqrt{1 + \dot{h}^2}} = D \tag{1.23}$$

which simplifies to

$$\frac{n}{\sqrt{1 + \dot{h}^2}} = D \tag{1.24}$$

Note that the physical meaning of D is the value of the index of refraction at the point where the light ray becomes horizontal so that $\dot{h} = 0$.

Squaring and rearranging we find

$$\frac{dh}{dx} = \sqrt{\frac{n^2}{D^2} - 1} \tag{1.25}$$

Thus

$$\boxed{x - x_0 = \int_{h_0}^{h} \frac{dh}{\sqrt{\frac{n^2}{D^2} - 1}}} \tag{1.26}$$

Explicit example: Consider a ray of light that begins moving horizontally ($\dot{h} = 0$) at $h = 0$ in an atmosphere where

$$n(h) = n_0 - \lambda h \tag{1.27}$$

where λ is some constant. We must solve the integral

$$x = \int \frac{dh}{\sqrt{\frac{(n_0 - \lambda h)^2}{D^2} - 1}} \tag{1.28}$$

This can be done by changing variables to

$$n_0 - \lambda h = D \cosh \phi \tag{1.29}$$

The integral becomes

$$x = -\int \frac{D}{\lambda} d\phi = -\frac{D}{\lambda}\phi + c \tag{1.30}$$

Returning to the original coordinates and requiring the boundary conditions $\dot{h}(x = 0) = 0$ and $h(x = 0) = 0$ gives the result

$$\boxed{h = \frac{n_0}{\lambda}\left(1 - \cosh \frac{\lambda x}{n_0}\right)} \tag{1.31}$$

- When λ is positive $n(h)$ decreases with altitude—this is what normally happens in the atmosphere. The plot of the form of the solution is given in figure 1.3.

  Thus if we look up at the Empire State building it will appear taller than it actually is.

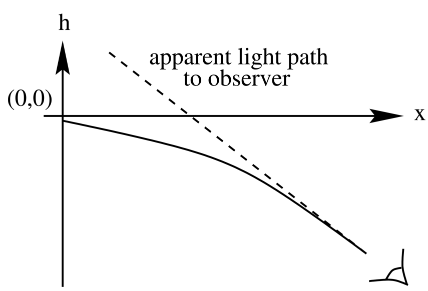

Figure 1.3. The solution for the light path $h(x)$ when $n(h)$ decreases with height.

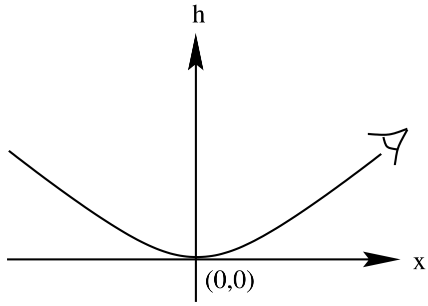

Figure 1.4. The solution for the light path $h(x)$ when $n(h)$ increases with height.

- If there is a temperature inversion then λ is negative so $n(h)$ increases with altitude. When we plot the form of the solution we get figure 1.4.
  We see 'the sky on the ground'—a mirage.

Exercise 1.2:

(a) Consider a fibre optic cable lying in the z direction. The cable is made of glass with index of refraction $n(r)$, where r is the radial distance from the centre of the cable. Working in cylindrical coordinates (r, θ, z) show that Fermat's principle implies that light travels on the path minimizing the quantity

$$\int_{z_1}^{z_2} f(r(z), \theta(z), r'(z), \theta'(z))dz = \int_{z_1}^{z_2} n(r)\sqrt{r'^2 + r^2\theta'^2 + 1}\,dz.$$

where a prime indicates differentiation with respect to z. z_1 and z_2 are the z-coordinates of the end points of the path.

(b) If a light ray initially has $\theta' = 0$ show, from the appropriate Euler–Lagrange equation, that the θ independence of f implies the path followed by the light is described by a constant value of θ.

