# The Principles of Dynamics

Dr C. T. Whelan.[1]

Lent 2000.

[1] LaTeXed by Tom Bentley

# Technicalities

## Copyright & Disclaimer



## Thanks

I would like to thank Dr. Colm Whelan for allowing me to produce these notes, and for his invaluable help in finding the many errors. Any remaining errors are mine and corrections should be sent to `tjb37@cam.ac.uk`.

## Books

By far the best book for this course is Goldstein's *Classical Mechanics*: It covers all the major parts of the course in a helpful and clear style. Landau and Lifshitz's *Mechanics* is also quite good, especially for the section on Adiabatic Invariants.

## Some Other Points to Note

It is a great shame that this course is so very avoidable in the Mathematical Tripos: The material it teaches you lies behind vast swathes of theoretical physics and applied mathematics, not least among which are:

- Quantum Mechanics,
- General Relativity.

However the theory has applications in differential equations, financial modelling, population modelling, and many other areas. In short, if you intend to study these things later, or if you simply want to understand mechanics at it's most natural level, this is a very good course to study.

# Contents

# Chapter 0

# Introduction: Sin and Death

**Synopsis:** A short historical introduction to serve as motivation for the course, and to put the material in context. The underlying Principle of Mechanics as being a *Variational Principle*.

**Newton's Laws** During the 17th century thinking man was obssed with Sin and Death. How, if God was perfect could He create a Universe which was not perfect? Why should there be so much suffering and apparent waste in the world? The answer was that it was all in some way necessary to lead to the final point: The salvation of Mankind. That path was determined by the end point.
Newton had given the world a mechanicsal universe, working according to a set of simple fixed laws and the whole majestic clockwork had no need for a Divine Hand to drive it.

Some of the greatest minds of the time were seduced into trying to find the underlying metaphysical reason for Newton's Laws: Trying, if not to find the hand of the creator then at least to find his finger prints on the Cosmos.
Leibnitz, in particlular, was determined to prove that all was for the best in the best possible world. He felt that the world we live in exhibits:

> 'The greatest simplicity in its premises and the greatest wealth in it phenomena.'

Leibitz had 3 major problems with Newtonian Mechanics

1. Occult Virtues: Leibnitz held that Newton had not explained 'Gravity' by postulating a 'Gravitational Force' - Forces are define in terms of directly measurable quantities (masses and velocities and their rates of change), ie as a property of their *motion*. Leibtitz felt that the underlying mechanism had not been found: He argued that Newtonian theory was a kinematical one, that is a science of motion. What he sought was a science of powers. Ie. Dynamics
   Leibnitz recognised that energy was conserved in certain mechanical systems and suggested that a principle of energy conservation might be the underlying one, from which all Laws of Motion could be derived. He deduced something like a potential energy function.

2. Action at a Distance: To get round this he postulated an ether of very fine particles. Much of his ideas on this subject anticipated what we would call a field

theory..

3. Absolute Space: How could the stars be treated as an absolute frame of reference? Leibitz argued that space was not a thing in itself, just a relation between objects in it. He claimed that all inertial frames should be as good as the next.

**Action** In this intellectual climate Maupertuis advanced an argument based on God's efficiency. He claimed that the Laws of Nature were acted out in a way where the least possible *action* was expended. He was unclear as to what 'action' was, but is had something to do with $mvs$.

**The Variational Principle** Euler liked the idea and defined the action of a particle moving from $A$ to $B$ as

$$\int_A^B mv \, \mathrm{d}s.$$

He postulated that for any given particle, the path taken was 'chosen' so that the action would be least. Actually, he always assumed the existance of a potential energy function $V(r)$ from which all forces were derived. Ie in our terms we are dealing with a conservative force.

