Classical Dynamics (2022-2023)*

Contents

*An evolution of notes prepared by Jack Davies and Chris Ford in Autumn 2020.

1 Dynamics

This aim of this module is study dynamics through variational principles which are often more efficient than direct solution of Newtonian equations of motion. However, ideas from Newtonian dynamics are important in Lagrangian and Hamiltonian mechanics. Basic concepts in the Newtonian approach are developed in this chapter.

1.1 Particles

Dynamics is concerned with the time-evolution of physical systems. That is if the state of a system is known at an 'initial' time what is the state of the system at a later time t? To make progress with this ambitious programme, approximations and idealisations are necessary. For example we may treat a planet as a point particle or ignore friction when analysing the motion of a pendulum. In this module we will be concerned with Classical Dynamics as opposed to Quantum Mechanics. However, some of the mathematical structure is similar to that used in Quantum Mechanics.

A key concept in Classical Dynamics is that of a *particle*. A particle is a massive object with zero size. It may have other physical characteristics such as electric charge. The state of the particle is specified by its position. This can be written as a position vector $\mathbf{r}$ or as coordinates ($\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$). A physical system can comprise a single particle or many particles. The *velocity* of a particle is the rate of change of its position

$$\mathbf{v} = \dot{\mathbf{r}} = \frac{dx}{dt}\mathbf{i} + \frac{dy}{dt}\mathbf{j} + \frac{dz}{dt}\mathbf{k}.$$

The speed v of the particle is the magnitude of the velocity. Throughout the module a dot denotes differentiation with respect to time t, e.g. $\dot{\mathbf{r}} = d\mathbf{r}/dt$.

The *linear momentum*, $\mathbf{p}$, of a particle is defined as

$$\mathbf{p} = m\mathbf{v}. \tag{1}$$

In relativistic mechanics a different definition is used

$$\mathbf{p} = \frac{m\mathbf{v}}{\sqrt{1 - \frac{v^2}{c^2}}}, \tag{2}$$

where c is the speed of light (if $v^2 \ll c^2$ this reduces to the previous definition). Lagrangian and Hamiltonian Mechanics allow more general relationships between momenta and velocities.

Newton's laws are a framework for describing the motion of systems of particles:

N1: A particle moves at constant velocity until acted upon by a force.

N2: The rate of change of momentum, $\mathbf{p}$, of a particle is equal to the force, $\mathbf{F}$, acting on the particle.

N3: The total momentum of system of an isolated system of particles is a constant of the motion.

Arguably, only the third is really a physical law[1] as the first two laws amount to the definition of force. If the forces on a system of particles are specified then Newton's 2nd law provides *equations of motion* for the system. Using the non-relativistic momentum Newton's second law takes the form

$$\mathbf{F} = m\mathbf{a},$$

where $\mathbf{a} = d\mathbf{v}/dt$ is the acceleration of the particle.

A particularly instructive example is the Kepler problem describing the gravitational interaction of two particles. According to Newton's Universal Law of Gravitation two particles with masses M and m are both subject to an attractive force of magnitude

$$F = \frac{GMm}{r^2},$$

where G is Newton's gravitational constant and r is the distance between the particles. Assuming the particle of mass M is fixed at the origin (if $M \gg m$ this is a consistent approximation) the force on the particle of mass m is

$$\mathbf{F} = -\frac{GMm\mathbf{r}}{r^3},$$

where the negative sign means that the force points toward the origin (the position of the other particle). Newton's 2nd law can be written as

$$m\ddot{\mathbf{r}} = -\frac{GMm\mathbf{r}}{r^3}. \tag{3}$$

The motion of the particle of mass m can be inferred by solving this (vector) equation of motion. Equation (3) can be recast as three coupled 2nd order ODEs.

The Kepler problem is a special case of a *central force* which is a position-dependent force of the form

$$\mathbf{F}(\mathbf{r}) = F(r)\frac{\mathbf{r}}{r},$$

where $r = \sqrt{x^2 + y^2 + z^2}$ is the distance of the particle from the origin. Here the force is

---

[1] As this is a real law it is sometimes wrong! For example the total momentum of a set of charged particles is not conserved. This shortcoming can be repaired by ascribing momentum to the electro-magnetic field (or radiation).

directed from origin towards particle (or other way round) and magnitude of force depends only on the distance from origin.

1.2 Energy

Suppose a particle moves from a position $\mathbf{r}$ to a position $\mathbf{r}+\delta\mathbf{r}$ while acted upon by a constant force $\mathbf{F}$. The *work done*, δW, by the force is defined to be the dot product

$$\delta W = \mathbf{F} \cdot \delta\mathbf{r}.$$

If the trajectory of the particle is given by $\mathbf{r} = \mathbf{r}(t)$ one has $\delta\mathbf{r} \approx \dot{\mathbf{r}}(t)\delta t$ for small δt and so $\delta W = \mathbf{F} \cdot \dot{\mathbf{r}}(t)\delta t$. The work done by a time-dependent force, $\mathbf{F}(t)$, over a finite time interval can be written as an integral

$$W = \int_{t_1}^{t_2} \mathbf{F}(t) \cdot \dot{\mathbf{r}}(t) \ dt.$$

Using N2 $\mathbf{F}(t) = m\ddot{\mathbf{r}}(t)$ (assuming non-relativistic speeds) the work integral can be written in the form

$$W = m \int_{t_1}^{t_2} \ddot{\mathbf{r}}(t) \cdot \dot{\mathbf{r}}(t) \ dt.$$

The integrand is a total derivative

$$\ddot{\mathbf{r}}(t) \cdot \dot{\mathbf{r}}(t) = \frac{d}{dt}\left(\frac{1}{2}\dot{\mathbf{r}}(t) \cdot \dot{\mathbf{r}}(t)\right) = \frac{1}{2}\frac{d}{dt}v^2(t).$$

Accordingly,

$$W = T(t_2) - T(t_1),$$

where T is the *kinetic energy*

$$T = \frac{1}{2}mv^2.$$

The change in the kinetic energy of a particle is equal to the work done on the particle. This fits with the elementary idea that 'energy is the capacity to do work'.

The discussion of the kinetic energy above made no assumptions about the nature of the forces. We now consider the special case of *conservative forces* which are derived from a potential energy function V. For a particle moving in three dimensions a conservative force has the form $\mathbf{F} = -\nabla V$ where $V = V(\mathbf{r}, t)$ is a function of position, $\mathbf{r}$, and time, t. In one dimension $F = -\partial V/\partial x$ where V depends on x and t. Note that conservative forces are position dependent. In one dimension any position dependent force is conservative as $F(x, t) = -\partial V(x, t)/\partial x$ can be integrated with respect to x to obtain $V(x, t)$. In higher dimensions not all position dependent forces are conservative. For example, in three dimensions a position dependent force $\mathbf{F}(\mathbf{r}, t)$ can be derived from a potential energy function only if $\nabla \times \mathbf{F} = 0$.

The Kepler problem involves a conservative force derived from the potential energy

$$V = -\frac{GMm}{\sqrt{x^2 + y^2 + z^2}}.$$

Another important example is terrestrial gravity where masses are subject to a downward force and in the absence of other forces have a constant acceleration of magnitude g (approximately 9.8ms^{-2}). The force on such a mass is $\mathbf{F} = -mg\mathbf{k}$ which is derived from the potential energy

$$V = mgz.$$

This ubiquitous formula appears in applications of terrestrial gravity such as the motion of pendulums and rigid bodies.

A simple one-dimensional example is the quadratic potential $V = \frac{1}{2}kx^2$ describing the displacement from equilibrium of a spring. Here N2 has the form

$$m\ddot{x} = -\frac{dV}{dx} = -kx.$$

This is the equation of a simple harmonic oscillator with angular frequency $\sqrt{k/m}$. Small oscillations about an equilibrium point can be approximated by a simple harmonic oscillator. For example, consider a particle of unit mass with potential energy $V(x) = -2x^2 + x^4$ giving the force $F(x) = -V'(x) = 4(x - x^3)$; there are equilibria at $x = 0$, $x = \pm 1$. Near $x = 1$ the motion resembles a harmonic oscillator as N2 is approximately $\ddot{y} = -\omega^2 y$ where $y = x - 1$ and $\omega^2 = 8$. The angular frequency of small oscillations about $x = 1$ (and $x = -1$) is $\sqrt{8}$.

If the potential energy has no explicit time-dependence (as in the examples above) then $E = T + V$ (referred to as the total energy) is a constant of the motion. In contrast to kinetic energy, potential energy is only meaningful if the forces acting on a system are conservative. For example, if a (velocity-dependent) damping force is added to the spring example

$$m\ddot{x} = -kx - m\gamma\dot{x},$$

where γ is a constant, the force is not conservative. Magnetic forces are another example of non-conservative forces; the force on a charged particle in a magnetic field depends on the velocity and cannot be written in the form $-\nabla V$.

1.3   Angular Momentum

If a particle is *not* acted upon by a force its momentum will remain constant. If acted on by a force $\mathbf{F}$ its momentum changes according to Newton's 2nd law

$$\frac{d\mathbf{p}}{dt} = \mathbf{F}.$$

There is an analogous result for *angular momentum* and *torque.*

Angular momentum is defined with respect to a fixed point (often chosen to be the origin).

The angular momentum of a particle with respect to the origin is defined to be the vector (cross product)

$$\mathbf{L} = \mathbf{r} \times \mathbf{p}.$$

(angular momentum with respect to position $\mathbf{a}$ is defined as $\mathbf{L} = (\mathbf{r} - \mathbf{a}) \times \mathbf{p}$).

The torque on a particle is defined with respect to a fixed point (often chosen to be the origin).

The torque on a particle with respect to the origin is defined to be the vector (cross product)

$$\mathbf{K} = \mathbf{r} \times \mathbf{F}.$$

(about a point with position vector $\mathbf{a}$, $\mathbf{K} = (\mathbf{r} - \mathbf{a}) \times \mathbf{F}$).

If $\mathbf{L}$ and $\mathbf{K}$ are defined with respect to the same point (henceforth taken to be the origin)

$$\frac{d\mathbf{L}}{dt} = \mathbf{K}.$$

To prove this take the definition of $\mathbf{L}$ and differentiate with respect to t:

$$\frac{d\mathbf{L}}{dt} = \dot{\mathbf{r}} \times \mathbf{p} + \mathbf{r} \times \dot{\mathbf{p}}$$

(one can apply the product rule to a vector cross product).

Now $\dot{\mathbf{r}} \times \mathbf{p} = m\dot{\mathbf{r}} \times \dot{\mathbf{r}} = 0$ so

$$\frac{d\mathbf{L}}{dt} = \mathbf{r} \times \dot{\mathbf{p}} = \mathbf{r} \times \mathbf{F},$$

using Newton's 2nd law.

1.4   Polar Coordinates

The position of a particle is described by three cartesian coordinates x, y and z. Cartesian coordinates are not always convenient for solving equations of motion. For example, solving the Kepler problem is clumsy using cartesian coordinates. In fact, motion in the Kepler problem is planar. This is true for any central force. Analysis of motion in a central force is effectively a $2d$ problem. Without loss of generality one may assume motion is in the xy

plane[2] (i.e. the $z = 0$ plane). Accordingly

$$\mathbf{F} = F(r)\frac{x\mathbf{i} + y\mathbf{j}}{r} \quad \text{with} \quad r = \sqrt{x^2 + y^2}.$$

We now write N2, $m\ddot{\mathbf{r}} = \mathbf{F}$, using polar coordinates[3] (the analysis also holds for non-central forces provided motion is planar). How to express this in polar coordinates? Consider the velocity

$$\mathbf{v} = \dot{\mathbf{r}} = \dot{x}(t)\mathbf{i} + \dot{y}(t)\mathbf{j}.$$

Now $x(t) = r(t)\cos\theta(t)$ so that $\dot{x} = \dot{r}\cos\theta - r\dot{\theta}\sin\theta$. Similarly $\dot{y} = \dot{r}\sin\theta + r\dot{\theta}\cos\theta$. Therefore

$$\mathbf{v} = (\dot{r}\cos\theta - r\dot{\theta}\sin\theta)\mathbf{i} + (\dot{r}\sin\theta + r\dot{\theta}\cos\theta)\mathbf{j}.$$

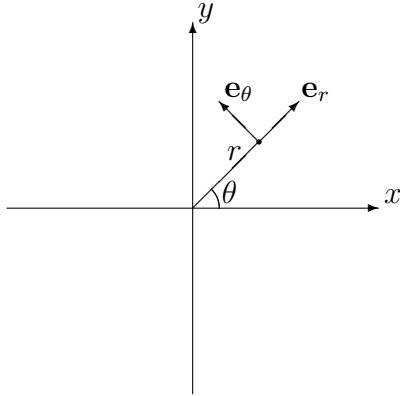This unwieldy formula uses basis vectors $\mathbf{i}$ and $\mathbf{j}$ whereas the coefficients are expressed in polar coordinates. Rewrite $\mathbf{v}$ as

$$\mathbf{v} = \dot{r}(\cos\theta\mathbf{i} + \sin\theta\mathbf{j}) + r\dot{\theta}(-\sin\theta\mathbf{i} + \cos\theta\mathbf{j}) \equiv \dot{r}\mathbf{e}_r + r\dot{\theta}\mathbf{e}_\theta.$$

where

$$\mathbf{e}_r = \cos\theta\mathbf{i} + \sin\theta\mathbf{j}, \quad \mathbf{e}_\theta = -\sin\theta\mathbf{i} + \cos\theta\mathbf{j}.$$

are 'new' basis vectors adapted to polar coordinates. $\mathbf{e}_r$ is a unit vector pointing in the direction of increasing r and $\mathbf{e}_\theta$ is a unit vector pointing in the direction of increasing θ. $\mathbf{e}_r$ and $\mathbf{e}_\theta$ are orthogonal to each other; $\mathbf{e}_r \cdot \mathbf{e}_r = \mathbf{e}_\theta \cdot \mathbf{e}_\theta = 1, \quad \mathbf{e}_r \cdot \mathbf{e}_\theta = 0.$

Unlike $\mathbf{i}$ and $\mathbf{j}$ the new basis vectors $\mathbf{e}_r$ and $\mathbf{e}_\theta$ are time dependent; by definition $\mathbf{e}_r$ and $\mathbf{e}_\theta$ depend on θ and unless θ is constant the basis vectors are time-dependent. Using the definition of $\mathbf{e}_r$ and $\mathbf{e}_\theta$ it follows that

$$\frac{d}{dt}\mathbf{e}_r = \dot{\theta}\,\mathbf{e}_\theta, \quad \frac{d}{dt}\mathbf{e}_\theta = -\dot{\theta}\,\mathbf{e}_r.$$

---

[2]Equivalently one may assume that $\mathbf{L}$ is parallel (or anti-parallel) to the z-axis.
[3]Polar coordinates r and θ are defined through $x = r\cos\theta$, $y = r\sin\theta$.