(c) Use the z independence of f to deduce that the first order differential equation for rays travelling paths with constant θ is

$$f - \frac{\partial f}{\partial r'}\, r' = \text{constant}$$

1.2 Newtonian dynamics

We have seen that the motion of light can be described by a 'principle of least time'. Is there an equivalent rule that would describe the motion of a particle in Newtonian dynamics? There is and it is enshrined as

Hamilton's principle: *A particle travels by the path between two points that minimizes the action.*

We need to know what the 'action' is. Let's write it first for one dimensional motion. The action is

$$\boxed{S[\text{path}] = \int_{t_a}^{t_b} L(x, \dot{x}, t)dt} \tag{1.32}$$

where the dot indicates differentiation with respect to the time, t. L is known as the *Lagrangian* and is given by

$$\boxed{L = \text{kinetic energy} - \text{potential energy} = T - V} \tag{1.33}$$

From appendix A, we know that the path that minimizes the action satisfies the Euler–Lagrange equation, analogous to the case of optics equation (1.11),

$$\boxed{\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) - \frac{\partial L}{\partial x} = 0} \tag{1.34}$$

which corresponds to equation (A.11) with q identified as x and s identified as t.

We can now check to see if any of this makes sense (!). For a non-relativistic particle in a one dimensional potential we have

$$L = T - V = \frac{1}{2}m\dot{x}^2 - V(x) \tag{1.35}$$

The Euler–Lagrange equation is therefore

$$\frac{d}{dt}(m\dot{x}) + \frac{\partial V}{\partial x} = 0 \tag{1.36}$$

which is Newton's second law since

$$F = -\frac{\partial V}{\partial x} \tag{1.37}$$

Note that the momentum of the particle is given by

$$p = m\dot{x} = \frac{\partial L}{\partial \dot{x}} \tag{1.38}$$

1.2.1 Multiple coordinates

Suppose that we now have several coordinates

$$q_i \qquad i = 1, \ldots, n \tag{1.39}$$

For example for one particle moving in three dimensions we might call $x = q_1$, $y = q_2$, $z = q_3$.

As discussed in appendix A, equation (A.11), for the n dimensional case we have to solve a set of n Euler–Lagrange equations—one associated with each coordinate,

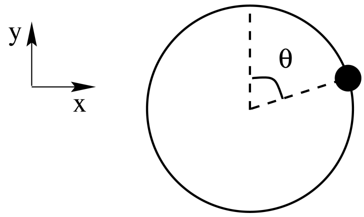

Figure 1.5. The coordinates describing a ball constrained to run around a hoop.

$$\boxed{\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0} \tag{1.40}$$

i.e. we need to write down n copies of the Euler–Lagrange equation, for $i = 1, 2, \ldots, n$ and try to solve them simultaneously.

Generalized coordinates
The reason that we have written the coordinates so generally as q_i rather than for example using x, y, z is that in some problems these are not the appropriate coordinates because of a *constraint*. A simple example to illustrate this is a ball on a wire hoop (figure 1.5).

The hoop stops the ball moving in the radial direction so the ball cannot be at any arbitrary (x, y). The sensible coordinate to use is the angle θ.

Such a reduced set of coordinates is called *generalized coordinates*.

Generalized momentum
A generalization of the idea of momentum can be defined in the spirit of equation (1.38). The *generalized momentum* associated with a generalized coordinate is given by

$$\boxed{p_i = \frac{\partial L}{\partial \dot{q}_i}} \tag{1.41}$$

1.2.2 Example: projectile motion

Consider the familiar problem of a projectile in a uniform gravitational field (figure 1.6).

We can obtain the normal Newtonian equations of motion from the Euler–Lagrange equations. We need expressions for the kinetic and potential energy of the system so we can build the Lagrangian. The kinetic energy is just

$$T = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}m\dot{y}^2 \tag{1.42}$$

and the potential energy

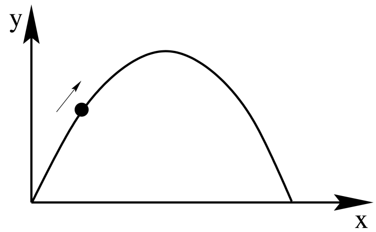

Figure 1.6. The motion of a projectile in the x, y plane subject to constant gravity in the vertical direction y.

$$V = mgy \tag{1.43}$$

So the Lagrangian is just

$$L = T - V = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}m\dot{y}^2 - mgy \tag{1.44}$$

Now we find the two Euler–Lagrange equations. The first associated with the x coordinate is

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) - \frac{\partial L}{\partial x} = 0 \tag{1.45}$$

which gives

$$\boxed{m\ddot{x} = 0} \tag{1.46}$$

The second equation associated with the y coordinate is

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{y}}\right) - \frac{\partial L}{\partial y} = 0 \tag{1.47}$$

which gives

$$\boxed{\ddot{y} = -g} \tag{1.48}$$

The two boxed equations are the standard Newtonian equations of motion.