To progress further he invented the *Calculus of Variations*[1], that is, a *necessary condition* to *extremise* of the integral

$$\int_A^B F(y, \frac{\mathrm{d}y}{\mathrm{d}t}, t) \, \mathrm{d}t$$

is that

$$\frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial F}{\partial \dot{y}}\right) = 0.$$

We can generalise this to a set of $N$ independent coordinates $y_n$:

$$\delta \int_A^B F(y_n, \dot{y}_n, t) \, \mathrm{d}t = 0$$

$$\Rightarrow \quad \frac{\partial F}{\partial y_n} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial F}{\partial \dot{y}_n}\right) = 0, \qquad 1 \leq n \leq N.$$

This is called the *Variational Principle*.
Having shown this Euler was able to show that if we had a conservative system (i.e. there exists a potential energy function $V$) then the path of a particle as deduced from the variational principle was precisely the same as Newton's Laws:

Consider

$$\int_A^B mv \, \mathrm{d}s = \int_{t(a)}^{t(b)} mv^2 \, \mathrm{d}t = \int_{t(a)}^{t(b)} 2T \, \mathrm{d}t$$

[1] See Methods IB if you need to review this - it forms a crucial rôle in this course.

where $T$ = kinetic energy = $\frac{m}{2}\left(\dot{x}^2 + \dot{y}^2\right)$. If there exists a potential energy function $V$ then

$$T + V = \text{ constant} = E \ \Rightarrow\ \mathsf{L} \equiv T - V = 2T - E.$$

We want to make the integral

$$\int_{t(a)}^{t(b)} \mathsf{L}(x, \dot{x}, y, \dot{y})\,\mathrm{d}t$$

stationary, but the Euler Lagrange equations imply that

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathsf{L}}{\partial \dot{x}}\right) - \frac{\partial \mathsf{L}}{\partial x} = 0 \\ \Rightarrow\quad &\frac{\mathrm{d}}{\mathrm{d}t}\left(m\dot{x}\right) = -\frac{\partial V}{\partial x}, \end{aligned}$$

and similarly for the $y$s. But these are Newton's Laws

$$\mathbf{F} = -\nabla V = \frac{\mathrm{d}}{\mathrm{d}t}\left(m\dot{\mathbf{x}}\right)$$

So the Variational Principle and energy conservation imply Newton's Law. Quite clever maybe, but it does it give us anything new?
Yes: The $y$'s in the E-L equations are implicitly dependent of any particluar coordinate system. We used Cartesian coordinates, but there was no reason to do this.

**Generalized Coordinates** Let us introduce *generalized coordinates*

$$\{q_1, ..., q_{3N}\}$$

If we have a system of $N$ particles (in 3 dimensions) free from constraints, it has $3N$ degrees of freedom, and we can choose to describe the motion in terms of *any* $3N$ independent variables $\{q_i\}_{i=1}^{3N}$. Usually these generalized coordinates will not form a convenient set of $N$ vectors in $\mathbb{R}^3$.
**Example:** Planetry Motion. We have a radial force

$$\frac{\mu m}{r^2}$$

$\Rightarrow$ the potential function $V(r)$ is radial. If we choose our coordinates to be $(r, \theta)$ we can define

$$\mathsf{L} \equiv T - V = \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2\right) - V(r)$$

Now if $\delta \int \mathsf{L}\,\mathrm{d}t = 0$ then the E-L equations for $\theta$ imply:

$$\begin{aligned} &\frac{\partial \mathsf{L}}{\partial \theta} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathsf{L}}{\partial \dot{\theta}} = 0 \\ \Rightarrow\quad &\frac{\mathrm{d}}{\mathrm{d}t}\left(mr^2\dot{\theta}\right) = 0 \\ \Rightarrow\quad &mr^2\dot{\theta} = \text{ constant } = l. \end{aligned}$$

We've derived the conservation of angular momentum, you'll recall from IA Dynamics that if $m$ is constant this implies Kepler's Second Law. Applying the E-L equations for $r$ implies:

$$\begin{aligned} \frac{\partial \mathsf{L}}{\partial r} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathsf{L}}{\partial \dot{r}} &= 0 \\ \Rightarrow \quad \frac{\mathrm{d}}{\mathrm{d}t}(m\dot{r}) - mr\dot{\theta}^2 + \frac{\partial V}{\partial r} &= 0 \end{aligned}$$

But now the conservation of angular momentum can be used to replace $\dot{\theta}$ to get

$$m\ddot{r} - \frac{l^2}{mr^3} = -\frac{\partial V}{\partial r}$$

i.e.