The acceleration is

$$\mathbf{a} = \dot{\mathbf{v}} = (\ddot{r} - r\dot{\theta}^2)\mathbf{e}_r + (r\ddot{\theta} + 2\dot{r}\dot{\theta})\mathbf{e}_\theta. \tag{4}$$

To write N2 in polar coordinates express the force $\mathbf{F}$ with respect to new basis vectors $\mathbf{e}_r$ and $\mathbf{e}_\theta$

$$\mathbf{F} = F_r\mathbf{e}_r + F_\theta\mathbf{e}_\theta.$$

Newton's 2nd law can be decomposed as follows

$$m(\ddot{r} - r\dot{\theta}^2) = F_r, \quad m(r\ddot{\theta} + 2\dot{r}\dot{\theta}) = F_\theta.$$

For a central force $F_\theta = 0$ so that $r\ddot{\theta} + 2\dot{r}\dot{\theta} = 0$. This implies that $r^2\dot{\theta}$ is a constant of the motion (conservation of angular momentum).

The acceleration formula (4) is extremely useful. However, it is specific to polar coordinates. What about other coordinate systems? We shall see that a Lagrangian approach allows one to obtain equations of motion in general coordinate systems without having to recompute the acceleration every time. Indeed the equations of motion can be formulated in a coordinate-independent form.

2 Lagrangian Mechanics

Lagrangian mechanics is developed in this chapter. A nice introduction to Lagrangian mechanics is given in chapter 19 of the 'Feynman Lectures on Physics', volume 2. Feynman begins with an account of how he was introduced to the Principle of Least Action by his high school teacher Mr Bader. The chapter starts with quite basic material but quickly moves on to advanced applications. Lagrangian mechanics is developed in chapter 2 of Goldstein, Safko and Poole and the first 15 sections of 'Mechanics', 3rd edition by L. D. Landau and E. M. Lifshitz.

2.1 Hamilton's Principle

In Lagrangian mechanics Newton's second law of motion is replaced with a variational principle 'Hamilton's Principle' or the 'Principle of Least Action'. This provides a coordinate independent formulation of the equations of motion. However, we will initially establish the connection between Newton's second law and Hamilton's Principle through cartesian coordinates.

Lagrangian mechanics is particularly simple for *conservative forces*, i.e., forces which are derived from a potential energy function V. Consider a particle of mass m moving in one dimension subject to a position dependent force $F(x,t) = -\partial V(x,t)/\partial x$ so that Newton's second law has the form

$$m\ddot{x} = -\frac{\partial V}{\partial x}.$$

This can be rewritten as

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{x}}\right) = -\frac{\partial V}{\partial x},$$

where $T = \frac{1}{2}m\dot{x}^2$ is the kinetic energy of the particle. As T does not depend on x and V does not depend on $\dot{x}$, it follows that

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) = \frac{\partial L}{\partial x}, \tag{5}$$

where

$$L(x,\dot{x},t) = T - V = \frac{1}{2}m\dot{x}^2 - V(x,t). \tag{6}$$

This is an Euler-Lagrange equation, and so the equation of motion is a stationarity condition

for the *action* integral

$$S = \int_{t_1}^{t_2} L(x, \dot{x}, t) \, dt,$$

with $x(t_1)$ and $x(t_2)$ fixed. If the potential energy is time-independent the Beltrami formula applies, giving

$$\dot{x}\frac{\partial L}{\partial \dot{x}} - L = \dot{x} \cdot m\dot{x} - \left(\frac{1}{2}m\dot{x}^2 - V\right) = \frac{1}{2}m\dot{x}^2 + V(x),$$

is constant. This is conservation of energy.

Hamilton's Principle states that the time evolution of a physical system from time t_1 to time t_2 is such that the action

$$S = \int_{t_1}^{t_2} L \, dt,$$

is stationary with respect to variations of the coordinates.

So far we have demonstrated that for a particle moving in one dimension subject to a conservative force, a suitable Lagrangian is $L = T - V$. In fact, this is true for any non-relativistic system subject to conservative forces. Consider how this works for a particle of mass m moving in the plane with cartesian coordinates x and y. Assuming the force is conservative

$$m\ddot{x} = -\frac{\partial V}{\partial x} \text{ and } m\ddot{y} = -\frac{\partial V}{\partial y}.$$

These equations of motion can be recast as two Euler-Lagrange equations

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

where

$$L(x, y, \dot{x}, \dot{y}, t) = \frac{1}{2}m\left(\dot{x}^2 + \dot{y}^2\right) - V(x, y, t).$$

The two Euler-Lagrange equations are stationarity conditions with respect to independent variations of x and y.

The form of the Euler-Lagrange equations is independent of the choice of coordinates. For example, consider the Kepler problem where $V(x, y) = -k/\sqrt{x^2 + y^2}$ with $k = GMm$. Converting the problem to polar coordinates (r, θ), the Euler-Lagrange equations are simply

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{r}}\right) - \frac{\partial L}{\partial r} = 0 \text{ and } \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right) - \frac{\partial L}{\partial \theta} = 0.$$

The equations are stationarity conditions for independent variations of r and θ. It is necessary

to rewrite the Lagrangian in the new coordinates. The potential energy can be written as $V = -k/r$ and the kinetic energy is[4]

$$T = \frac{m}{2} \left(\dot{x}^2 + \dot{y}^2 \right) = \frac{m}{2} \left(\dot{r}^2 + r^2 \dot{\theta}^2 \right).$$

In polar coordinates the Lagrangian is

$$L = \frac{1}{2} m (\dot{r}^2 + r^2 \dot{\theta}^2) + \frac{k}{r} = L(r, \dot{r}, \dot{\theta}), \tag{7}$$

and the equations of motion are

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{r}} \right) - \frac{\partial L}{\partial r} = \frac{d}{dt} m \dot{r} - m r \dot{\theta}^2 + \frac{k}{r^2} = 0,$$

and

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\theta}} \right) - \frac{\partial L}{\partial \theta} = \frac{d}{dt} m r^2 \dot{\theta} = 0.$$

The second equation can be integrated to give $m r^2 \dot{\theta} =$ constant; this is conservation of angular momentum.

In a Newtonian approach, these equations can be obtained via the acceleration formula (4). Here the force is $\mathbf{F} = -k \mathbf{e}_r / r^2$. The Lagrangian approach has the advantage that we don't have to use vectors - all we need is the form of the *scalar* quantity $L = T - V$ in the coordinate system we are using. For example, in spherical polar coordinates (r, θ, ϕ), where we have that $x = r \sin \theta \cos \phi$, $y = r \sin \theta \sin \phi$ and $z = r \cos \theta$ the kinetic energy is
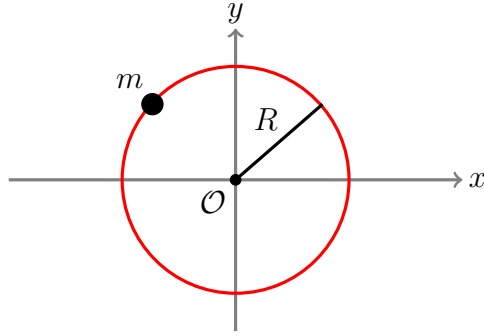
$$T = \frac{1}{2} m \left(\dot{r}^2 + r^2 \dot{\theta}^2 + r^2 \sin^2 \theta \ \dot{\phi}^2 \right). \tag{8}$$

We now consider the Lagrangian formulation of mechanical systems subject to one or more constraints. For example, consider a bead of mass m bead moving without friction along a circular wire. The constraint for this problem is $x^2 + y^2 = R^2$, where the circular wire has radius R. Without the constraint, the Lagrangian is given by

$$L = \frac{1}{2} m (\dot{x}^2 + \dot{y}^2) - V(x, y),$$

where V describes the external force on the bead (e.g. if the only external force is gravity $V = mgy$)

---

[4]This can be obtained through an explicit calculation of $\dot{x}^2 + \dot{y}^2$. Alternatively, use $\mathbf{v} = \dot{r} \mathbf{e}_r + r \dot{\theta} \mathbf{e}_\theta$ to compute $\mathbf{v} \cdot \mathbf{v}$.

Now enforce the constraint with a Lagrange multiplier[5]

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) - V(x, y) + \lambda(t)(x^2 + y^2 - R^2).$$

The Euler-Lagrange equations are

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) - \frac{\partial L}{\partial x} = m\ddot{x} + \frac{\partial V}{\partial x} - 2\lambda x = 0,$$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{y}}\right) - \frac{\partial L}{\partial y} = m\ddot{y} + \frac{\partial V}{\partial y} - 2\lambda y = 0,$$

$$\frac{\partial L}{\partial \lambda} = x^2 + y^2 - R^2 = 0.$$

The x-component of the force can be read off from the first equation $F_x = m\ddot{x} = -\partial V/\partial x + 2\lambda(t)x$. Here $2\lambda(t)x$ is the x-component of the constraint force. The constraint force is perpendicular to the wire. The Lagrange multipler $\lambda(t)$ fixes the magnitude and direction (towards or away from the origin) of the constraint force.

The constrained minimisation of the action can be accomplished without a Lagrange multiplier by first 'solving' the constraint. Minimise

$$s = \int_{t_1}^{t_2} (T - V)\, dt,$$

while restricting to trajectories which satisfy the constraint $r = R$. Converting to polar coordinates and assuming $r = R$,

$$s = \int_{t_1}^{t_2} \left[\frac{1}{2}mR^2\dot{\theta}^2 - V(R, \theta)\right] dt,$$

where $V(r, \theta)$ is the potential energy in polar coordinates. For gravity where $V = mgy =$

---

[5]Note that unlike for the hanging rope problem discussed in the primer on Calculus of Variations the Lagrange multiplier is a function of time. This is because the constraint holds for every time.

13

$mgr\sin\theta$, we have

$$s = \int_{t_1}^{t_2} \underbrace{\left[\frac{1}{2}mR^2\dot\theta^2 - mgR\sin\theta\right]}_{L=L(\theta,\dot\theta)} dt,$$

Since the Lagrangian is a function of θ and $\dot\theta$ only, we only have one Euler-Lagrange equation, namely

$$\frac{d}{dt}\left(\frac{\partial L}{\partial\dot\theta}\right) - \frac{\partial L}{\partial\theta} = mR^2\ddot\theta + mgR\cos\theta = 0,$$

so that $\ddot\theta = -g\cos\theta/R$.

The constraint $x^2 + y^2 = R^2$ or $r = R$ is an example of a *holonomic* constraint.

A holonomic constraint is where some function of the coordinates and time is fixed to be zero.

For the bead on a hoop this is $x^2 + y^2 - R^2 = 0$ or in polar coordinates $r - R = 0$. In three dimensions a holonomic constraint has the form $\varphi(x,y,z,t) = 0$. For example, a particle constrained to a sphere $\varphi(x,y,z,t) = x^2 + y^2 + z^2 - R^2 = 0$ or $r = R$ using polar coordinates.

In general, for a particle subject to an external (conservative) force $\mathbf{F} = -\nabla V$ and a holonomic constraint $\varphi = 0$, one can consider either

$$\text{(a)} \quad s = \int_{t_1}^{t_2} (T - V)\, dt,$$

restricting attention to trajectories satisfying the constraint by eliminating one of the coordinates

OR

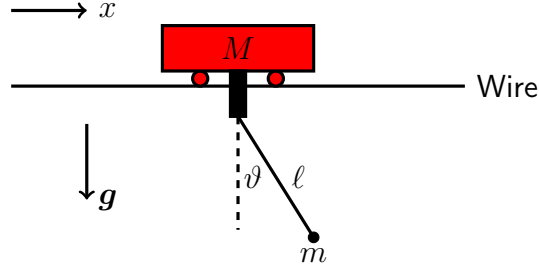$$\text{(b)} \quad S = \int_{t_1}^{t_2} \left[T - V + \lambda(t)\varphi\right] dt.$$

Usually, option (a) is the most practical. For motion without gravity on a sphere of radius R, use spherical polar coordinates and simply set $r = R$ in the potential energy and in the kinetic energy formula (8).

$$s = \int_{t_1}^{t_2} \underbrace{\left[\frac{1}{2}mR^2(\dot\theta^2 + \sin^2\theta\,\dot\phi^2) - V(R,\theta,\phi)\right]}_{L=L(\theta,\phi,\dot\theta,\dot\phi)} dt,$$

which means we can reduce the number of Euler-Lagrange equations to two.

14

Another example involving holonomic constraints is a simple pendulum on a trolley as shown in the diagram below

Let x be the horizontal position of the trolley, assume the mass of the trolley and the pendulum bob are given by M and m, respectively. Furthermore, assume the bob is suspended under gravity g by a string or limb of length ℓ, at an angle θ to the vertical. Suppose the position of the bob is (X, Y). Neglecting the mass of the limb, the potential energy of the bob is given by $V_b = mgY = mg(-\ell \cos \theta)$. We need to find the kinetic energy of the bob and the trolley. The kinetic energy of the trolley is simply $T_t = M\dot{x}^2/2$, while the kinetic energy of the bob is slightly more complicated to evaluate. Since $X = x + \ell \sin \theta$ and $Y = -\ell \cos \theta$

$$\dot{X} = \dot{x} + \ell \dot{\theta} \cos \vartheta, \quad \dot{Y} = \ell \dot{\theta} \sin \theta,$$

and the kinetic energy of the bob is

$$T_b = \frac{1}{2} m (\dot{X}^2 + \dot{Y}^2) = \frac{1}{2} m \left[\dot{x}^2 + 2\ell \dot{x} \dot{\theta} \cos \theta + \ell^2 \dot{\theta}^2 \right].$$

Thus, the full Lagrangian takes the form

$$L = T_t + T_b - V_b$$
$$= \frac{1}{2}(M + m)\dot{x}^2 + \frac{1}{2} m \left[2\ell \dot{x} \dot{\theta} \cos \theta + \ell^2 \dot{\theta}^2 \right] + mg\ell \cos \theta.$$

Since the Lagrangian is independent of x, $\partial L / \partial \dot{x} = (M + m)\dot{x} + m\ell \dot{\theta} \cos \theta$ is a constant. This constant is p_x, the x-component of the momentum of the system. Accordingly,

$$\dot{x} = \frac{p_x - m\ell \dot{\theta} \cos \theta}{M + m}. \tag{9}$$

The Euler-Lagrange equation for θ is

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}} \right) - \frac{\partial L}{\partial \theta} = \frac{d}{dt}\left(m\ell^2 \dot{\theta} + m\ell \dot{x} \cos \theta \right) + \sin \theta (m\ell \dot{x} \dot{\theta} + mgl) = 0.$$

Using (9) to eliminate $\dot{x}$ yields a complicated ODE for θ. Instead consider *small oscillations* so that we approximate (9) with

$$\dot{x} = \frac{p_x - m\ell\dot{\theta}}{M + m}.$$

and in this approximation the ODE for θ is[6]

$$\frac{d}{dt}\left[m\ell^2\dot{\theta} + \frac{m\ell(p_x - m\ell\dot{\theta})}{M + m}\right] + mg\ell\theta = 0,$$

which simplifies to

$$\ddot{\theta} = -\frac{(M + m)g}{M\ell}\theta$$

Hence, the angular frequency for small oscillation is

$$\omega = \sqrt{\frac{(M + m)g}{M\ell}}.$$

2.2 Non-Conservative Forces

So far we have focussed on conservative forces which can be described through Lagrangians of the form $L = T - V$. However, there are other Lagrangians. For example, the motion of a free relativistic particle moving in three dimensions is governed by the Lagrangian

$$L = -mc^2\sqrt{1 - \frac{v^2}{c^2}} = -mc^2\sqrt{1 - \frac{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}{c^2}},$$

where c is the speed of light. Assuming $v/c \ll 1$, then

$$\left(1 - \frac{v^2}{c^2}\right)^{1/2} \approx 1 - \frac{v^2}{2c^2} - \frac{v^4}{8c^4} + \cdots$$

so that $L \approx \frac{1}{2}mv^2 - mc^2$; the first term is the kinetic energy of the particle and the second term is minus the rest energy $(E = mc^2)$ of the particle. In this case the action is mc^2 times minus the time measured from the perspective of a particle, or proper time. In this case the 'Principle of Least Action' can be reformulated as the 'Principle of Maximal Ageing'.