Hopefully you're starting to see the power of this technique now—the kinetic and potential energies of a system are fairly easy to work out and then we just do some maths. There's not all that resolving forces business! The next problem is an example that would be very hard by the standard methodology.

1.2.3 Example 2: double pendulum

Consider a double pendulum as shown in figure 1.7.

It would be pretty hard work to determine all the forces in play here. However, the Lagrangian technique means we only have to calculate the energies of the two masses to get to the equations of motion.

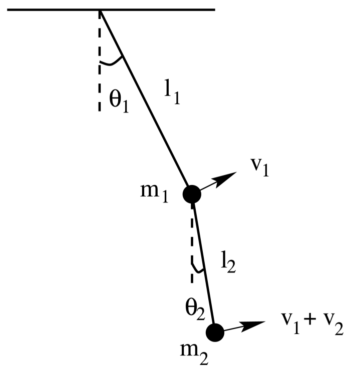

Figure 1.7. The coordinates relevant for a double pendulum.

The first mass has a velocity $\vec{v}_1$ with magnitude $l_1\dot{\theta}_1$ ($v = \omega r$). The second mass has both this motion plus a second contribution from the swing of the second pendulum $\vec{v}_2$ with magnitude $l_2\dot{\theta}_2$. The total velocity of the second mass is therefore

$$\vec{v}_{\text{tot}} = \vec{v}_1 + \vec{v}_2 \tag{1.49}$$

so

$$\begin{aligned} v_{\text{tot}}^2 &= (\vec{v}_1 + \vec{v}_2) \,.\, (\vec{v}_1 + \vec{v}_2) \\ &= (l_1\,\dot{\theta}_1)^2 + (l_2\,\dot{\theta}_2)^2 + 2l_1\,\dot{\theta}_1\,l_2\,\dot{\theta}_2\cos(\theta_2 - \theta_1) \end{aligned} \tag{1.50}$$

where $\theta_2 - \theta_1$ is the angle between $\vec{v}_1$ and $\vec{v}_2$.

Thus the total kinetic energy of the system is

$$T = \frac{1}{2}m_1 l_1^2\dot{\theta}_1^2 + \frac{1}{2}m_2[(l_1\dot{\theta}_1)^2 + (l_2\dot{\theta}_2)^2 + 2l_1\dot{\theta}_1 l_2\dot{\theta}_2\cos(\theta_2 - \theta_1)] \tag{1.51}$$

The potential energy is determined by the heights of the masses

$$V = -m_1 g l_1 \cos\theta_1 - m_2 g(l_1\cos\theta_1 + l_2\cos\theta_2) \tag{1.52}$$

and the Lagrangian is

$$L = T - V \tag{1.53}$$

There are two Euler lagrange equations—one associated with θ_1

$$\begin{aligned} \frac{d}{dt}\Big[m_1 l_1^2\dot{\theta}_1 + m_2 l_1^2\dot{\theta}_1 + m_2 l_1 l_2\dot{\theta}_2\cos(\theta_2 - \theta_1)\Big] - m_2 l_1 l_2\dot{\theta}_1\dot{\theta}_2\sin(\theta_2 - \theta_1) \\ + (m_1 + m_2) g l_1 \sin\theta_1 = 0 \end{aligned} \tag{1.54}$$

and one with θ_2

$$\frac{d}{dt}\Big[m_2 l_2^2\dot{\theta}_2 + m_2 l_1 l_2\dot{\theta}_1\cos(\theta_2 - \theta_1)\Big] + m_2 l_1 l_2\dot{\theta}_1\dot{\theta}_2\sin(\theta_2 - \theta_1) + m_2 g l_2 \sin\theta_2 = 0 \tag{1.55}$$

These are pretty messy (but that was the point!). Things simplify a bit if we assume that both θ_1 and θ_2 are small and expand to linear order. We then get

$$\begin{aligned}(m_1 + m_2)l_1^2\ddot{\theta}_1 + m_2 l_1 l_2 \ddot{\theta}_2 &= -(m_1 + m_2)g l_1 \theta_1 \\ m_2 l_2^2 \ddot{\theta}_2 + m_2 l_1 l_2 \ddot{\theta}_1 &= -m_2 g l_2 \theta_2\end{aligned} \tag{1.56}$$

These coupled equations in fact have normal mode solutions of the form

$$\begin{aligned}\ddot{\theta}_1 &= -\omega^2\theta_1 \\ \ddot{\theta}_2 &= -\omega^2\theta_2\end{aligned} \tag{1.57}$$

i.e. the two pendulums oscillate with the same frequency.