$$\begin{aligned} m\ddot{r} &= -\frac{\mathrm{d}}{\mathrm{d}r}\left(V + \frac{1}{2}\frac{l^2}{mr^2}\right) \\ \Rightarrow \quad m\ddot{r}\dot{r} \equiv \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}m\dot{r}^2\right) &= -\frac{\mathrm{d}}{\mathrm{d}r}\left(V + \frac{1}{2}\frac{l^2}{mr^2}\right)\frac{\mathrm{d}r}{\mathrm{d}t} \\ \Rightarrow \quad \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}m\dot{r}^2 + \frac{l^2}{2mr^2} + V(r)\right) &= 0 \end{aligned}$$

This is the conservation of energy.

**The Differences between Analytic and Vectorial Mechanics**

- In analytic dynamics the equations of motion can be deduced from a single unifying principle. In vectorial mechanics we have Newton's Laws
- In vectorial mechanics we look at the motion of the individual particles that make up the system. In analytic dynamics we treat the system as a whole
- It frequently happens that there are constraints on the system (eg. in a rigid body we have the constraint that the distances between particles remains fixed) In the Newtonian point of view the must ascribe forces to these constraints. In analytic mechanics we don't care about these forces, it is enough to know the constraints.

# Chapter 1

# From Newton to Lagrange

**Synopsis:** A brief recap of dynamics, followed by the development of the Lagrangian formalism. Lagrange's Equations must take various forms depending on the nature of the forces (conservative, derivable from a velocity dependent potential, or even more general), and Constraints (holonomic, monogenic, etc.). Hamilton's Principle. Conserved quantities.

## 1.1 Summary of Newton's Laws

If $\mathbf{r}$ is the radius vector of a particle wrt some origin, then the *velocity*, $\mathbf{v}$ is

$$\mathbf{v} = \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}$$

The *linear momentum* $\mathbf{P}$ is $\mathbf{P} = m\mathbf{v}$. Newton's first two laws imply

$$\mathbf{F} = \frac{\mathrm{d}}{\mathrm{d}t}(m\mathbf{v}) \tag{1.1}$$

A reference frame in which (1) holds is called *inertial* or *Gallelian*.
Newton's Third Law:

> 'To every action there is an equal and opposite reaction.'

What does this mean? Is it always true?
Suppose we have two particles $i$ and $j$ and suppose $i$ exerts a force $\mathbf{F}_{ij}$ on $j$. Then we can translate NIII to read $\mathbf{F}_{ij} = \mathbf{F}_{ji}$ Ie the force $j$ exerts of $i$ is equal and opposite. This is the 'weak' formulation of the law.
If $\mathbf{F}_{ij} = -\mathbf{F}_{ji}$ and the forces act along the line connecting the particles, we have a central force. This is the strong form of the law, and it holds for many forces in Nature eg. Gravity, Electrostatics.
Consider the example of the Biot-Savart Law between moving charges:

1. If we have two charges moving with parallel velocity vectors that are not perpendicular to the line joining the two particles

   Then the weak form holds, but not the strong form.

2. Consider two charges moving instantaneously such that their velocity vectors are perpendicular

   The 2nd charge exerts a non-zero force on the first while experiencing no 'reaction' force at all.

### 1.1.1 Momentum, Work and Forces

If the force acting on the particle is zero then

$$\frac{\mathrm{d}}{\mathrm{d}t}(m\mathbf{v}) = 0 \Rightarrow m\mathbf{v} = \text{ const } = \mathbf{P}$$

Define the *angular momentum* of a particle about $O$ to be

$$\mathbf{L} = \mathbf{r} \wedge \mathbf{P}$$

Define the *moment* of the force, or the *torque*, to be

$$\begin{aligned} \mathbf{N} = \mathbf{r} \wedge \mathbf{F} &= \mathbf{r} \wedge \frac{\mathrm{d}}{\mathrm{d}t}(m\mathbf{v}) \\ &= \frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{r} \wedge m\mathbf{v}) \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{L} \end{aligned}$$