Another Lagrangian comes from consideration of magnetic forces. A particle with mass m and electric charge q in a magnetic field $\boldsymbol{B} = \boldsymbol{B}(\boldsymbol{r})$ satisfies the Lorentz force law

$$\boldsymbol{F} = q(\boldsymbol{v} \times \boldsymbol{B}).$$

---

[6]Here θ and its derivatives are treated as small, e.g. $\dot{\theta}\sin\theta \approx \dot{\theta}\,\theta$ is neglected as a second order term.

This force is not conservative as it is velocity dependent. A (non-relativistic) Lagrangian which gives rise to this equation of motion is

$$L = \frac{1}{2}mv^2 + q\boldsymbol{v} \cdot \boldsymbol{A} = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) + q(\dot{x}A_x + \dot{y}A_y + \dot{z}A_z). \tag{10}$$

Here $\boldsymbol{A}(x, y, z) = A_x\mathbf{i} + A_y\mathbf{j} + A_z\mathbf{k}$ is a vector potential for the magnetic field defined through $\mathbf{B} = \nabla \times \mathbf{A}$. A vector potential can be introduced because the magnetic field satisfies $\nabla \cdot \mathbf{B} = 0$ (this is one of Maxwell's equation). The relativistic version of this Lagrangian is

$$L = -mc^2\sqrt{1 - \frac{v^2}{c^2}} + q\boldsymbol{v} \cdot \boldsymbol{A}.$$

A weakness of Lagrangian methods is that there is not a natural formulation of dissipative effects such as friction (see problem sheet 2).

2.3   Generalised Coordinates

We now introduce the notion of generalised coordinates. Consider a system described by N coordinates q_i for $i = 1, 2, \ldots, N$. Assume that any holonomic constraints[7] have been eliminated, say $N + P$ coordinates with P constraints - solve constraints to produce N coordinates. The generalised velocities are given by $\dot{q}_i$. Further assume the time evolution is governed by the Lagrangian

$$L = L(q_1, q_2, \ldots, q_N, \dot{q}_1, \dot{q}_2, \ldots, \dot{q}_N, t) = L(q_i, \dot{q}_i, t),$$

a function of $2N + 1$ variables. The action integral is

$$S = \int_{t_1}^{t_2} L(q_i, \dot{q}_i, t)\, dt$$

and there are N Euler-Lagrange equations of the form

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0 \text{ for } i = 1, 2, \ldots, N.$$

In elementary mechanics momentum is defined through the vector equation $\mathbf{p} = m\mathbf{v}$. In Lagrangian mechanics N generalised momenta or canonical momenta are defined through the Lagrangian via

$$p_i = \frac{\partial L}{\partial \dot{q}_i}, \text{ for } i = 1, 2, \ldots, N.$$

---

[7]With $N + 1$ coordinates $r_1, r_2,...,r_{N+1}$ a holonomic constraint has the form, $\varphi(r_1, r_2, ..., r_{N+1}, t) = 0$, which allows one coordinate to be eliminated leaving N coordinates $q_1, q_2,..., q_N$.

A simple example is a particle of mass m moving in three dimensions where for cartesian coordinates the Lagrangian is $L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - V(x, y, z)$. Here the canonical momenta are

$$p_x = \frac{\partial L}{\partial \dot{x}} = m\dot{x}, \quad p_y = \frac{\partial L}{\partial \dot{y}} = m\dot{y}, \quad p_z = \frac{\partial L}{\partial \dot{z}} = m\dot{z},$$

which is exactly what is expected. However, the generalised momenta are not necessarily components of the linear momentum, $\boldsymbol{p} = m\boldsymbol{v}$. To see this, consider a particle in a magnetic field. Using the Lagrangian (10)

$$p_x = \frac{\partial L}{\partial \dot{x}} = m\dot{x} + qA_x, \quad p_y = \frac{\partial L}{\partial \dot{y}} = m\dot{y} + qA_y, \quad p_z = \frac{\partial L}{\partial \dot{z}} = m\dot{z} + qA_z,$$

For the Kepler problem in polar coordinates the Lagrangian (7) gives

$$p_r = \frac{\partial L}{\partial \dot{r}} = m\dot{r}, \quad p_\theta = \frac{\partial L}{\partial \dot{\theta}} = mr^2\dot{\theta}.$$

Here p_θ is the angular momentum.

If the Lagrangian is independent of a generalised coordinate the coordinate is said to be *cyclic*. It follows from that Euler-Lagrange equations that the momentum corresponding to a cyclic coordinate is a constant of the motion. For example, if $\partial L/\partial q_1 = 0$ then q_1 is cyclic and $p_1 = \partial L/\partial \dot{q}_1$ is constant.

Using the canonical momentum, we can express the Beltrami formula in a different way. For a one dimensional Lagrangian $L(q, \dot{q})$ with no explicit time-dependence, this formula would read as.

$$H = p\dot{q} - L = \text{constant}.$$

In general, consider a time-independent Lagrangian $L(q_i, \dot{q}_i)$ for $i = 1, 2, \cdots, N$. Then

$$H = \sum_{i=1}^{N} p_i \dot{q}_i - L = \text{constant}, \tag{11}$$

if $q_i(t)$ is a solution of the equations of motion. The proof is an exercise on problem sheet 3. For N coordinates, there is only one conserved quantity from Beltrami unless the Lagrangian can be separated into independent parts, e.g., $L(q_1, q_2, \dot{q}_1, \dot{q}_2) = L_1(q_1, \dot{q}_1) + L_2(q_2, \dot{q}_2)$ in two dimensions. In this case, we can apply Beltrami to the individual Lagrangians separately.

2.4   Preview of Noether's theorem

Noether's theorem is based on the idea that conservation laws are associated with 'symmetries'. The simplest case is the statement that if a coordinate, say q_1, is cyclic then the corresponding momentum $p_1 = \partial L/\partial \dot{q}_1$ is a constant of the motion. For example, in the Kepler problem θ is cyclic and $p_\theta = mr^2\dot{\theta}$ is a conserved quantity (the angular momentum). Although this result is very useful it depends on a specific coordinate choice, whereas the Euler-Lagrange equations are coordinate independent. In Cartesian coordinates, the Lagrangian for the Kepler problem is given by

$$L = \frac{m}{2}\left(\dot{x}^2 + \dot{y}^2\right) + \frac{k}{\sqrt{x^2 + y^2}}.$$

Neither x nor y is cyclic so it is not obvious from this form of the Lagrangian that $yp_x - xp_y$ is constant. However, the Lagrangian is invariant under the rotation

$$x' = x\cos s + y\sin s, \quad y' = y\cos s - x\sin s.$$

For small s this reduces to $x' = x + ys$, $y' = y - xs$ (neglecting higher order terms). This can also be written in the form

$$\delta x = x' - x = ys, \quad \delta y = y' - y = -xs.$$

More generally, if the Lagrangian is invariant under the deformation

$$\delta q_i = q_i' - q_i = sf_i(q_1, q_2, ..., q_N, t),$$

then

$$\sum_{i=1}^{N} p_i f_i = \text{constant of motion}.$$

There are more general (and complicated) statements of Noether's theorem which use the action rather than the Lagrangian. In section 4 of these notes, a simple but general statement of this result is developed using Hamiltonian methods.

3 Hamiltonian Mechanics

In the previous chapter Hamilton's Principle was developed using a Lagrangian function. This chapter considers an alternative approach using the Hamiltonian. The discussion here is close to that in chapters 8 and 9 of Goldstein, Safko and Poole (see also sections 40 to 46 of Landau and Lifshitz).

3.1 The Hamiltonian

In Lagrangian mechanics the time evolution of a system is governed by a Lagrangian, $L(q_i, \dot{q}_i, t)$, which is a function of the generalised coordinates, the generalised velocities and time. The idea behind Hamiltonian mechanics is to formulate dynamics using generalised coordinates q_i and momenta p_i. We say that $(q_i, \dot{q}_i)$ are points in *configuration space* and (q_i, p_i) are points in *phase space*. The equations

$$p_j = \frac{\partial L(q_i, \dot{q}_i, t)}{\partial \dot{q}_j}, \tag{12}$$

express the momenta as functions of q_i, $\dot{q}_i$ and t. Suppose that these equations can be inverted to express the generalised velocities as functions of q_i, p_i and t. We shall see that as in (12) the velocities can be expressed as derivatives of a single function, the *Hamiltonian* $H(q_i, p_i, t)$, through

$$\dot{q}_j = \frac{\partial H(q_i, p_i, t)}{\partial p_j}.$$

Furthermore, the Hamiltonian governs the motion in phase space just as the Lagrangian governs motion in configuration space. Mathematically, the Hamiltonian is the *Legendre transform* of the Lagrangian *with respect to the velocities*. Essentially, the Legendre transform executes the required inversion of (12).

There now follows a brief explanation of the Legendre transform. Let us start with the simplest case of a function of one variable f. The equation

$$\lambda = f'(x),$$

has the 'solution'

$$x = g'(\lambda),$$

where g is the Legendre transform of f.

The Legendre transform for a function f of a single variable x is defined through

$$g(\lambda) = \sup_{x \in \text{dom}(f)} \left[\lambda x - f(x)\right].$$

To maximise the function $h(x, \lambda) = \lambda x - f(x)$ with respect to x, we must find when this quantity is stationary, that is $\partial h/\partial x = \lambda - f'(x) = 0$ so that

$$g(\lambda) = \lambda x - f(x)\big|_{\lambda = f'(x)}.$$

The function g exists if f is *convex*. Differentiating the above formula for g

$$g'(\lambda) = x + \lambda \frac{dx}{d\lambda} - f'(x) \frac{dx}{d\lambda}.$$

the second and third terms cancel when x is eliminated through $\lambda = f'(x)$. Applying a Legendre transform twice gives the original function.

A simple example is the function $f(x) = \frac{1}{2}cx^2$, where c is a positive constant. Now

$$h(x, \lambda) = \lambda x - f(x) = \lambda x - \frac{1}{2}cx^2,$$

which can be maximised with respect to x by setting $\lambda = f'(x) = cx$. Inserting $x = \lambda/c$ into h yields

$$g(\lambda) = h(\lambda/c, \lambda) = \frac{\lambda^2}{c} - \frac{1}{2}c\left(\frac{\lambda}{c}\right)^2 = \frac{\lambda^2}{2c}.$$

Another example is $f(x) = \sqrt{1 + x^2}$ (the graph is a branch of a hyperbola). Now

$$\lambda = f'(x) = \frac{x}{\sqrt{1 + x^2}},$$

which can be inverted to yield $x = \lambda/\sqrt{1 - \lambda^2}$. Inserting this into $h(x, \lambda) = \lambda x - f(x)$ yields $g(\lambda) = -\sqrt{1 - \lambda^2}$ so the graph of the Legendre transform is a semi-circle!

In N-dimensions, consider a function f with N variables x_i ($i = 1, ..., N$) then the Legendre transform is defined by

$$g(\lambda_i) = \sup_{(x_1, ..., x_N) \in \text{dom}(f)} \left[\sum_{j=1}^{N} \lambda_j x_j - f(x_i)\right].$$

Then, it follows that $x_i = \partial g/\partial \lambda_i$, where g is the Legendre transform of f.

Since $H(q_i, p_i, t)$ is the Legendre transform of the Lagrangian $L(q_i, \dot{q}_i, t)$ with respect to

$\dot{q}_i$, we can write

$$H(q_i, p_i, t) = \sum_{j=1}^{N} p_j \dot{q}_j - L(q_i, \dot{q}_i, t),$$

where the $\dot{q}_i$ are eliminated using $p_i = \partial L / \partial \dot{q}_i$.

A simple example is $L = \frac{1}{2} m \dot{x}^2 - V(x)$. Here $p = \partial L / \partial \dot{x} = m \dot{x}$ and so $\dot{x} = \frac{p}{m}$. From which, we can express the Hamiltonian as

$$H = p\dot{x} - L = \frac{p^2}{m} - \frac{1}{2} m \left(\frac{p}{m} \right)^2 + V(x) = \frac{p^2}{2m} + V(x),$$

which is the total energy. Further examples are considered on problem sheet 4.

3.2 Hamilton's Equations

In the Lagrangian framework with Lagrangian $L(q_i, \dot{q}_i, t)$ and momenta $p_i = \partial L / \partial \dot{q}_i$ the equations of motion are the Euler-Lagrange equations,

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) - \frac{\partial L}{\partial q_i} = 0 \quad \text{or} \quad \dot{p}_i = \frac{\partial L}{\partial q_i}.$$

It turns out that $\partial L / \partial q_i = -\partial H / \partial q_i$ (see problem sheet 4). This leads to Hamilton's equations

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \quad \text{and} \quad \dot{p}_i = -\frac{\partial H}{\partial q_i}, \tag{13}$$

for $i = 1, 2, \ldots, N$. This is a system of $2N$ first-order differential equations. It is possible to take these equations as a starting point without using a Lagrangian.