To find ω you can try substituting in the form of the solution in equations (1.57) into (1.56). You'll find two simultaneous equations for θ_1 and θ_2 with two solutions. You'll find in one case θ_1/θ_2 is positive and in the other it is negative. So in one case the pendulums swing together and in the other case in opposite directions.

Exercise 1.3: If a system with generalized coordinate q has the Lagrangian

$$L = \frac{1}{2}\dot{q}^2 - q^3$$

what is the Euler–Lagrange equation describing the system?

Exercise 1.4: If a system with generalized coordinates ξ and ψ has the Lagrangian

$$L = \frac{1}{2}\dot{\xi}^2 + \cos\xi\dot{\psi} - \xi e^{\psi}$$

what are the Euler–Lagrange equations describing the system?

Exercise 1.5: Two blocks of equal mass M are connected by a flexible string of length ℓ. One block is placed on a smooth horizontal table and the other block hangs over the edge. Using the length z of string hanging over the edge as a generalized coordinate, write down the Lagrangian and use the Euler–Lagrange equation to find the acceleration of the hanging mass in the following cases:

(i) The mass of the string is negligible.
(ii) The string is heavy with mass m distributed uniformly along it.

Exercise 1.6:

(a) Show that for a non-relativistic, free particle of mass m travelling with constant velocity v the action S describing its motion reduces to

$$S = mvd/2$$

where d is the distance travelled. This was a form for the action proposed by Maupertuis who believed it reflected the simplicity and economy of the Creator-God ….

(b) Consider such a particle rolling on a table in the x, y plane with speed v_1. Along the y-axis there is a height discontinuity in the table which the particle can move over at the cost of potential energy which reduces its velocity to v_2. If the particle starts at (x_1, y_1) to the left of the y-axis and ends to the right at (x_2, y_2) show that the action for it passing across the y-axis at arbitrary y (assuming it travels in a straight line except when it crosses the y-axis) is given by

$$S = mv_1\sqrt{x_1^2 + (y - y_1)^2} + mv_2\sqrt{x_2^2 + (y - y_2)^2}$$

By minimizing the action deduce the relation

$$v_1 \sin\theta_1 = v_2 \sin\theta_2$$

where the angles are the angles between the particle's direction of motion and the x-axis before and after it crosses the y-axis. Contrast this result with Snell's law for light.

1.3 Conservation laws

Finally let's look at one of the most surprising pieces of insight to come out of the Lagrangian formalism—that is a deeper understanding of conservation laws. The mathematics of this is discussed in more detail in appendix B.

1.3.1 Ignorable coordinates

If the Lagrangian does not depend on some coordinate q_i it is called an *ignorable coordinate*. Then $\partial L/\partial q_i = 0$ and its associated generalized momentum is conserved as we can see from the Euler–Lagrange equation

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = \frac{dp_i}{dt} = 0 \tag{1.58}$$

so

$$p_i = \text{constant} \tag{1.59}$$

This is clearly a mathematical fact but there is a deeper interpretation. If L only depends on $\dot{q}_i$ not q_i itself then we can shift

$$q_i \rightarrow q_i + \text{const} \tag{1.60}$$

and leave the Lagrangian, L, (and hence the physics) invariant. This is a *symmetry*—translation invariance in the q_i direction.

Thus we learn that the true relation is

$$\boxed{\text{symmetry} \rightarrow \text{conserved momentum}}$$

This is a new insight we have not seen before in Newtonian mechanics.