So if the total torque is zero the angular momentum is constant/conserved.
The work done by an external force upon a particle in going from A to B is

$$W_{AB} = \int_A^B \mathbf{F} \cdot \mathrm{d}\mathbf{s},$$

but $\mathbf{v} = \dot{\mathbf{s}}$ so

$$\begin{aligned} W_{AB} &= \int_{t_A}^{t_B} m \frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} \cdot \mathbf{v}\, \mathrm{d}t \\ &= \frac{m}{2} \int_{t_A}^{t_B} \frac{\mathrm{d}}{\mathrm{d}t} v^2\, \mathrm{d}t \\ &= \frac{m}{2}\left(v_A^2 - v_B^2\right) \end{aligned}$$

$\Rightarrow$ work done = change in kinetic energy.

**Definition:** If the force field is st. the work done is the same for any path then we have a *conservative system*.

This is true iff

$$\oint \mathbf{F} \cdot \mathrm{d}\mathbf{s} = 0$$

Which implies there exists a function $V(r)$ st. $\mathbf{F} = -\nabla V$.

$$\int_A^B \mathbf{F} \cdot \mathrm{d}\mathbf{s} = -V_B + V_A$$

which implies $W_{AB} = V_A - V_B = T_B - T_A \Rightarrow T_A + V_A = T_B + V_B$ ie. energy is conserved.

### 1.1.2 Systems of Particles

Suppose we have a system of $N$ particles. We distinguish between the external applied force and the internal forces between particles.
Newton's first two laws become

$$\frac{\mathrm{d}}{\mathrm{d}t}P_i = \sum_j \mathbf{F}_{ji} + \mathbf{F}_i^{\text{ext}} \tag{1.2}$$

(note $\mathbf{F}_{ii} = 0$.)
Applying NIII in the weak form means that the $\sum \mathbf{F}_{ij}$ term cancels.
Define the *centre of mass* by

$$\mathbf{R} = \frac{\sum_i m_i r_i}{\sum_i m_i} = \frac{\sum_i m_i r_i}{M}$$

Then

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\sum_i \mathbf{P}_i &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_i mi\mathbf{v}_i\right) \\
&= \frac{\mathrm{d}^2}{\mathrm{d}t^2}\left(\sum m_i r_i\right) \\
&= M\frac{\mathrm{d}^2\mathbf{R}}{\mathrm{d}t^2} \\
&= \sum_i \mathbf{F}_i^{\text{ext}} \\
\Rightarrow \mathbf{P} &= M\dot{R} \qquad (1.3)
\end{aligned}$$

So the momentum of the system is the same as the momentum of it's centre of mass.
Note that to get this result we have only required the weak form of NIII.
Now consider the total torque of the system

$$\begin{aligned}
\sum_i r_i \wedge \dot{\mathbf{P}}_i &= \sum_i \frac{\mathrm{d}}{\mathrm{d}t}\left(r_i \wedge \dot{\mathbf{P}}_i\right) \\
&= \dot{\mathbf{L}} \\
&= \sum_i r_i \wedge \mathbf{F}_i^{\text{ext}} + \sum_{ij} r_i \wedge \mathbf{F}_{ji}
\end{aligned}$$

But if we now assume the strong form of NIII

$$\begin{aligned}
r_i \wedge \mathbf{F}_{ji} + \mathbf{r} \wedge \mathbf{F}_{ij} &= (r_i - \mathbf{r}_j) \wedge \mathbf{F}_{ji} \\
&= 0 \qquad (1.4)
\end{aligned}$$

So

$$\mathbf{N}^{\text{ext}} = \frac{\mathrm{d}\mathbf{L}}{\mathrm{d}t} \tag{1.5}$$

This is the *conservation of angular momentum*
Define

$$\begin{aligned}
&\mathbf{r}_i' = r_i - \mathbf{R} \\
\Rightarrow \quad &\mathbf{v}_i = \mathbf{v}_i' + \mathbf{V}
\end{aligned}$$

ie. Working in the CoM frame. Now

$$\mathbf{L} = \sum_i \mathbf{R} \wedge m\mathbf{v}_i + \sum_i \mathbf{r}'_i \wedge m_i \mathbf{v}'_i + \left( \sum_i m_i \mathbf{r}'_i \right) \wedge \mathbf{v} + \mathbf{R} \wedge \frac{\mathrm{d}}{\mathrm{d}t} \left( \sum_i m_i \mathbf{r}'_i \right)$$