For example, consider the Hamiltonian

$$H = \frac{p^2}{2m} + V(x).$$

Hamilton's equations of motion for this Hamiltonian are given by

$$\dot{q} = \frac{\partial H}{\partial p} = \frac{p}{m}, \quad \dot{p} = -\frac{\partial H}{\partial x} = -\frac{dV}{dx}.$$

We can derive Hamilton's equations of motion from an action,

$$S = \int_{t_1}^{t_2} \left[\sum_{i=1}^{N} p_i \dot{q}_i - H \right] dt.$$

Note that if the momenta are eliminated via $p_i = \partial L / \partial \dot{q}_i$ the integrand coincides with the

Lagrangian L. Here, $(q_i(t), p_i(t))$ are arbitrary trajectories in phase space. Assume $q_i(t_1)$ and $q_i(t_2)$ are fixed. There is no need to impose this on the $p_i(t)$[8]. Consider a small variation of the momenta $p_i(t)$, i.e., make the replacement $p_i(t) \rightarrow p_i(t) + \delta p_i(t)$. Then leads to a change in the action

$$\delta S = \int_{t_1}^{t_2} \left[\sum_{i=1}^{N} \delta p_i \dot{q}_i - \delta H \right] dt = \int_{t_1}^{t_2} \left[\sum_{i=1}^{N} \delta p_i \dot{q}_i - \sum_{i=1}^{N} \frac{\partial H}{\partial p_i} \delta p_i \right] dt = \sum_{i=1}^{N} \int_{t_1}^{t_2} \delta p_i \left(\dot{q}_i - \frac{\partial H}{\partial p_i} \right) dt,$$

where δH denotes the change in H induced by the variations of the momenta. Now if we require $\delta S = 0$, we obtain Hamilton's equation, $\dot{q}_i = \partial H / \partial p_i$. Similarly, if we consider small variations of the generalised coordinates, i.e., make the replacement $q_i(t) \rightarrow q_i(t) + \delta q_i(t)$, then δS becomes

$$\delta S = \int_{t_1}^{t_2} \left[\sum_{i=1}^{N} p_i \delta \dot{q}_i - \delta H \right] dt = \sum_{i=1}^{N} \int_{t_1}^{t_2} \left(p_i \frac{d \delta q_i}{dt} - \frac{\partial H}{\partial q_i} \delta q_i \right) dt.$$

Using integration by parts

$$\delta S = \sum_{i=1}^{N} \left[p_i \delta q_i \Big|_{t_1}^{t_2} - \int_{t_1}^{t_2} \left(\dot{p}_i + \frac{\partial H}{\partial q_i} \right) \delta q_i \, dt \right] = - \sum_{i=1}^{N} \int_{t_1}^{t_2} \left(\dot{p}_i + \frac{\partial H}{\partial q_i} \right) \delta q_i \, dt,$$

since $\delta q_i(t_1) = \delta q_i(t_2) = 0$. Hence, imposing that $\delta S = 0$, we arrive at the second set of Hamilton's equations, namely $\dot{p}_i = -\partial H / \partial q_i$.

For example, consider the Kepler problem with Lagrangian

$$L = \frac{1}{2} m(\dot{r}^2 + r^2 \dot{\theta}^2) + \frac{k}{r}.$$

Here $p_r = m\dot{r}$ and $p_\theta = mr^2 \dot{\theta}$ so that $\dot{r} = p_r/m$ and $\dot{\theta} = p_\theta/(mr^2)$. Inserting into $p_r \dot{r} + p_\theta \dot{\theta} - L$ yields

$$H(r, \theta, p_r, p_\theta) = \frac{p_r^2}{2m} + \frac{p_\theta^2}{2mr^2} - \frac{k}{r}. \tag{14}$$

Hamilton's equations are

$$\dot{r} = \frac{\partial H}{\partial p_r} = \frac{p_r}{m}, \quad \dot{\theta} = \frac{\partial H}{\partial p_\theta} = \frac{p_\theta}{mr^2},$$

$$\dot{p}_r = -\frac{\partial H}{\partial r} = \frac{p_\theta^2}{mr^3} - \frac{k}{r^2}, \quad \dot{p}_\theta = -\frac{\partial H}{\partial \theta} = 0.$$

The final equation is just the result that θ is cyclic.

---

[8]In the discussion of canonical transformations in section 3.3 $\delta p_i(t_1) = \delta p_i(t_2) = 0$ will be imposed.

In the Lagrangian framework, the quantity

$$H = \sum_{i=1}^{N} p_i \dot{q}_i - L$$

is a constant of motion provided L has no explicit time-dependence, i.e., $\partial L/\partial t = 0$. Whereas, in the Hamiltonian framework, H is a constant of motion if $\partial H/\partial t = 0$. It is easy to prove that the Hamiltonian satisfies

$$\frac{dH}{dt} = \frac{\partial H}{\partial t}.$$

By the chain rule,

$$
\begin{aligned}
\frac{dH}{dt} &= \sum_{i=1}^{N} \left(\frac{\partial H}{\partial p_i} \dot{p}_i + \frac{\partial H}{\partial q_i} \dot{q}_i \right) + \frac{\partial H}{\partial t} \\
&= \sum_{i=1}^{N} \left[\frac{\partial H}{\partial p_i} \left(-\frac{\partial H}{\partial q_i} \right) + \frac{\partial H}{\partial q_i} \frac{\partial H}{\partial p_i} \right] + \frac{\partial H}{\partial t} = \frac{\partial H}{\partial t},
\end{aligned}
$$

using Hamilton's equations.

More generally, let $A = A(p_i, q_i, t)$ be a function on phase space. Then the chain rule yields the time derivative formula,

$$\frac{dA}{dt} = \sum_{i=1}^{N} \left(\frac{\partial A}{\partial p_i} \dot{p}_i + \frac{\partial A}{\partial q_i} \dot{q}_i \right) + \frac{\partial A}{\partial t} = \sum_{i=1}^{N} \left(-\frac{\partial A}{\partial p_i} \frac{\partial H}{\partial q_i} + \frac{\partial A}{\partial q_i} \frac{\partial H}{\partial p_i} \right) + \frac{\partial A}{\partial t} = \{A, H\} + \frac{\partial A}{\partial t}, \tag{15}$$

where we have defined the *Poisson bracket*,

$$\{A, B\} = \sum_{i=1}^{N} \left(\frac{\partial A}{\partial q_i} \frac{\partial B}{\partial p_i} - \frac{\partial A}{\partial p_i} \frac{\partial B}{\partial q_i} \right). \tag{16}$$

3.3 Poisson Brackets

The Poisson bracket has the following properties:

(i) Anti-symmetry:

$$\{A, B\} = -\{B, A\}.$$

In particular, $\{A, A\} = 0$ and so

$$\frac{dH}{dt} = \{H, H\} + \frac{\partial H}{\partial t} = \frac{\partial H}{\partial t}.$$

(ii) Leibniz formula:

$$\{A, BC\} = \{A, B\}C + B\{A, C\}.$$

(iii) Jacobi identity:

$$\{A, \{B, C\}\} + \{B, \{C, A\}\} + \{C, \{A, B\}\} = 0.$$

In terms of Poisson brackets, we can rewrite Hamilton's equations as

$$\dot{q}_i = \frac{\partial H}{\partial p_i} = \{q_i, H\}, \qquad \dot{p}_i = -\frac{\partial H}{\partial q_i} = \{p_i, H\}. \tag{17}$$

The *fundamental Poisson brackets*

$$\{q_i, q_j\} = \{p_i, p_j\} = 0 \text{ and } \{q_i, p_j\} = \delta_{ij}. \tag{18}$$

follow from the definition (16).

For an example of Poisson bracket calculations consider the Hamiltonian for a particle moving in the plane subject to a conservative force

$$H(r, \theta, p_r, p_\theta, t) = \frac{p_r^2}{2m} + \frac{p_\theta^2}{2mr^2} + V(r, \theta),$$

where r and θ are polar coordinates. One of Hamilton's equations is

$$\dot{p}_\theta = \{p_\theta, H\} = \left\{p_\theta, \frac{p_r^2}{2m}\right\} + \left\{p_\theta, \frac{p_\theta^2}{2mr^2}\right\} + \{p_\theta, V(r, \theta)\} = -\{V, p_\theta\}.$$

A short calculation gives $\{V, p_\theta\} = \partial V/\partial\theta$.

3.4   Canonical Transformations

The Lagrangian formulation of mechanics is invariant under point transformation, i.e., under the change of coordinates

$$q_i' = q_i'(q_1, q_2, ..., q_N, t),$$

the Euler-Lagrange equations become

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

where $L' = L'(q_i', \dot{q}_i', t)$ is the original Lagrangian expressed in the new coordinates and velocities.

.Now consider the transformation of coordinates and momenta,

$$Q_i = Q_i(q_1, ..., q_N, p_1, ..., p_N, t) \text{ and } P_i = P_i(q_1, ..., q_N, p_1, ..., p_N, t),$$

for $i = 1, 2, \ldots, N$, where the new equations of motion preserve the Hamiltonian form, i.e.,

$$\dot{Q}_i = \frac{\partial K}{\partial P_i} \text{ and } \dot{P}_i = -\frac{\partial K}{\partial Q_i}, \tag{19}$$

with the new Hamiltonian denoted by $K = K(Q_i, P_i, t)$. Note that in Lagrangian mechanics, $L' = L$, but it need not be the case that $K = H$. We shall see that *Canonical Transformations* preserve the Hamiltonian form of the equations of motion. Canonical transformations can be defined without reference to the Hamiltonian:

A canonical transformation is a change of variables, $Q_i = Q_i(q_j, p_j, t)$ and $P_i = P_i(q_j, p_j, t)$, which preserves fundamental Poisson brackets, i.e.,

$$\{Q_i, Q_j\} = \{P_i, P_j\} = 0 \text{ and } \{Q_i, P_j\} = \delta_{ij}.$$

Note that point transformations are canonical transformations. A simple example of a canonical transformation that is not a point transformation is

$$Q = \frac{q}{p} \text{ and } P = \frac{p^2}{2}.$$

To see this, we evaluate the fundamental Poisson brackets:

$$\{Q, P\} = \frac{\partial Q}{\partial q}\frac{\partial P}{\partial p} - \frac{\partial Q}{\partial p}\frac{\partial P}{\partial q} = \frac{1}{p} \cdot p - \left(-\frac{q}{p^2}\right) \cdot 0 = 1.$$

The change of variables

$$Q = \frac{q}{p} \text{ and } P = \frac{p^2}{2} + q$$

is not a canonical transformation since

$$\{Q, P\} = \frac{1}{p} \cdot p - \left(-\frac{q}{p^2}\right) \cdot 1 = 1 + \frac{q}{p^2} \neq 1.$$

We can also define canonical transformations via the action

$$S = \int_{t_1}^{t_2} \left[\sum_{i=1}^{N} p_i \dot{q}_i - H\right] dt. \tag{20}$$

It is clear that the Hamiltonian form of the equations of motion are preserved if

$$\left(\sum_{i=1}^{N} p_i \frac{dq_i}{dt} - H \right) dt = \sum_{i=1}^{N} p_i dq_i - H dt = \sum_{i=1}^{N} P_i dQ_i - K dt,$$

since imposing $\delta S = 0$ yields

$$\dot{Q}_i = \frac{\partial K}{\partial P_i} \text{ and } \dot{P}_i = -\frac{\partial K}{\partial Q_i}.$$

However, this differential condition is very restrictive. Consider the weaker condition

$$\sum_{i=1}^{N} p_i dq_i - H dt = \sum_{i=1}^{N} P_i dQ_i - K dt + dF, \tag{21}$$

where $F = F(q_i, p_i, t)$. Note that the dF term does not affect the equations of motion since

$$\int_{t_1}^{t_2} \frac{dF}{dt} dt = F(q_i(t_2), p_i(t_2), t_2) - F(q_i(t_1), p_i(t_1), t_1).$$

Therefore F drops out of the calculation of δS provided $p_i(t_1)$ and $p_i(t_2)$ are fixed.

We now outline the proof that the differential condition

$$\sum_{i=1}^{N} p_i dq_i - H dt = \sum_{i=1}^{N} P_i dQ_i - K dt + dF$$

yields a canonical transformation. First we shall prove this result for $N = 1$. In this case, the differential condition is

$$p dq - H dt = P dQ - K dt + dF.$$

Applying the chain rule, the differentials are

$$dQ = \frac{\partial Q}{\partial q} dq + \frac{\partial Q}{\partial p} dp + \frac{\partial Q}{\partial t} dt, \qquad dF = \frac{\partial F}{\partial q} dq + \frac{\partial F}{\partial p} dp + \frac{\partial F}{\partial t} dt,$$

and so we have

$$p dq - H dt = \left(P \frac{\partial Q}{\partial q} + \frac{\partial F}{\partial q} \right) dq + \left(P \frac{\partial Q}{\partial p} + \frac{\partial F}{\partial p} \right) dp + \left(P \frac{\partial Q}{\partial t} + \frac{\partial F}{\partial t} - K \right) dt.$$

Equating coefficients

$$p = P \frac{\partial Q}{\partial q} + \frac{\partial F}{\partial q}, \quad 0 = P \frac{\partial Q}{\partial p} + \frac{\partial F}{\partial p}, \quad H = K - P \frac{\partial Q}{\partial t} - \frac{\partial F}{\partial t}.$$

If we differentiate the first of these equations with respect to p and the second with respect to q, then subtract them from one another (integrability condition), we obtain

$$1 = \frac{\partial P}{\partial p}\frac{\partial Q}{\partial q} - \frac{\partial P}{\partial q}\frac{\partial Q}{\partial p} = \{Q, P\}.$$

Hence, since $\{Q, Q\} = \{P, P\} = 0$ automatically, the differential condition results in a canonical transformation. Now, for $N > 1$, the same argument to equate coefficients can be used to obtain

$$p_i = \sum_{j=1}^{N} p_j \frac{\partial Q_j}{\partial q_i} + \frac{\partial F}{\partial q_i} \ \text{ and } \ 0 = \sum_{j=1}^{N} p_j \frac{\partial Q_j}{\partial p_i} + \frac{\partial F}{\partial p_i}.$$

Now applying the integrability condition to this system of equations, the following results are obtained:

$$(q_i, q_j)_{Q,P} = (p_i, p_j)_{Q,P} = 0 \ \text{ and } \ (q_i, p_j)_{Q,P} = \delta_{ij},$$

where we define the Lagrange bracket[9] by

$$(u, v)_{q,p} = \sum_{i=1}^{N} \left(\frac{\partial q_i}{\partial u}\frac{\partial p_i}{\partial v} - \frac{\partial q_i}{\partial v}\frac{\partial p_i}{\partial u} \right).$$

For $N = 1$,

$$(q, p)_{Q,P} = \{Q, P\}.$$

It can be shown that the Lagrange bracket conditions are equivalent to the fundamental Poisson bracket results (see section 9.5 of Goldstein, Safko and Poole).