1.3.2 Energy conservation

Consider the case that L does not depend explicitly on t. This implies that a quantity known as the *Hamiltonian* is conserved. The *Hamiltonian* is defined as

$$H = \sum_i \frac{\partial L}{\partial \dot{q}_i}\dot{q}_i - L \tag{1.61}$$

To prove that it is conserved we explicitly calculate

$$\frac{dH}{dt} = \sum_i \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right)\dot{q}_i + \sum_i \frac{\partial L}{\partial \dot{q}_i}\ddot{q}_i - \frac{\partial L}{\partial t} - \sum_i \frac{\partial L}{\partial q_i}\dot{q}_i - \sum_i \frac{\partial L}{\partial \dot{q}_i}\ddot{q}_i = 0 \tag{1.62}$$

using the Euler–Lagrange equations and $\frac{\partial L}{\partial t} = 0$.

In simple systems the Hamiltonian is just the total energy of the system as we can see for example in one dimension where

$$L = \frac{1}{2}m\dot{x}^2 - V(x) \tag{1.63}$$

so using the definition above

$$H = \frac{1}{2}m\dot{x}^2 + V(x) \tag{1.64}$$

In conclusion here we have learnt that time translation invariance implies energy conservation.

1.3.3 Example: central forces

Consider a particle moving subject to a central force i.e. in a potential $V(r)$ (figure 1.8).

The kinetic energy of the particle is

$$T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) = \frac{1}{2}m\dot{r}^2 + \frac{1}{2}mr^2\dot{\theta}^2 \tag{1.65}$$

thus

$$L = \frac{1}{2}m\dot{r}^2 + \frac{1}{2}mr^2\dot{\theta}^2 - V(r) \tag{1.66}$$

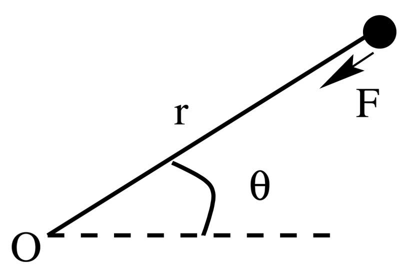

Figure 1.8. The coordinates relevant for a particle moving in a central potential.

There is an Euler–Lagrange equation associated with the r coordinate

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{r}}\right) - \frac{\partial L}{\partial r} = 0 \tag{1.67}$$

giving

$$m\ddot{r} = mr\dot{\theta}^2 - \frac{\partial V}{\partial r} \tag{1.68}$$

Plus a second equation for θ, which since L is independent of θ, is just

$$\frac{d}{dt}(mr^2\dot{\theta}) = 0 \tag{1.69}$$

which tells us that angular momentum is conserved.

The Hamiltonian is also conserved and is given here by

$$H = \frac{1}{2}m\dot{r}^2 + \frac{1}{2}mr^2\dot{\theta}^2 + V \tag{1.70}$$

which is the total energy.

Exercise 1.7: If a system with generalized coordinates x and y has the action

$$S = \int\left(\frac{1}{2}\dot{x}^2 + \frac{1}{2}\dot{y}^2 + \cos\dot{y} - x\right)dt$$

what quantities are conserved?

1.3.4 Hamiltonian and energy

Finally it is worth stressing that the Hamiltonian is not always the energy of the system. As an example consider a bead on a hoop that is being rotated at a fixed angular velocity ω, as shown in figure 1.9.

To be explicit, the hoop is in a vertical plane near the surface of the Earth, where that vertical plane is subject to a steady rotation about a fixed axis passing through the centre of the hoop, due to an external turning force or torque. The fact that the energy is not conserved is due to the fact that the turning force, which is required to maintain the steady rate of rotation, is external to the system. However, we shall show that, even in this case, the Hamiltonian is conserved, even though the Hamiltonian cannot be identified with the energy.

The single coordinate θ (which is a function of time t) as shown in the diagram is sufficient to describe the position of the bead so this is a good generalized coordinate. The kinetic energy is given by

$$T = \frac{1}{2}m(a^2\dot{\theta}^2 + a^2\sin^2\theta\omega^2) \tag{1.71}$$

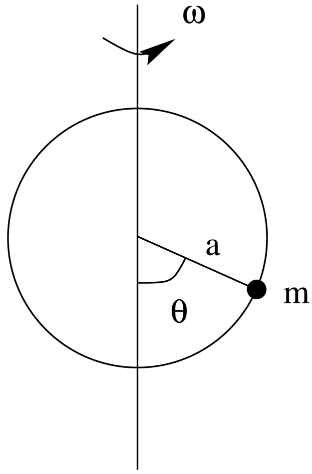

Figure 1.9. A bead on a hoop which is rotating at angular speed ω.