But the last two terms are zero.
Now $\sum m_i \mathbf{r}'_i$ define the radius vector of the centre of mass in a coord system with it's origin at the CoM. Ie.

$$\sum_i m_i \mathbf{r}'_i = 0$$

$$\begin{aligned} \mathbf{L} &= \sum_i \mathbf{R} \wedge m_i \mathbf{V} + \sum_i r_i \wedge m\mathbf{v}'_i \\ \text{or} \quad & \mathbf{R} \wedge M\mathbf{V} + \sum \mathbf{r}'_i \wedge m_i \mathbf{v}'_i \end{aligned}$$

The total angular momentum about a point $O$ is the angular momentum of the system concentrated at the CoM plus the angular momentum about the centre.

### 1.1.3 Energy

We wish to calculate the work done by all the forces in moving the system from an initial configuration $A$ to a final one $B$.

$$\begin{aligned} W_{AB} &= \sum_i \int_A^B \mathbf{F}_i \cdot \mathrm{d}\mathbf{s}_i \\ &= \sum_i \int_A^B \mathbf{F}^{\text{ext}} \cdot \mathrm{d}\mathbf{s} + \sum_{ij} \int_A^B \mathbf{F}_{ji} \cdot \mathrm{d}\mathbf{s}_i \end{aligned}$$

Recall that $\mathbf{F}_{ii} = 0$ then

$$\begin{aligned} W_{AB} &= \sum_i \int_A^B m_i \dot{\mathbf{v}}_i \cdot \mathbf{v}_i \, \mathrm{d}t \\ &= \sum_i \int_A^B \mathrm{d} \left( \frac{1}{2} m v_i^2 \right) \\ \Rightarrow W_{AB} &= T_B - T_A \end{aligned}$$

Now we want to transform to the CoM frame:

$$\begin{aligned} T &= \frac{1}{2} \sum_i m_i \left( \mathbf{V} + \mathbf{v}'_i \right) \cdot \left( \mathbf{V} + \mathbf{v}'_i \right) \\ &= \frac{1}{2} \sum_i m_i V^2 + \frac{1}{2} \sum_i m_i {v'}_i^2 \\ \Rightarrow T &= \frac{1}{2} M V^2 + \frac{1}{2} \sum_i m_i {v'}_i^2 \end{aligned}$$

If we can assume that the external forces can be derived from a potential energy function then the first term in ?? can be written

$$\begin{aligned}\sum_i \int_A^B \mathbf{F}_i^{\text{ext}} \cdot \, \mathrm{d}\mathbf{s}_i &= -\sum_i \int_A^B (\nabla_i \mathbf{V}_i) \cdot \, \mathrm{d}\mathbf{s} \\ &= -\sum_i \mathbf{V}_i \Big|_A^B\end{aligned}$$

Where we are using $\mathbf{V}$ for the potential.
If, now, the internal forces are conservative then the 'mutual' forces on the $i$th and $j$th particles can be obtained from a potential function $\mathbf{V}_{ij}$
If the strong form of the action-reaction law holds then $\mathbf{V}_{ij}$ can only be a function of the distance between the $i$th and $j$th particles,

$$\mathbf{V}_{ij} = \mathbf{V}_{ij}(|r_i - \mathbf{r}_j|)$$

$$\mathbf{F}_{ji} = -\nabla_i \mathbf{V}_{ij} = \nabla_j \mathbf{V}_{ij} = -\mathbf{F}_{ij} \tag{1.6}$$

If the $\mathbf{V}_{ij}$ were also functions of the difference of some other variable (e.g. velocity) then the forces would still be equal and opposite but not necessarily lie along the line connecting the two particles. If ?? holds then $\nabla_i \mathbf{V}_{ij} = (r_i - \mathbf{r}_j) f$ where $f$ is some other function.
Now when all the forces are conservative the second term in ?? becomes a sum over terms of the form

$$-\int_A^B (\nabla_i \mathbf{V}_{ij}) \cdot \, \mathrm{d}\mathbf{s}_i + \int_A^B (\nabla_j \mathbf{V}_{ij}) \cdot \, \mathrm{d}\mathbf{s}_j$$