In the above discussion F was assumed to be a function of the 'old' variables q_i, p_i and t. It can also be written in terms of the 'new' variables Q_i, P_i and t. However, it is also useful to write it in a mixed form. For example, consider $F = F(q_i, Q_i, t)$ as a function of the old and new coordinates. Then for $N = 1$, $F = F(q, Q, t)$ and

$$dF = \frac{\partial F}{\partial q}dq + \frac{\partial F}{\partial Q}dQ + \frac{\partial F}{\partial t}dt,$$

and so equating coefficients gives us the equations

$$p = \frac{\partial F}{\partial q}, \ \ 0 = P + \frac{\partial F}{\partial Q}, \ \ H = K - \frac{\partial F}{\partial t}.$$

Recall that $Q = q/p$, $P = p^2/2$ is a canonical transformation. The associated generating

---

[9]Lagrange brackets are not examinable.

function can be obtained by writing p and P as a function of q and Q

$$p = \frac{q}{Q} = \frac{\partial F}{\partial q}, \quad P = \frac{q^2}{2Q^2} = -\frac{\partial F}{\partial Q}.$$

These equations can be integrated to give the generating function $F = q^2/(2Q)$.

We can define a canonical transformation through the differential condition

$$\sum_{i=1}^{N} p_i dq_i - H dt = \sum_{i=1}^{N} P_i dQ_i - K dt + dF, \tag{22}$$

with $F = F(q_i, Q_i, t)$. It follows from equating coefficients of dq_i, dQ_i and dt that

$$p_i = \frac{\partial F}{\partial q_i}, \quad P_i = -\frac{\partial F}{\partial Q_i} \quad \text{and} \quad K = H + \frac{\partial F}{\partial t},$$

where F is the generating function of the canonical transformation.

It is not possible to derive every canonical transformation from a generating function $F(q_i, Q_i, t)$. For example, consider the trivial (identity) transformation $(Q, P) = (q, p)$. In this case, we cannot write

$$p = \frac{\partial F}{\partial q} \quad \text{and} \quad P = -\frac{\partial F}{\partial Q}.$$

We may consider alternative mixed forms. There are four possibilities[10] for this, namely

$$F_1(q_i, Q_i), \quad F_2(q_i, P_i), \quad F_3(p_i, Q_i), \quad F_4(p_i, P_i).$$

We can rewrite the differential condition using the product rule

$$d\left(\sum_{i=1}^{N} Q_i P_i\right) = \sum_{i=1}^{N} P_i dQ_i + \sum_{i=1}^{N} Q_i dP_i,$$

so that

$$\sum_{i=1}^{N} p_i dq_i - H dt = -\sum_{i=1}^{N} Q_i dP_i - K dt + dF_2,$$

where

$$F_2 = F + \sum_{i=1}^{N} Q_i P_i,$$

---

[10]Excluding generating functions such as $F(q_1, p_2, Q_1, Q_2, t)$.

then equating coefficients yields the equations

$$p_i = \frac{\partial F_2}{\partial q_i}, \quad Q_i = \frac{\partial F_2}{\partial P_i}, \quad \text{and} \quad K = H + \frac{\partial F_2}{\partial t}.$$

There are similar formulas for type 3 and type 4 generating functions.

Returning to the canonical transformation $(Q, P) = (q/p, p^2/2)$, the type 2 generating function is $F_2 = q\sqrt{2P}$. We saw that the identity transformation $(Q, P) = (q, p)$. does not have a type 1 generating function. It is generated by the type 2 function $F_2(q, P) = qP$.

Remark: The Poisson bracket is defined as

$$\{A, B\} = \{A, B\}_{q_i, p_i} = \sum_{i=1}^{N} \left(\frac{\partial A}{\partial q_i} \frac{\partial B}{\partial p_i} - \frac{\partial A}{\partial p_i} \frac{\partial B}{\partial q_i} \right),$$

but we can use any set of canonical variables to compute Poisson brackets, i.e., $\{A, B\}_{q_i, p_i} = \{A, B\}_{Q_i, P_i}$.

3.5   Time Evolution as a Canonical Transformation

Shifting time does not affect the Hamiltonian form of the equations of motion, i.e.,

$$Q_i = q_i(t) = Q_i(q_j, p_j, t), \quad P_i = p_i(t) = P_i(q_j, p_j, t),$$

where q_i and p_i are generalised coordinates and momenta at $t = 0$ (or some other initial time). The new variables defined through shifting time define a canonical transformation.

For small t we have

$$Q_i = q_i + t\{q_i, H\} + \mathcal{O}(t^2), \quad P_i = q_i + t\{p_i, H\} + \mathcal{O}(t^2),$$

using $\dot{q}_i = \{q_i, H\}$ and $\dot{p}_i = \{p_i, H\}$. This set of variables defines a canonical transformation since

$$\{Q_i, P_j\} = \{q_i, p_j\} + t \left[\{q_i, \{p_j, H\}\} + \{\{q_i, H\}, p_j\} \right] + \mathcal{O}(t^2)$$
$$= \delta_{ij} - t\{H, \{q_i, p_j\}\} + \mathcal{O}(t^2)$$
$$= \delta_{ij} - t\{H, \delta_{ij}\} + \mathcal{O}(t^2)$$
$$= \delta_{ij} + \mathcal{O}(t^2),$$

as required, using the Jacobi identity. Similarly, up to terms of order t^2, we have $\{Q_i, Q_j\} = \{P_i, P_j\} = 0$, and so this new set of variables defines a canonical transformation.

Liouville's Theorem states that the Jacobian of a canonical transformation is unity, i.e., canonical transformations preserve volumes in phase space.

For $N = 1$,

$$dQ \, dP = \det \begin{bmatrix} \partial Q/\partial q & \partial Q/\partial p \\ \partial P/\partial q & \partial P/\partial p \end{bmatrix} dq \, dp = \{Q, P\} \, dq \, dp = dq \, dp,$$

for a canonical transformation. The $N > 1$ case is considered in the next section. As time evolution is a canonical transformation, we can conclude that time evolution preserves phase space volume.

Consider simple harmonic motion where the Hamiltonian is

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2,$$

with mass m and angular frequency ω. Further assume that $m = \omega = 1$, i.e.,

$$H = \frac{p^2}{2} + \frac{x^2}{2}.$$

Hamilton's equations become

$$\dot{x} = \frac{\partial H}{\partial p} = p \text{ and } \dot{p} = -\frac{\partial H}{\partial x} = -x,$$

Here trajectories in phase space are circles with the same frequency for all orbits. Time evolution amounts to a continuous rotation in phase space.

3.6   Appendix: Proof of Liouville's Theorem

Liouville's theorem is essentially obvious for $N = 1$. For $N > 1$ it is not. The argument here is an expanded version of the one given by Landau and Lifshitz. In the case that $N > 1$, we need to show that the determinant of a $2N \times 2N$ matrix M is unity. To do this, split M into four $N \times N$ block matrices, i.e.,

$$M = \begin{bmatrix} \underbrace{\partial Q_i/\partial q_j}_{N \times N} & \partial Q_i/\partial p_j \\ \partial P_i/\partial q_j & \partial P_i/\partial p_j \end{bmatrix},$$

for $i, j = 1, 2, \ldots, N$. Then apply a canonical transformation through two non-canonical transformations, i.e., $(q_i, p_i) \to (q_i, P_i) \to (Q_i, P_i)$. Under the first change of variables, we get the matrix M_1:

$$M_1 = \begin{bmatrix} \partial q_i/\partial q_j & \partial q_i/\partial p_j \\ \partial P_i/\partial q_j & \partial P_i/\partial p_j \end{bmatrix} = \begin{bmatrix} \mathbf{1}_N & \mathbf{0}_N \\ \partial P_i/\partial q_j & \partial P_i/\partial p_j \end{bmatrix},$$

where $\mathbf{0}_N$ and $\mathbf{1}_N$ are $N \times N$ zero and identity matrices, respectively. Follow a similar

procedure to find the matrix M_2. Then,

$$\det M = \det(M_1 M_2) = \det M_1 \det M_2 = \det \frac{\partial P_i}{\partial p_j} \det \frac{\partial Q_i}{\partial q_j}.$$

Now assume the existence of a type 2 generating function $F = F_2(q_i, P_i)$. Recall this means we have

$$p_i = \frac{\partial F}{\partial q_i} \text{ and } Q_i = \frac{\partial F}{\partial P_i},$$

and so

$$\frac{\partial p_i}{\partial P_j} = \frac{\partial^2 F}{\partial P_j \partial q_i} \text{ and } \frac{\partial Q_j}{\partial q_i} = \frac{\partial^2 F}{\partial q_i \partial P_j},$$

which tells us that

$$\det M_1 = \frac{1}{\det M_2}.$$

4 Integrability and Symmetries

This chapter uses Hamiltonian methods to discuss more advanced topics including Hamilton-Jacobi theory, action angle variables, integrability and Noether's theorem. Sections 4.1 and 4.2 closely follow chapter 10 of Goldstein, Safko and Poole (see also sections 47 to 50 of Landau and Lifshitz).

4.1 The Hamilton Jacobi Equation

The phase space coordinates of a system can be written as a function of time t and the phase space coordinates at $t = 0$ (or some other reference time), i.e.,

$$q_i(t) = q_i(Q_j, P_j, t), \quad p_i(t) = p_i(Q_j, P_j, t)$$

where $Q_j = q_j(t = 0)$, $P_j = p_j(t = 0)$. This canonical transformation is the inverse of the transformation considered in section 3.5. As the new variable are, by construction, constants of the motion, the new Hamiltonian is zero, i.e.

$$\dot{Q}_i = \frac{\partial K}{\partial P_i}, \quad \dot{P}_i = -\frac{\partial K}{\partial Q_i},$$

with $K(Q_i, P_i, t) = 0$.

We would now like to construct a canonical transformation so that the new Hamiltonian is zero. To do this, write this transformation through a type 2 generating function $F_2 = F_2(q_i, P_i, t)$, i.e.,

$$p_i = \frac{\partial F_2}{\partial q_i}, \quad Q_i = \frac{\partial F_2}{\partial P_i}, \quad K = H + \frac{\partial F_2}{\partial t}.$$

Setting $K = 0$ gives

$$H + \frac{\partial F_2}{\partial t} = 0,$$

which can be written in the form

$$H\left(q_i, \frac{\partial F_2}{\partial q_i}, t\right) + \frac{\partial F_2}{\partial t} = 0,$$

This generating function is usually denoted S (Hamilton's Principal Function) instead of F_2

$$H\left(q_i, \frac{\partial S}{\partial q_i}, t\right) + \frac{\partial S}{\partial t} = 0. \tag{23}$$

This is the Hamilton Jacobi equation - a first order PDE. Solutions of the Hamilton-Jacobi

equation have the form $S = S(q_i, \alpha_1, \alpha_2, ..., \alpha_{N+1}, t)$ where α_1, α_2,...,α_{N+1} are arbitrary constants ($N + 1$ since the PDE has $N + 1$ derivatives). One of the constants, say α_{N+1}, is additive (adding a constant to S does not change the Hamilton-Jacobi equation). Consider $S(q_i, \alpha_i, t)$, $i = 1, 2, ..., N$ where the N constants are non-additive. A solution of the Hamilton Jacobi equation $S = S(q_i, \alpha_i, t)$ including N non-additive constants is called a *complete solution*. Identifying these constants as new momenta P_i. As $K = 0$ the new coordinates

$$Q_i = \frac{\partial S}{\partial P_i} = \frac{\partial S}{\partial \alpha_i},$$

are constants of the motion. Call the new coordinates β_i

$$\beta_i = \frac{\partial S}{\partial \alpha_i}. \tag{24}$$

Instead of directly solving Hamilton's equations ($2N$ coupled ODEs) solve the Hamilton-Jacobi equation (a first order PDE in $N + 1$ variables). Given a complete solution $S(q_i, \alpha_i, t)$ one can, in principle, invert $\beta_i = \partial S/\partial \alpha_i$ to obtain $q_i = q_i(\alpha_j, \beta_j, t)$ thereby solving the equations of motion.

How can one solve the PDE (23)? The method of characteristics reduces the problem of solving such a PDE to the solution of a system of ODEs. This system of ODEs is Hamilton's equations. Solving the Hamilton-Jacobi equation through the method of characteristics is not an alternative to directly solving Hamilton's equations. However, a form of separation of variables can be used to integrate the Hamilton-Jacobi equation. Write S as a *sum* of functions of fewer than $N + 1$ variables. For example, consider the Hamiltonian for a simple harmonic oscillator

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2.$$

It is easy to see that this leads to $\ddot{x} = -\omega^2 x$. Setting $m = 1$ and $\omega = 1$ gives

$$H = \frac{1}{2}(p^2 + x^2),$$

and the Hamilton-Jacobi equation is

$$\frac{1}{2}\left[\left(\frac{\partial S}{\partial x}\right)^2 + x^2\right] + \frac{\partial S}{\partial t} = 0.$$

Consider a solution of the form

$$S = W(x) - f(t).$$

The Hamilton-Jacobi equation is then

$$\frac{1}{2}\left[W'(x)^2 + x^2\right] - f'(t) = 0,$$

or

$$\frac{1}{2}\left[W'(x)^2 + x^2\right] = f'(t) = \alpha = \text{ constant,}$$

This is constant since the first term is independent of t and the second term is independent of x. Accordingly, $f(t) = \alpha t$, dropping an additive constant. $W(x)$ satisfies the ODE

$$W'(x) = \pm\sqrt{2\alpha - x^2},$$

so that

$$W(x) = \pm\int \sqrt{2\alpha - x^2}\, dx = \pm\left[\frac{1}{2}x\sqrt{2\alpha - x^2} + \alpha\sin^{-1}\frac{x}{\sqrt{2\alpha}}\right],$$

again ignoring an additive constant. Now $S(x, \alpha, t) = W(x) - \alpha t$. The new coordinate is

$$\beta = \frac{\partial S}{\partial\alpha} = \frac{\partial W}{\partial\alpha} - t = \pm\sin^{-1}\frac{x}{\sqrt{2\alpha}} - t.$$

Note that it is simpler to compute $\partial W/\partial\alpha$ by differentiating the integral

$$\frac{\partial}{\partial\alpha}\int \sqrt{2\alpha - x^2}\, dx = \int \frac{\mathrm{d}x}{\sqrt{2\alpha - x^2}} = \sin^{-1}\frac{x}{\sqrt{2\alpha}} + \text{constant,}$$

rather than (tediously) differentiating the explicit formula for W. Therefore

$$\pm\sin^{-1}\frac{x}{\sqrt{2\alpha}} = \beta + t,$$

or

$$x(t) = \sqrt{2\alpha}\sin(t + \beta),$$

which is indeed the general solution for the simple harmonic oscillator. Here $\sqrt{2\alpha}$ is the amplitude and β is a phase. The 'new momentum' α is the energy.

The solution of the simple harmonic oscillator via the Hamilton-Jacobi method is rather clumsy. However, in certain more complicated cases it reduces the problem to a series of integrations. The separation of the time coordinate works for any Hamiltonian, $H(q_i, p_i)$, which has no explicit time-dependence. In this case

$$S(q_i, \alpha_i, t) = W(q_i, \alpha_i) - \alpha_1 t.$$

Here the Hamilton-Jacobi equation can be written in the form

$$H\left(q_i, \frac{\partial W}{\partial q_i}\right) = \alpha_1.$$

As in the previous example, the constant α_1 can be identified as the energy. W is known as *Hamilton's Characteristic Function.* Hamilton's Principal Function, S, generates a canonical

transformation such that the new Hamiltonian is zero. Hamilton's Characteristic Function (HCF) generates a canonical transformation so that the new Hamiltonian is α_1 and the new coordinates, $Q_i = \partial W / \partial \alpha_i$ are cyclic.[11] The canonical transformation generated by HCF is time-independent.