and the potential energy by

$$V = -mga\cos\theta \tag{1.72}$$

Thus

$$L = \frac{1}{2}m(a^2\dot{\theta}^2 + a^2\sin^2\theta\omega^2) + mga\cos\theta \tag{1.73}$$

Since L does not depend on t the Hamiltonian is conserved. In particular

$$H = \frac{1}{2}ma^2\dot{\theta}^2 - \frac{1}{2}ma^2\sin^2\theta\omega^2 - mga\cos\theta \tag{1.74}$$

Although H is conserved the total energy of the system is not since to keep the hoop rotating a constant external torque must be applied, thereby doing work on the system.

Appendix A. Calculus of variation

In this appendix we derive the Euler–Lagrange equation from the calculus of variation, using a general notation which is applicable both to optics and dynamics.

Consider a set of curves $q(s)$ between two points (q_1, s_1) and (q_2, s_2) in the s, q plane (we will only consider curves where the trajectory is single valued at each value of s)[1] (see figure A.1).

Imagine we are interested in one curve that minimizes the quantity

$$S[q(s)] = \int_{s_1}^{s_2} L(q, \dot{q}, s)ds \tag{A.1}$$

[1] For example, in the case of light in two dimensions we identify $q(s)\rightarrow y(x)$, while for a particle in one dimension we identify $q(s)\rightarrow x(t)$, or for a simple pendulum we identify $q(s)\rightarrow \theta(t)$.

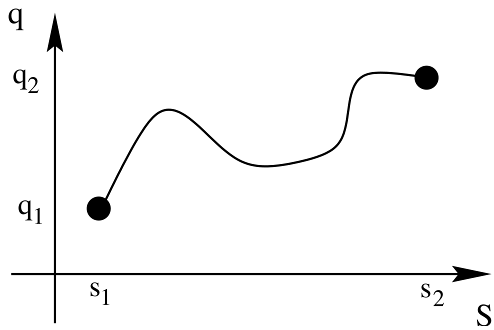

Figure A.1. An arbitrary path in the $q - s$ plane between fixed end points.

L is just a number at each point on a given curve determined by the values of q and s at that point and the gradient $\dot{q} = dq/ds$. The integral sums these numbers along the line.

If the curve that minimizes S is $\bar{q}(s)$ we can write the other curves as deviations from it

$$q(s) = \bar{q}(s) + \delta q(s) \tag{A.2}$$

subject to the boundary conditions

$$\delta q(s_1) = \delta q(s_2) = 0 \tag{A.3}$$

The value of S for these curves varies from the value for $\bar{q}(s)$ by

$$\delta S = S[\bar{q} + \delta q] - S[\bar{q}] \tag{A.4}$$

Since $\bar{q}(s)$ is the *minimum* though $\delta S = 0$ to lowest order in δq.

Let's calculate $S[\bar{q} + \delta q]$ to order δq:

$$\begin{aligned} S[\bar{q} + \delta q] &= \int_{s_1}^{s_2} L(\bar{q} + \delta q, \dot{\bar{q}} + \delta\dot{q}, s)ds \\ &\simeq \int_{s_1}^{s_2} \left(L(\bar{q}, \dot{\bar{q}}, s) + \delta\dot{q}\frac{\partial L}{\partial \dot{q}} + \delta q\frac{\partial L}{\partial q} + \cdots \right) ds \\ &\simeq S[\bar{q}] + \int_{s_1}^{s_2} \left(\delta\dot{q}\frac{\partial L}{\partial \dot{q}} + \delta q\frac{\partial L}{\partial q} \right) ds + \mathcal{O}(\delta q^2) \end{aligned} \tag{A.5}$$

Integrating the second term by parts ($u = \partial L/\partial\dot{q}$, $dv/ds = \delta\dot{q}$ etc)

$$\int_{s_1}^{s_2} \delta\dot{q}\frac{\partial L}{\partial \dot{q}}\, ds = \left[\delta q\frac{\partial L}{\partial \dot{q}} \right]_{s_1}^{s_2} - \int_{s_1}^{s_2} \delta q\frac{d}{ds}\left(\frac{\partial L}{\partial \dot{q}} \right) ds \tag{A.6}$$

The first term vanishes since δq vanishes at the ends of the path.