But, in Cartesian coordinates $\mathrm{d}\mathbf{s}_i - \mathrm{d}\mathbf{s}_j = \mathrm{d}\mathbf{r}_i - \mathrm{d}\mathbf{r}_j \equiv \mathrm{d}\mathbf{r}_{ij}$, so the term for the $i$-$j$th pair is

$$\int (\nabla_{ij} \mathbf{V}_{ij}) \cdot \, \mathrm{d}\mathbf{r}_{ij}$$

Then the total work due to the internal forces is

$$-\frac{1}{2} \sum_{i,j,i \neq j} \int_A^B (\nabla_{ij} \mathbf{V}_{ij}) \cdot \, \mathrm{d}\mathbf{r}_{ij} = -\frac{1}{2} \sum_{i,j,i \neq j} \mathbf{V}_{ij} \Big|_A^B \tag{1.7}$$

(note the factor of $\frac{1}{2}$ is present because we're double counting in the indices.)
If both the external and the internal forces can be derived from potentials, and internal forces are radial, then we can define a total potential energy

$$\mathbf{V} = \sum_i \mathbf{V}_i + \frac{1}{2} \sum_{ij} \mathbf{V}_{ij} \tag{1.8}$$

such that the total energy is conserved.
The 2nd term in ?? is the internal potential energy of the system. In general $\mathbf{V}_{ij}$ is not constant and can change as the system changes with time. But a special case

is... **Definition:** A *rigid body* is a system of particles in which the distances $|\mathbf{r}_{ij}|$ are constant and cannot vary with time. ie

$$|\mathbf{r}_{ij}|^2 = |r_i - \mathbf{r}_j|^2 = c_{ij}^2, \text{ const } \forall i, j, t$$

But then

$$\frac{\mathrm{d}}{\mathrm{d}t} (\mathbf{r}_{ij})^2 = 0 \Rightarrow \mathbf{r}_{ij} \cdot \mathrm{d}\mathbf{r}_{ij} = 0$$

Thus if Newton's Law holds in the strong form: $\mathbf{F}_{ij} \cdot \mathrm{d}\mathbf{r}_{ij} = 0$, and internal forces do no work.

### 1.1.4 Constraints

Take as examples:

- Rigid bodies
- Gas molecules in a container
- A particle moving on a solid sphere

**Definition:**If the condition of constraint is such that it can be written in the form $f(\mathbf{r}_1, ..., \mathbf{r}_n, t) = 0$ then we have a *holonomic* constraint. An example is the rigid body. Constraints which cannot be written this way are called *non-holonomic*, eg a gas in a container. If the constraint contains time explicitly then it is said to be *rheonomous*, if it does not it is called *scleronomous*, eg. bead on rigid wire is subject to the latter type of constraint, but if the bead is on a moving wire then we have the former type.

Note that this means a holonomic constraint must allow us to eliminate some variables

Very often the constraint can be written $\sum_i g_i(x_1, ..., x_n)\, \mathrm{d}x_i = 0$ then the constraint will be holonomic. If an integrating function exists $f(x_1, ..x_n)$ such that $g_i = \frac{\partial f}{\partial x_i}$ ie the constraint is holonomic only if

$$\frac{\partial f g_i}{\partial x_j} = \frac{\partial f}{\partial x_j} g_i + f \frac{\partial g_i}{\partial x_j} = g_j g_i + f \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial f g_j}{\partial x_i} \tag{1.9}$$

**Example:**The Rolling Disc.

Other examples of non-holonomic constraints are a particle on the sphere, and all constraints depending on higher derivatives

## 1.2 D'Alembert's Principle and Lagrange's Equations

### 1.2.1 Two preliminary lemmas

**The cancellation of the dots** If we have a function $x = x(q_i, \dot{q}_i)$ then

$$\dot{x} = \sum_i \frac{\partial x}{\partial q_i} \dot{q}_i$$

So that

$$\frac{\partial \dot{x}}{\partial \dot{q}_i} = \frac{\partial x}{\partial q_i}$$

**The interchange of the d and the ∂**

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial x}{\partial q_i}\right) = \sum_j \frac{\partial^2 x}{\partial q_i \partial q_j}\dot{q}_j = \frac{\partial}{\partial \dot{q}_i}\sum_j \frac{\partial \dot{x}}{\partial q_j}\dot{q}_j = \frac{\partial \ddot{x}}{\partial \dot{q}_i}$$

So that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial x}{\partial q_i}\right) = \frac{\partial \dot{x}}{\partial q_i}$$

by the cancellation of the dots.