Assume that $N > 1$. A coordinate, say q_1, is called separable if Hamilton's Principal Function (HPF) can be written in the form

$$S(q_i, \alpha_i, t) = S_1(q_1, \alpha_i, t) + S'(q_2, q_3, ..., q_N, \alpha_i, t).$$

Any cyclic coordinate is separable. If, say q_2, is cyclic then HPF can be written as

$$S = \alpha_2 q_2 + S'(q_1, q_3, ..., q_N, \alpha_1, \alpha_2, ..., \alpha_N, t),$$

and the constant α_2 is the conserved momentum $p_2 = \partial S / \partial q_2$. The converse is not true as a separable coordinate is not necessarily cyclic. The Hamilton-Jacobi equation is *completely separable* if all the coordinates are separable, that is

$$S = \sum_{j=1}^{N} S_j(q_j, \alpha_1, \alpha_2, ..., \alpha_N, t).$$

For example, motion in a central force is described by the Hamiltonian

$$H = \frac{p_r^2}{2m} + \frac{p_\theta^2}{2mr^2} + V(r).$$

θ is cyclic and hence separable. The Hamilton-Jacobi equation

$$\frac{1}{2m} \left[\left(\frac{\partial S}{\partial r} \right)^2 + \frac{1}{r^2} \left(\frac{\partial S}{\partial \theta} \right)^2 \right] + V(r) + \frac{\partial S}{\partial t} = 0,$$

is completely separable. H has no explicit t dependence ('t is cyclic') so that $S = W - \alpha_1 t$ with

$$W = W_r(r) + \alpha_\theta \theta.$$

The Hamilton-Jacobi equation for W_r is

$$\frac{1}{2m} W'(r)^2 + \frac{\alpha_\theta^2}{2mr^2} + V(r) = \alpha_1,$$

---

[11] $\beta_i = \partial S / \partial \alpha_i = \partial(W - \alpha_1 t)/\partial \alpha_i$. Q_2, Q_3,...,Q_N are the same as β_2, β_3,..., β_N. However, $\beta_1 = Q_1 - t$ so that $Q_1 = \beta_1 + t$ is not a constant of the motion.

which can be integrated to give

$$W_r(r) = \pm \int \sqrt{2m\big[\alpha_1 - V(r)\big] - \frac{\alpha_\theta^2}{r^2}}\, dr.$$

In principle this provides HPF and solves the equations of motion. The new coordinates are

$$\beta_\theta = \frac{\partial S}{\partial \alpha_\theta} = \frac{\partial}{\partial \alpha_\theta}(W_r + \alpha_\theta \theta - \alpha_1 t) = \frac{\partial W_r}{\partial \alpha_\theta} + \theta$$

$$= \mp \int \frac{\alpha_\theta\, dr}{r^2 \sqrt{2m\big[\alpha_1 - V(r)\big] - \alpha_\theta^2/r^2}} + \theta.$$

$$\beta_1 = \frac{\partial S}{\partial \alpha_1} = \frac{\partial}{\partial \alpha_1}(W_r + \alpha_\theta \theta - \alpha_1 t)$$

$$= \pm \int \frac{m\, dr}{\sqrt{2m\big[\alpha_1 - V(r)\big] - \alpha_\theta^2/r^2}} - t.$$

The same symbol, S, is used for Hamilton's Principal Function and the action S. Consider the total derivative of HPF

$$\frac{dS}{dt} = \sum_{i=1}^{N} \frac{\partial S}{\partial q_i}\frac{dq_i}{dt} + \frac{\partial S}{\partial t} = \sum_{i=1}^{N} p_i \dot{q}_i - H = L.$$

In chapter 2 the action was defined as the integral of L for trajectories with fixed end points. The Hamilton Jacobi is a PDE satisfied by the extremal action as a function of the coordinates, q_i, and time, t, of the end point. The connection between the action and Hamilton-Jacobi equation is discussed by Goldstein, Safko and Poole (section 10.1).

4.2   Action Angle Variables

In the Hamilton-Jacobi approach the (non-additive) integration constants α_i are interpreted as 'new' momenta. There is some arbitrariness in the choice of these constants (they could be taken to be the momenta at a fixed reference time or identified with the energy or conserved momenta). If the following conditions hold there is a natural choice for these momenta known as *action variables*:

  (i) H has no explicit time dependence,
  (ii) The Hamilton-Jacobi equation is completely separable,
  (iii) The dynamics is oscillatory (rotation or libration).

Libration is where the motion is periodic in a coordinate, say q_1, and its conjugate momentum p_1. An example is the simple harmonic oscillator with Hamiltonian $H = \frac{1}{2}(p^2 + x^2)$
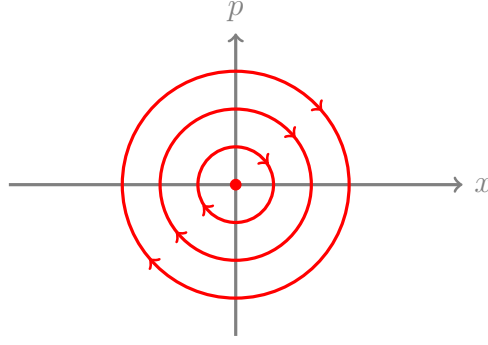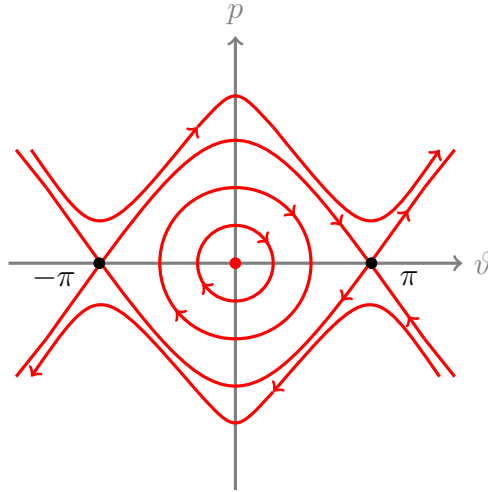
Figure 1: Phase diagram for the simple harmonic oscillator.

The simple pendulum exhibits both libration and rotation. Here the Hamiltonian is

$$H = \frac{p^2}{2} - \cos\theta,$$

where the mass, length and the acceleration due to gravity are set to unity.

For $-1 < E < 1$, where the trajectories resemble those of the harmonic oscillator, there is libration. For $E > 1$ the momentum p is periodic but the θ coordinate is not (here θ shifts by $\pm 2\pi$ during each rotation).

Action variables are defined through

$$J_i = \oint p_i \, dq_i, \quad i = 1, ..., N \quad \text{(no sum)} \tag{25}$$

where the integral is over a complete libration or rotation. Given assumption (i) the momenta can be written $p_i = \partial W / \partial q_i$ where W is Hamilton's Characteristic Function. From assumption (ii) we have

$$W = \sum_{i=1}^{N} W_i(q_i, \alpha_1, \alpha_2, ..., \alpha_N),$$

so that

$$p_i = \frac{\partial W_i(q_i, \alpha_1, \alpha_2, ..., \alpha_N)}{\partial q_i},$$

and

$$J_i = \oint \frac{\partial W_i(q_i, \alpha_1, \alpha_2, ..., \alpha_N)}{\partial q_i} \, dq_i,$$

and so J_i is a constant as it is function of the constant momenta α_1, α_2,...,α_N.

Here we assumed that the α_i are constants of the motion which are identified as new momenta. We can replace these (unspecified) momenta with the action angle variables J_i. That is we can take the J_i to be new momenta and write HCF in the form

$$W = \sum_{i=1}^{N} W_i(q_i, J_1, J_2, ..., J_N).$$

This generates a canonical transformation for which the new coordinates

$$w_i = \frac{\partial W}{\partial J_i}$$

are cyclic. These new coordinates are called angle variables. (w_i, J_i) are the *action-angle variables*.

Consider how this works for the simple harmonic oscillator with Hamiltonian $H = \frac{1}{2}(p^2 + x^2)$ and from the previous section $W = \int \sqrt{2\alpha - x^2} \, dx$, where α was taken to be the energy $\frac{1}{2}(p^2 + x^2)$. Here

$$J = \oint p \, dx = 2\pi\alpha,$$

which can be obtained by explicitly computing the integral

$$J = 4 \int_0^{\sqrt{2\alpha}} \sqrt{2\alpha - x^2} \, dx,$$

(the factor of 4 as the integral represents one quarter of the trajectory) or just by noting that J is the area enclosed by the circular trajectory. Accordingly, HCF can be written as

$$W = \int \sqrt{\frac{J}{\pi} - x^2} \, dx.$$

The new coordinate, w, is obtained through

$$w = \frac{\partial W}{\partial J} = \frac{1}{2\pi} \int \frac{dx}{\sqrt{J/\pi - x^2}} = \frac{1}{2\pi} \sin^{-1} \frac{x}{\sqrt{2\alpha}} + \text{constant}$$

$$= \frac{1}{2\pi} \sin^{-1} \frac{x}{\sqrt{x^2 + p^2}} + \text{constant} = -\frac{\theta}{2\pi} + \text{different constant}.$$

That is up to the factor of -2π the new coordinate, w, is the angle θ [12] Note that the periodicity in the new coordinate is unity (increasing w by one corresponds to one libration in the original coordinates).

In general, the periodicity of all angle variables is one.

To derive this integrate $\dot{w}_i$ over a full period

$$\int_{t_1}^{t_2} \dot{w}_i \, dt = \int_{t_1}^{t_2} \frac{d}{dt} \left[\frac{\partial W_i(q_i, J_1, ..., J_N)}{\partial J_i} \right] dt = \int_{t_1}^{t_2} \frac{\partial^2 W_i}{\partial q_i \partial J_i} \frac{dq_i}{dt} \, dt$$

$$= \int_{q_i(t_1)}^{q_i(t_2)} \frac{\partial^2 W_i}{\partial J_i \partial q_i} \, dq_i = \frac{\partial}{\partial J_i} \int_{q_i(t_1)}^{q_i(t_2)} \frac{\partial W_i}{\partial q_i} \, dq_i = \frac{\partial}{\partial J_i} J_i = 1.$$

Therefore w_i increases by 1 during a full libration or rotation.

Hamilton's Characteristic function $W(q_i, J_i)$ generates a canonical transformation for which the new coordinates, w_i, are cyclic. That is the new Hamiltonian, K, does not depend on the angles w_i. Using Hamilton's equations

$$\dot{w}_i = \frac{\partial K}{\partial J_i} = \nu_i = \text{constant}.$$

Suppose that w_i executes one libration (or rotation) in time T_i. Integrating from $t = 0$ to $t = T_i$ yields $1 = \nu_i T_i$ so that $\nu_i = 1/T_i$. That is the constant $\nu_i = \partial K / \partial J_i$ is the frequency.

In other words, to compute frequencies, express the Hamiltonian as a function of the action variables, J_i, and differentiate.

An example is the simple pendulum (see problem sheet 7). A further example is the Kepler problem

$$H = \frac{p_r^2}{2} + \frac{p_\theta^2}{2r^2} - \frac{k}{r}.$$

As θ is cyclic, HCF has the form $W = W_r + \alpha_\theta \theta$ so that $p_\theta = \partial W / \partial \theta = \alpha_\theta$. Therefore

$$J_\theta = \int_0^{2\pi} p_\theta \, d\theta = 2\pi \alpha_\theta.$$

Now compute

$$J_r = 2 \int_{r_1}^{r_2} \sqrt{2\alpha_1 - \frac{\alpha_\theta^2}{r^2} + \frac{2k}{r}} \, dr,$$

as a function of $J_\theta = 2\pi \alpha_\theta$ and $\alpha_1 = K$. Here r_1 and r_2 are the minimum and maximum values of r during the libration in r. The factor of 2 reflects that the libration comprises two parts (where r is increasing and decreasing). Then invert to write K as a function of J_r

---

[12]Landau and Lifshitz define action variables through $2\pi I = \oint p \, dx$. With this convention angle variable is $-\theta$. We will stick with $J = \oint p \, dx$ as in Goldstein, Safko and Poole.

and J_θ. One can show that $\nu_r = \nu_\theta$. This is just the result the the Kepler orbits are closed $(\alpha_1 < 0)$.

4.3 Adiabatic Invariance

A further key property of the action variables is that they are *adiabatic invariants.* Although the derivation of this property is rather subtle the idea is straightforward to apply. In the discussion below $N = 1$.

Consider a Hamiltonian of the form $H(q, p, \lambda)$ where λ is a parameter. If action-angle variables (w, J) can be introduced the Hamiltonian or energy may be expressed as a function of J and λ. Alternatively, the action variable, J, can be expressed as a function of energy, E, and the parameter λ, that is $J = J(E, \lambda)$. If the parameter λ is a function of time so that the Hamiltonian has the form $H(q, p, \lambda(t))$ then E and $J(E, \lambda)$ are no longer constants of the motion. However, it turns out that $J(E(t), \lambda(t))$ is 'almost' a constant of the motion provided that λ changes 'slowly'.

A classic example is the lengthening pendulum problem. The Hamiltonian is the same as for a simple pendulum except that the length, ℓ, is no longer assumed to be constant. Taking the pendulum bob to have unit mass the Hamiltonian is

$$H = \frac{p^2}{2\lambda(t)^2} - \lambda(t) \cos q,$$

where $\lambda(t) = \ell(t)$ in units where $g = 1$. For constant $\lambda > 0$ the motion is periodic and the action variable J exists and is a constant of the motion. The action-angle variables

$$w = w(q, p, \lambda), \quad J = J(q, p, \lambda), \tag{26}$$

define a canonical transformation from the old phase space variables (q, p) to the new action-angle variables (w, J). If the constant λ is replaced by a function of t the transformation (26) is still canonical even though H and J are not constants of the motion. If λ changes 'slowly' the motion is approximately periodic. For a single oscillation the change in λ, J and E will be small. However, over many oscillations λ may change appreciably (e.g. it may double). Over many oscillations E may also change appreciably. However, it turns out that over many oscillations the change in J remains small. This remarkable property of action variables is called *adiabatic invariance*.