Thus

$$S[\bar{q} + \delta q] - S[\bar{q}] = -\int_{s_1}^{s_2} \delta q\left(\frac{d}{ds}\left(\frac{\partial L}{\partial \dot{q}} \right) - \frac{\partial L}{\partial q} \right) ds + \cdots \tag{A.7}$$

This is only zero (at order δq) if

$$\boxed{\frac{d}{ds}\left(\frac{\partial L}{\partial \dot{q}}\right) - \frac{\partial L}{\partial q} = 0} \tag{A.8}$$

This is the *Euler–Lagrange equation.*

In general, we will want to solve problems in more than one dimension. For example, there may be several such generalized coordinates, q_i corresponding to three dimensions, x, y, z or multiple angles θ_i. The above formalism is easily adapted for such cases. The definition of the action above in terms of the Lagrangian ($L = T - V$) remains the same, however we now have several coordinates

$$q_i \qquad i = 1, \ldots, n \tag{A.9}$$

In the derivation above of the Euler–Lagrange equation, it is straightforward to take into account deviations in the path in all of these coordinates. We would find that the change in the action of a path close to the minimizing path would have the form

$$\Delta S = -\int_{s_1}^{s_2} \sum_i \delta q_i \left(\frac{d}{ds}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i}\right) ds \tag{A.10}$$

At the minimum the coefficients of each δq_i must vanish independently so we get a set of Euler–Lagrange equations—one associated with each coordinate

$$\boxed{\frac{d}{ds}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0} \tag{A.11}$$

Exercise 1.8: Work through the above derivation in the case where L depends on two coordinates q and p. What two equations must then be satisfied by the minimizing curve?

Appendix B. Mathematics of conservation laws

Under certain circumstances the Euler–Lagrange equation simplifies from a second order equation to a first order equation. This has important applications in Newtonian dynamics, where the physical interpretation is the connection between symmetry and conservation laws, although here we just focus on the mathematics.

There are two particularly interesting special cases:

(1) If $L(q, \dot{q}, s)$ is independent of the coordinate q

$$\frac{d}{ds}\left(\frac{\partial L}{\partial \dot{q}}\right) = 0 \tag{B.1}$$

So

$$\boxed{\frac{\partial L}{\partial \dot{q}} = \text{constant}} \tag{B.2}$$

(2) If $L(q, \dot{q}, s)$ is independent of the coordinate s

$$\frac{dL}{ds} = \frac{\partial L}{\partial q}\dot{q} + \frac{\partial L}{\partial \dot{q}}\ddot{q} \tag{B.3}$$

using the Euler–Lagrange equation gives

$$\frac{dL}{ds} = \frac{d}{ds}\left(\frac{\partial L}{\partial \dot{q}}\right) + \frac{\partial L}{\partial \dot{q}}\ddot{q} \tag{B.4}$$

which is just

$$\frac{d}{ds}\left[L - \dot{q}\frac{\partial L}{\partial \dot{q}}\right] = 0 \tag{B.5}$$

so that

$$\boxed{L - \dot{q}\frac{\partial L}{\partial \dot{q}} = \text{constant}} \tag{B.6}$$

Exercise 1.9: This is an exercise in using calculus of variation outside of optics or dynamics. A smooth curved wire connects the origin to the lower point (x_1, y_1). A bead on the wire slides without friction from rest at the upper to the lower point under the influence of gravity. Its mechanical energy is conserved as it moves along the wire. Choose down to be the positive y direction.

(a) Show that the time, T, required for the bead's journey is

$$T = \frac{1}{\sqrt{2g}}\int_0^{x_2}\sqrt{\frac{(1 + y'^2)}{y}}\,dx$$

(b) Given that the integrand of the above integral is independent of x show that the curve $y(x)$ making T stationary satisfies the differential equation

$$\frac{dy}{dx} = \sqrt{\frac{(b - y)}{y}}$$

(c) Change the dependent variable from y to ϕ where $y = b\sin^2\phi/2$ and show that the above can be integrated to give the *brachistochrone*

$$x = b/2(\phi - \sin\phi)$$