### 1.2.2 D'Alembert's Principle

**Definition:** A *virtual displacement* of a system refers to a change in the configuration of the system as a result of an arbitrary infinitesimal displacement $\delta r_i$ consistent with the forces and constraints at time $t$. The displacement is called virtual so as to distinguish it from an actual displacement occurring in time during which the forces and constraints can vary.
Suppose the system is in equilibrium, ie. the total force on each particle is zero, $\mathbf{F}_i = 0$. Then clearly

$$\mathbf{F}_i \cdot \delta r_i = 0$$

so as not to affect constraints and forces. If we decompose $\mathbf{F}$ as $\mathbf{F}_i = \mathbf{F}_i^{\text{ext}} + \mathbf{f}_i$ then

$$\sum_i \mathbf{F}_i^{\text{ext}} \cdot \delta \mathbf{r}_i + \sum_i \mathbf{f}_i^{\text{ext}} \cdot \delta \mathbf{r}_i = 0$$

We now make the assumption that constraint forces do no work (ie the second term is zero) under the virtual displacement. Ie we assume we have a rigid body. Then

$$\sum_i \mathbf{F}_i^{\text{ext}} \cdot \delta r_i = 0 \tag{1.10}$$

This is the *Principle of Virtual Work*, or what some authors call *D'Alemberts Principle*,[1] ie. The condition for he equilibrium of a system is that the virtual work of the applied forces is zero..
Consider a system described by $n$ generalised coordinates. Let us assume al constraints are holonomic. We remark that $\{q_i\}$ may be less in number than the total number $3N$ of degrees of freedom of the system (constraints).
Now the work can be done in an infinitesimal displacement will be proportional to the elements $\mathrm{d}q_i$,

$$\mathrm{d}W = \sum_r Q_r \, \mathrm{d}q_r,$$

[1] We shall reserve this for a later result

$Q_r$ is then defined as the *generalised force*.
Consider now a system of $N$ particles, let $\mathbf{F}_i$ be the force on the $i$th particle, let $\mathbf{P}_i$ be its momentum. From Newton

$$\mathbf{F}_i - \dot{\mathbf{P}}_i = 0$$

$\Rightarrow$

$$\sum_i \left(\mathbf{F}_i - \dot{\mathbf{P}}_i\right) \cdot \delta r_i = 0$$

where $\delta \mathbf{r}_i$ is a virtual displacement. But $\mathbf{F}_i = \mathbf{F}_i^{\text{ext}} + \mathbf{f}_i$ so

$$\sum_i \left(\mathbf{F}_i^{\text{ext}} + \mathbf{f}_i - \dot{\mathbf{P}}_i\right) \cdot \delta \mathbf{r}_i = 0$$

We make the assumption that forces of constraint do no work , ie $\sum \mathbf{f}_i \cdot \delta r_i = 0$ and we obtain

$$\sum_i \left(\mathbf{F}_i^{\text{ext}} - \dot{\mathbf{P}}_i\right) \cdot \delta \mathbf{r}_i = 0, \tag{1.11}$$

what we shall call D'Alemberts Principle - this is the dynamic principle of virtual work.