Adiabatic invariance is an approximate conservation law (how good an approximation depends on how slow λ changes). The origin of this property can be traced back to an exact property of the action variable in the case that λ is constant:

$$\int_0^T \left(\frac{\partial J}{\partial \lambda}\right)_{q,p} dt = 0, \tag{27}$$

where $T = (\partial J/\partial E)_\lambda$ is the period (for constant λ the motion is periodic). Here the q,p subscripts indicate that in taking the partial derivative with respect to λ the old phase space coordinates are treated as constants. The result (27) is the statement that the average value of $(\partial J/\partial \lambda)_{q,p}$ is zero. For example consider the Hamiltonian

$$H = \frac{1}{2}p^2 + \frac{1}{2}\lambda q^2.$$

If λ is a positive constant this is a simple harmonic oscillator with angular frequency $\sqrt{\lambda}$. A simple calculation gives

$$J = \oint p \, dq = \frac{2\pi E}{\sqrt{\lambda}} = \frac{\pi p^2}{\sqrt{\lambda}} + \pi \sqrt{\lambda} q^2.$$

Now

$$\left(\frac{\partial J}{\partial \lambda}\right)_{q,p} = -\frac{\pi}{2}\lambda^{-3/2}p^2 + \frac{\pi}{2}\lambda^{-1/2}q^2 = \pi\lambda^{-3/2}(-T + V),$$

which has average value zero; for a simple harmonic oscillate the average value of the kinetic energy equals the average potential energy (see Problem Sheet 1).

Now allow λ to depend on time and consider the action variable $J = J(q, p, \lambda(t))$ defined through the canonical transformation (26). Now

$$\frac{dJ}{dt} = \{J, H\} + \left(\frac{\partial J}{\partial t}\right)_{q,p} = 0 + \dot{\lambda}(t)\left(\frac{\partial J}{\partial \lambda}\right)_{q,p}.$$

Consider the change in J over from $t = 0$ to $t = T_0$ where T_0 is the period for the initial value of λ.

$$J(T_0) - J(0) = \int_0^{T_0} \dot{\lambda}(t)\left(\frac{\partial J}{\partial \lambda}\right)_{q,p} dt.$$

Using the mean value theorem for integrals

$$J(T_0) - J(0) = \dot{\lambda}(\tau)\int_0^{T_0} \left(\frac{\partial J}{\partial \lambda}\right)_{q,p} dt,$$

for some τ between 0 and T_0. If $\dot{\lambda}$ is small then both the pre-factor and the integral is also small based on (27). Over many oscillations λ may change significantly, but J remains close to its initial value.

Equation (27) can be obtained by evaluating the integrand

$$\left(\frac{\partial J}{\partial \lambda}\right)_{q,p} = \left(\frac{\partial J}{\partial E}\right)_\lambda \left(\frac{\partial E}{\partial \lambda}\right)_{q,p} + \left(\frac{\partial J}{\partial \lambda}\right)_E,$$

using $J(q, p, \lambda) = J(E(q, p, \lambda), \lambda)$. To show that this integrates to zero over a period write[13]

$$\left(\frac{\partial E}{\partial \lambda}\right)_{q,p} = -\left(\frac{\partial p}{\partial \lambda}\right)_{q,E}\left(\frac{\partial E}{\partial p}\right)_{q,\lambda} = -\left(\frac{\partial p}{\partial \lambda}\right)_{q,E}\frac{dq}{dt}.$$

Accordingly,

$$\int_0^T \left(\frac{\partial J}{\partial \lambda}\right)_{q,p} dt = -\left(\frac{\partial J}{\partial E}\right)_\lambda \oint \left(\frac{\partial p}{\partial \lambda}\right)_{q,E} dq + T\left(\frac{\partial J}{\partial \lambda}\right)_E,$$

which vanishes as the integral on the right hand side is $(\partial J/\partial \lambda)_E$ and $T = (\partial J/\partial E)_\lambda$.

4.4 Integrability and Noether's Theorem

Treating Hamilton's Characteristic Function, $W(q_i, \alpha_i)$, as a type 2 generating function we have

$$p_j = \frac{\partial W(q_i, \alpha_i)}{\partial q_j}.$$

Assuming this can be inverted to express the α_i as functions of q_i and p_i then we have N functions with the properties

$$\{\alpha_i, \alpha_j\} = 0, \quad \text{and} \quad \{\alpha_i, H\} = 0, \tag{28}$$

the former hold as the α are 'new momenta' and the latter follows since the α_i are constants of the motion.

Instead of going through Hamilton-Jacobi theory one can consider Hamiltonian systems for which (28) holds. This leads to a definition:

A system with Hamiltonian $H(q_1, ..., q_N, p_1, ..., p_N)$ is called Liouville integrable if there exist N independent functions[14] $\alpha_i(q_1, q_2, ..., q_N, p_1, ..., p_N)$ with the properties

$$\{\alpha_i, \alpha_j\} = 0, \quad \text{and} \quad \{\alpha_i, H\} = 0.$$

In Lagrangian mechanics, the momentum associated with a cyclic coordinate is conserved. This result carries over to Hamiltonian mechanics. A coordinate, q_1 say, is called cyclic if $\partial H/\partial q_1 = 0$. If this holds then $\dot{p}_1 = -\partial H/\partial q_1 = 0$. A more general form of this result is as follows:

---

[13]Use the reciprocal relation: if $f(x, y, z) =$ constant then $\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x \left(\frac{\partial z}{\partial x}\right)_y = -1$ with $x = E$, $y = \lambda$ and $z = p$; q is treated as a constant.

[14]Independent means that the differentials $d\alpha_i$ are linearly independent.

If the Hamiltonian is invariant (meaning $H(q'_i, p'_i, t) = H(q_i, p_i, t)$) under the deformation

$$q'_i = q_i + s f_i(q_1, ..., q_N) + \mathcal{O}(s^2), \quad p'_i = p_i - s \sum_{j=1}^{N} p_j \frac{\partial f_i(q_1, ..., q_N)}{\partial q_j} + \mathcal{O}(s^2),$$

where s is a small deformation parameter, then

$$\sum_{i=1}^{N} p_i f_i(q_1, ..., q_N)$$

is a constant of the motion. This is exactly the result from section 2.4 converted into Hamiltonian form.

Note that the conserved quantity is linear in the momenta. This is because the deformation is essentially an active point transformation. In addition the deformation is a canonical transformation. We can exploit this to give a more general (and simpler) statement of Noether's theorem. This more general formulation also includes the 'conservation of energy' as a special case.

Consider the transformation

$$Q_i = q_i + s\{q_i, \alpha\} + \mathcal{O}(s^2), \quad P_i = p_i + s\{p_i, \alpha\} + \mathcal{O}(s^2), \tag{29}$$

where s is a small deformation parameter and $\alpha = \alpha(q_1, ..., q_N, p_1, ..., p_N, t)$ is a function of the phase space variables and time. Up to terms of order s^2 the above is a canonical transformation. The proof of this uses the Jacobi identity - a similar argument was used to show that time evolution defines a canonical transformation. Noether's theorem states that if $K(Q_i, P_i, t) = H(Q_i, P_i, t)$ then α is a constant of the motion. The proof is short (see Problem Sheet 7).

By definition a canonical transformation preserves the form of Hamilton's equation. However, under such a transformation the Hamiltonian function changes. A *symmetry* is a continuous deformation (and canonical transformation) which does not change the Hamiltonian function. Noether's theorem states that there is a conserved quantity, α, associated with the symmetry. Moreover, this quantity directly enters the formula (29) for the deformation (for small s). α is also called the *generator* of the symmetry.

For the Kepler problem, consider the deformation generated by $\alpha = p_\theta$

$$r' = r + s\{r, \alpha\} = r, \quad \theta' = \theta + s\{\theta, \alpha\} = \theta + s,$$

$$p'_r = p_r + s\{p_r, \alpha\} = p_r, \quad p'_\theta = p_\theta + s\{p_\theta, \alpha\} = p_\theta.$$

The deformation is a shift of the angle θ by s which is simply a rotation with angle s. Under this deformation the form of the Kepler Hamiltonian is unchanged. That is rotation is a symmetry

generated by the angular momentum p_θ. A more complicated symmetry is generated by the components of the Laplace-Runge-Lenz vector.

If the Hamiltonian has no explicit time-dependence then it is itself a constant of the motion. This is 'conservation of energy' or the 'Beltrami formula'. This result is included in the Hamiltonian formulation of Noether's theorem. To see this take $\alpha(q_i, p_i, t) = H(q_i, p_i, t)$. For small s, $K(Q_i, P_i, t) = H(Q_i, P_i, t + s)$. If the Hamiltonian has no explicit time dependence $K(Q_i, P_i) = H(Q_i, P_i)$. One can say that if time-translation is a symmetry then there exists a conserved quantity, the Hamiltonian or energy.

5 Rigid Bodies

This short chapter is an introduction to the theory of rigid bodies; see Chapters 4 and 5 of Goldstein, Safko and Poole and sections 31 to 35 of Landau and Lifshitz.

A rigid body is a collection of particles subject to the (holonomic) constraint that the distance between constituent particles is fixed.

5.1 Motion of a Rigid Body

The motion of a rigid body is governed by the Lagrangian $L = T - V$ assuming all forces are conservative. The potential energy describes the external forces on the body (not constraint forces).

Now assume that the external forces are purely gravitational.

Consider one constituent particle of mass m with potential energy

$$V(\mathbf{r}) = m\Phi(\mathbf{r}).$$

Here $\Phi(\mathbf{r})$ is the gravitational potential - the force on the particle is

$$\mathbf{F} = -\nabla V = -m\nabla\Phi.$$

For terrestrial gravity in cartesian coordinates (x, y, z) one has $\Phi = gz$ where g is the (constant) acceleration due to gravity. In the gravitational field of a spherically symmetric body of mass M and radius R, the potential is

$$\Phi = -\frac{GM}{r}, \quad r > R.$$

For a number of particles the potential energy is

$$V = \sum_n m^{(n)}\Phi(\mathbf{R}^{(n)}),$$

where $m^{(n)}$ and $\mathbf{R}^{(n)}$ are the mass and position of the nth particle, respectively. For a continuous body the sum becomes an integral

$$V = \int_{\text{body}} \rho(\mathbf{r})\Phi(\mathbf{r})\, d^3\mathbf{r},$$

where ρ is the mass density.

Now consider the *approximation*

$$V \approx M\Phi(\mathbf{R}),$$

where M is the total mass and $\mathbf{R}$ is the position of the centre of mass of the body:

$$M = \sum_n m^{(n)}, \quad \mathbf{R} = \frac{1}{M}\sum_n m^{(n)}\mathbf{R}^{(n)}.$$

This approximation is exact if the gravitational potential is linear in the coordinates (as in terrestrial gravity where ϕ is proportional to z).
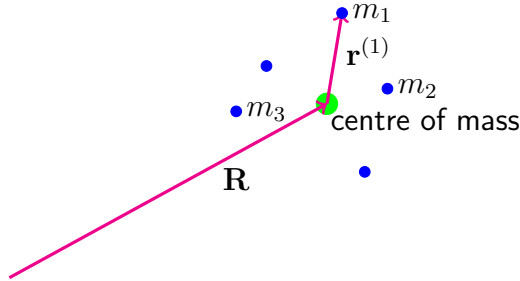
We now consider the kinetic energy of a rigid body

$$T = \frac{1}{2}\sum_n m^{(n)}\mathbf{V}^{(n)}\cdot\mathbf{V}^{(n)},$$

where $\mathbf{V}^{(n)} = d\mathbf{R}^{(n)}/dt$ is the velocity of the nth particle. The kinetic energy can be split into 'translational' and 'rotational' parts. The positions of the constituent particles may be written in the form

$$\mathbf{R}^{(n)} = \mathbf{R} + \mathbf{r}^{(n)},$$

where $\mathbf{r}^{(n)}$ is the position of the nth particle relative to the centre of mass.

It is straightforward to show that

$$T = T_{\text{trans}} + T_{\text{rot}},$$

where

$$T_{\text{trans}} = \frac{1}{2}M\mathbf{V}\cdot\mathbf{V}, \quad T_{\text{rot}} = \frac{1}{2}\sum_n m^{(n)}\mathbf{v}^{(n)}\cdot\mathbf{v}^{(n)},$$

where $\mathbf{V} = d\mathbf{R}/dt$ is the velocity of the centre of mass and $\mathbf{v}^{(n)} = d\mathbf{r}^{(n)}/dt$ is the velocity of the nth particle relative to the centre of mass.

The relative velocity $\mathbf{v}^{(n)}$ can be written in the form

$$\mathbf{v}^{(n)} = \boldsymbol{\omega}\times\mathbf{r}^{(n)}, \tag{30}$$

where $\boldsymbol{\omega}$ is the *angular velocity* of the rigid body.

Note that $|\mathbf{r}^{(n)}|$, the distance between the centre of mass and the nth particle, is constant. Therefore $\mathbf{r}^{(n)}(t + \delta t)$ will be a rotation of $\mathbf{r}^{(n)}(t)$, i.e. $\mathbf{r}^{(n)}(t + \delta t) = \mathcal{R}\mathbf{r}^{(n)}(t)$. The rotation, $\mathcal{R}$, is the *same* for all constituent particles. Now assume δt is small. A small rotation of a vector, $\mathbf{U}$, about an axis in the direction of the unit vector $\mathbf{n}$ can be expressed though a vector product

$$\mathcal{R}(\mathbf{n}, \delta\theta)\mathbf{U} = \mathbf{U} + \delta\theta\mathbf{n} \times \mathbf{U},$$

so that

$$\mathbf{r}^{(n)}(t + \delta t) - \mathbf{r}^{(n)}(t) = \delta\theta\mathbf{n} \times \mathbf{r}^{(n)}.$$

Dividing through by δt and taking the limit $\delta t \to 0$ yields (30) where the direction of $\boldsymbol{\omega}$ is the instantaneous axis of rotation and its magnitude is the rate of rotation about this axis.

The rotational kinetic energy is

$$T_{\text{rot}} = \frac{1}{2}\sum_n m^{(n)}|\boldsymbol{\omega} \times \mathbf{r}^{(n)}|^2 = \frac{1}{2}\sum_n m^{(n)}\left[\omega^2|\mathbf{r}^{(n)}|^2 - (\boldsymbol{\omega} \cdot \mathbf{r}^{(n)})^2\right], \tag{31}$$

using $|\boldsymbol{\omega} \times \mathbf{r}|^2 = \omega^2 r^2 - (\boldsymbol{\omega} \cdot \mathbf{r})^2$. In components (31) takes the form

$$T_{\text{rot}} = \frac{1}{2}\sum_n m^{(n)}\omega_i \left(x_k^{(n)}x_k^{(n)}\delta_{ij} - x_i^{(n)}x_j^{(n)}\right)\omega_j, \tag{32}$$

where the vectors are given with respect to some orthonormal basis, $\boldsymbol{\omega} = (\omega_1, \omega_2, \omega_3)$ and $\mathbf{r}^{(n)} = (x_1^{(n)}, x_2^{(n)}, x_3^{(n)})$. The summation convention is used in (32). Equation (32) can be written in the more compact form

$$T_{\text{rot}} = \frac{1}{2}\omega_i I_{ij}\omega_j, \tag{33}$$

where

$$I_{ij} = \sum_n m^{(n)}\left(x_k^{(n)}x_k^{(n)}\delta_{ij} - x_i^{(n)}x_j^{(n)}\right), \tag{34}$$

are the components of the *inertia tensor*. For a continuous body,

$$I_{ij} = \int_{\text{body}} \rho(\mathbf{r})\left(r^2\delta_{ij} - x_i x_j\right)\, d^3\mathbf{r}. \tag{35}$$

The diagonal elements are called moments of intertia and can be written in the form

$$I_{11} = \int_{\text{body}} \rho(\mathbf{r})\left(x_2^2 + x_3^2\right)\, d^3\mathbf{r}, \quad I_{22} = \int_{\text{body}} \rho(\mathbf{r})\left(x_3^2 + x_1^2\right)\, d^3\mathbf{r},$$

and similarly for I_{33}. The off-diagonal elements of I_{ij} are called products of inertia

$$I_{12} = - \int_{\text{body}} \rho(\mathbf{r}) x_1 x_2 \, d^3 \mathbf{r},$$

and similarly for the other off-diagonal elements. Note that $I_{ij} = I_{ji}$, that is the inertia tensor is symmetric which can be diagonalised by an orthogonal transformation. Equivalently, by a rotation of axes one can bring the off diagonal elements to zero. When this is done the moments of inertia are called principal moments of inertia, written I_1, I_2, I_3. Axes for which the products of inertia vanish are called *principal axes*. In this case the rotational energy formula simplifies to

$$T_{\text{rot}} = \frac{1}{2} \left(I_1 \omega_1^2 + I_2 \omega_2^2 + I_3 \omega_3^2 \right). \tag{36}$$

This formula is the rotational analogue of the standard kinetic energy formula $T = \frac{1}{2} m v^2$.

The Lagrangian for a rigid body is in principle $L = \frac{1}{2} M \mathbf{V} \cdot \mathbf{V} + T_{\text{rot}} - V$. Although we have expressed T_{rot} in terms of the angular velocity, the components of the angular velocity are not time derivatives of generalised coordinates. We require generalised coordinates for a rigid body. To exploit the above formulas for T_{rot} we would need to express the components of the angular velocity in terms of the coordinates and time derivatives of the coordinates.
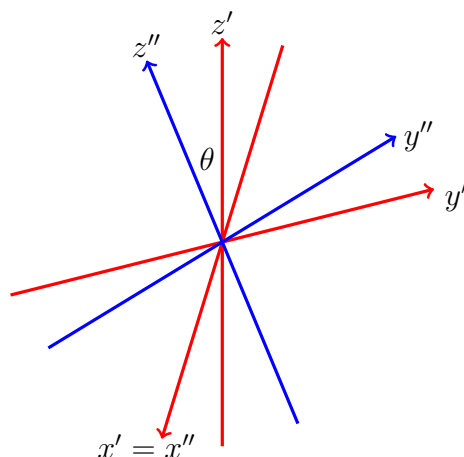
6 coordinates are required to describe a rigid body: 3 to specify the centre of mass and 3 to specify the orientation of the body. *Euler Angles* (ϕ, θ, ϕ) are a set of three coordinates to describe the orientation of a rigid body. Different conventions for Euler angles are in use. We will use the zxz convention which is used by both Goldstein and Landau and Lifshitz. Start with xyz axes where the origin is at the centre of mass. The idea is now to rotate the axes into new axes $x_1 x_2 x_3$ which are principal axes. Build this rotation out of 3 separate rotations with angles ϕ, θ and ψ:

Rotate about the z axis with angle ϕ. Call the new axes x', y' and z'

Rotate about the x' axis with angle θ. Call the new axes x'', y'' and z''. Note that the angles ϕ and θ are spherical polar angles describing the orientation of the new z'' axis with

repect to the original xyz axes.

Rotate about the z'' axis with angle ψ. Call new axes x_1, x_2, x_3. It is shown in the appendix that the components of the angular velocity are

$$\omega_1 = \dot{\phi}\sin\theta\sin\psi + \dot{\theta}\cos\psi, \quad \omega_2 = \dot{\phi}\sin\theta\cos\psi - \dot{\theta}\sin\psi, \quad \omega_3 = \dot{\phi}\cos\theta + \dot{\psi}, \qquad (37)$$
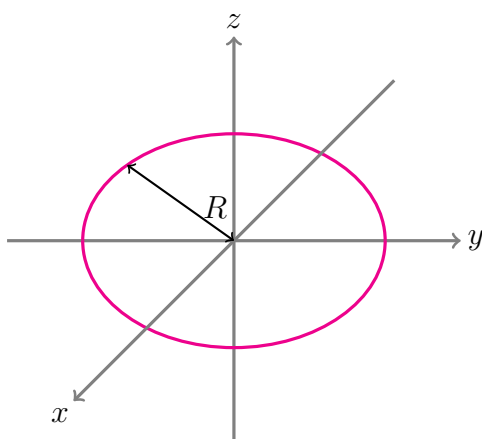
are the components of angular velocity in terms of the Euler angles.

For very symmetric rigid bodies the principal axes are not unique. For a sphere of uniform density any axis through the centre is a principal axis. Here

$$I_1 = I_2 = I_3 = \frac{2}{5}MR^2.$$

Another simple example is a uniform disc. Here

$$I_3 = \frac{1}{2}MR^2, \quad I_1 = I_2 = \frac{1}{2}I_3.$$

5.2 The Symmetric Top

In the previous section we considered the motion of a *freely rotating* rigid body. Here there are 6 generalised coordinates describing the position of the centre of mass and the orientation of the rigid body. The number of generalised coordinates is reduced if the motion of the rigid body is constrained. The simplest example is rotation about a fixed, not necessarily principal, axis (e.g., a flywheel). Taking the z-axis as the fixed axis of rotation

$$T = \frac{1}{2} I \omega^2, \quad I = \int_{\text{body}} (x^2 + y^2) \, dx \, dy \, dz.$$

Here ω is $d\phi/dt$ where ϕ is the angle required to describe the orientation of the rigid body. In this case the constraint reduces the number of generalised coordinates from 6 to 1.

A more interesting example of constrained rotation is where the rigid body moves with one point fixed thereby reducing the number of coordinates from 6 to 3. If the fixed point is the centre of mass the analysis is the same same as in the previous section with $\mathbf{R} = \mathbf{0}$. The total kinetic energy is

$$T = T_{\text{rot}} = \frac{1}{2} \omega_i I_{ij} \omega_j$$

where

$$I_{ij} = \int_{\text{body}} \rho(\mathbf{r}) \left(r^2 \delta_{ij} - x_i x_j \right) \, \mathsf{d}^3 \mathbf{r}.$$

In section 5.1 it was crucial that the origin be the centre of mass as otherwise the splitting of T into translational and rotational parts would not be valid. As in section 5.1 the $x_1 x_2 x_3$ axes can be chosen as principal axes (which intersect at the fixed point not the centre of mass) so that

$$T = \frac{1}{2} (I_1 \omega_1^2 + I_2 \omega_2^2 + I_3 \omega_3^2),$$

where the components of the angular velocity with respect to the principal axes are given by equation (37). Here there is no translational part and the origin does not have to be the centre of mass. An example is a symmetric top. Consider a body symmetric about the x_3 axis such as a cone of uniform density.
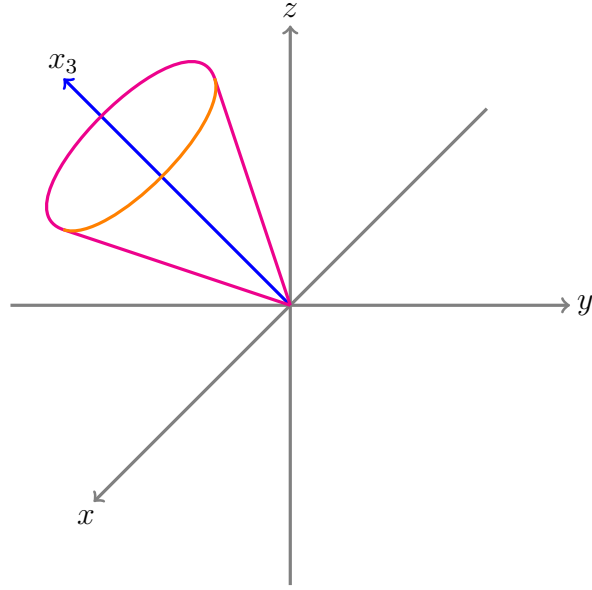
Due to the symmetry $I_1 = I_2$ so that the formula for T simplifies to

$$T_{\text{rot}} = \frac{I_1}{2} \left(\dot{\theta}^2 + \sin^2 \theta \dot{\phi}^2 \right) + \frac{I_3}{2} (\dot{\psi} + \dot{\phi} \cos \theta)^2,$$

as the cross terms in ω_1^2 and ω_2^2 cancel.

The Lagrangian for a symmetric top fixed at one point is $L = T_{\text{rot}} - M\Phi(\mathbf{R})$. Here $\Phi = gz = gl \cos \theta$. where l is the distance between the origin and the centre of mass of the cone. Accordingly,

$$L = \frac{I_1}{2} \left(\dot{\theta}^2 + \sin^2 \theta \, \dot{\phi}^2 \right) + \frac{I_3}{2} (\dot{\psi} + \dot{\phi} \cos \theta)^2 - Mgl \cos \theta.$$

Here ϕ and ψ are cyclic. The momenta

$$p_\phi = I_1 \sin^2 \theta \, \dot\phi + I_3(\dot\psi + \dot\phi \cos\theta)\cos\theta, \quad p_\psi = I_3(\dot\psi + \dot\phi \cos\theta) = I_3\omega_3,$$

are conserved. It follows that

$$\dot\phi = \frac{p_\phi - p_\psi \cos\theta}{I_1 \sin^2 \theta}.$$

The total energy is

$$E = T_{\text{rot}} + V = \frac{I_1}{2}\dot\theta^2 + \frac{(p_\phi - p_\psi \, \cos\theta)^2}{2I_1 \sin^2 \theta} + \frac{p_\psi^2}{2I_3} + Mgl\cos\theta.$$

This can be written in the form

$$E = \frac{I_1}{2}\dot\theta^2 + U_{\text{eff}}(\theta),$$

where U_{eff} is an effective potential. This admits a solution where θ and $\dot\phi$ are constant (precession). In general there are two values of θ for which $\dot\theta = 0$.
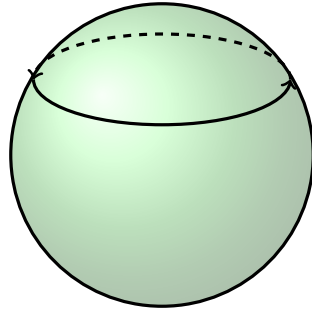

Figure 2: Precession, i.e., $\theta = $ constant and $\dot\phi = $ constant.
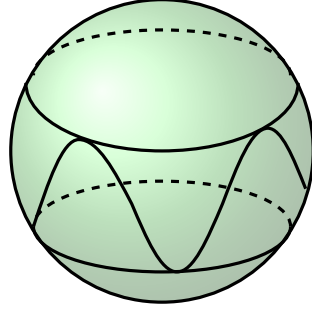
Figure 3: Nutation, i.e., $\dot{\phi}$ is not constant (does not change sign).

5.3 Appendix: Angular Velocity and Euler Angles

The rotational kinetic energy of a rigid body can be written as

$$T_{\text{rot}} = \frac{1}{2}(I_1\omega_1^2 + I_2\omega_2^2 + I_3\omega_3^2),$$

where I_1, I_2, I_3 are the three principal moments of inertia and ω_1, ω_2, ω_3 are the components of the angular velocity with respect to the moving body frame (where the axes are chosen to be principal). In terms of the Euler angles ϕ, θ, ψ

$$\omega_1 = \dot{\phi}\sin\theta\sin\psi + \dot{\theta}\cos\psi, \quad \omega_2 = \dot{\phi}\sin\theta\cos\psi - \dot{\theta}\sin\psi, \quad \omega_3 = \dot{\phi}\cos\theta + \dot{\psi}.$$

This note outlines the derivation of the above expressions.

Euler angles are defined by obtaining the body frame via three rotations of the original frame. The coordinates with respect to the body frame, denoted x_i', are related to the coordinates with respect to a fixed inertial frame, denoted x_i, by the matrix equation

$$x_i' = A_{ij}x_j,$$

where A_{ij} are the entries of a 3×3 rotation matrix A. This can be written as a product of three elementary rotation matrices:

$$A = BCD,$$

where

$$D = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix}, \quad B = \begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

As $x_i = (A^{-1})_{ij}x_j'$, $v_i = \dot{x}_i = d(A^{-1})_{ij}/dt\, x_j$. Now

$$\frac{d}{dt}A^{-1} = -A^{-1}\dot{A}A^{-1},$$

53

so that
$$v = -A^{-1}\dot{A}A^{-1}x' = -A^{-1}\dot{A}x.$$

The components of the angular velocity are given through the matrix equation

$$-A^{-1}\dot{A} = \begin{pmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{pmatrix}$$

Here the angular velocity components are with respect to the intertial frame. With respect to the moving frame $v' = -\dot{A}A^{-1}x'$ so that

$$-\dot{A}A^{-1} = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix}$$

The calculation of the angular velocity can be split into three - assume that only one angle, say ψ, depends on time and treat the other two as constants. Treating ϕ and θ as constants

$$-\dot{A}A^{-1} = -\dot{D}CBB^{-1}C^{-1}D^{-1} = -\dot{D}D^{-1}$$

$$= -\dot{\psi} \begin{pmatrix} -\sin\psi & \cos\psi & 0 \\ -\cos\psi & -\sin\psi & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -\dot{\psi} & 0 \\ \dot{\psi} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

so that $\omega_1 = \omega_2 = 0$ and $\omega_3 = \dot{\psi}$. Setting ψ and ϕ to be constant gives

$$-\dot{A}A^{-1} = -D\dot{C}BB^{-1}C^{-1}D^{-1} = -D\dot{C}C^{-1}D^{-1}$$

$$= -\dot{\theta} \begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & -\sin\theta & \cos\theta \\ 0 & -\cos\theta & -\sin\theta \end{pmatrix}$$

$$\times \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \dot{\theta} \begin{pmatrix} 0 & 0 & -\sin\psi \\ 0 & 0 & -\cos\psi \\ \sin\psi & \cos\psi & 0 \end{pmatrix},$$

so that $\omega_2 = -\dot{\theta}\sin\psi$, $\omega_1 = \dot{\theta}\cos\psi$, $\omega_3 = 0$. A similar argument shows that for constant ψ and θ, $\omega_1 = \dot{\phi}\sin\theta\sin\psi$, $\omega_2 = \dot{\phi}\sin\theta\cos\psi$, $\omega_3 = \dot{\phi}\cos\theta$.