### 1.2.3 Lagrange's Equations

continuing from above:

$$\begin{aligned}
\sum_i \dot{\mathbf{P}}_i \cdot \delta \mathbf{r}_i &= \sum_i m_i \ddot{r}_i \cdot \delta \mathbf{r}_i \\
&= \sum_{ij} m_i \ddot{\mathbf{r}}_i \cdot \left(\frac{\partial r_i}{\partial q_j} \cdot \delta q_j\right) \\
&= \sum_j \left(\sum_i \frac{\mathrm{d}}{\mathrm{d}t}\left(m_i \dot{r}_i \frac{\partial r_i}{\partial q_j} - m_i \dot{\mathbf{r}}_i \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathbf{r_i}}{\partial q_j}\right)\right)\right) \delta q_i
\end{aligned}$$

But as we've seen

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial r_i}{\partial q_j}\right) = \frac{\partial \mathbf{v}_i}{\partial q_j}, \quad \text{and} \quad \frac{\mathrm{d}\mathbf{v}_i}{\mathrm{d}\dot{q}_j} = \frac{\mathrm{d}r_i}{\mathrm{d}q_j}$$

So

$$\begin{aligned}
\sum_i m_i \ddot{\mathbf{r}}_i \frac{\partial r_i}{\partial q_j} &= \sum_i \left[\frac{\mathrm{d}}{\mathrm{d}t}\left(m \mathbf{v}_i \frac{\partial \mathbf{v}_i}{\partial \dot{q}_j}\right) - m_i \mathbf{v}_i \frac{\partial \mathbf{v}_i}{\partial q_j}\right] \\
&= \sum_j \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial}{\partial \dot{q}_j}\left(\sum_i \frac{1}{2} m_i v_i^2\right)\right) - \frac{\partial}{\partial q_j}\left(\sum_i \frac{1}{2} m_i v_i^2\right) \delta q_j
\end{aligned}$$

Now let us make use of the fact that we have holonomic constraints - we can define our coordinates $\{q_i\}$ such that they form a complete set

$$r_i = r_i(q_1, ...q_n, t)$$

So

$$\mathbf{v}_i = \frac{\mathrm{d}r_i}{\mathrm{d}t} = \sum_k \frac{\partial r_i}{\partial q_k}\frac{\mathrm{d}q_k}{\mathrm{d}t} + \frac{\partial r_i}{\partial t}$$

Hence $\delta \mathbf{r}_i = \sum_j \frac{\partial r_i}{\partial q_j}\delta q_j$ since $\delta r_i$ is indept of time. So we have now that:

$$\sum_i \mathbf{F}_i^{\text{ext}} \cdot \delta r_i = \sum_{ij} \mathbf{F}_i^{\text{ext}} \cdot \frac{\partial r_i}{\partial q_j}\delta q_j$$

We now define the *Generalised Forces* corresponding to our generalised coords as

$$Q_j = \sum_i \mathbf{F}_i^{\text{ext}} \cdot \frac{\partial r_i}{\partial q_j}$$

So using $T = \frac{1}{2}\sum_i v_i^2 - KE$ we an write

$$\sum_{ij}\left[\left\{\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\mathrm{d}T}{\mathrm{d}q_j}\right\} - Q_j\right]\delta q_j = 0$$

which is just D'Alemberts principle again!
Since this s true for any virtual displacement and the $q_j$s are independent (holonomic) $\Rightarrow$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} = Q_j \tag{1.12}$$

Assume we are dealing with a conservative system ie that $\mathbf{F}_i^{\text{ext}} = -\nabla_i \mathbf{V}$ then

$$Q_j = \sum_i \mathbf{F}_i^{\text{ext}} \cdot \frac{\partial r_i}{\partial q_j} = -\sum_i (\nabla_i \mathbf{V}) \cdot \frac{\partial r_i}{\partial q_j}$$

ie $Q_j = -\frac{\partial \mathbf{V}}{\partial q_j}$ so substitution into equation ?? $\Rightarrow$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial}{\partial q_j}(T - \mathbf{V}) = 0$$

But we know that $\frac{\partial \mathbf{V}}{\partial \dot{q}_j} = 0$ so define the *Lagrangian*, $\mathsf{L}$ by

$$\mathsf{L} = T - \mathbf{V}$$

then

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathsf{L}}{\partial \dot{q}_i}\right) - \frac{\partial \mathsf{L}}{\partial q_i} = 0}$$

**Note 1.**
- *This is set of $n$ second order* odes*, the $\partial$'s are there only as part of the notation*
- *The solution of LEs will involve finding $2n$ functions eg. at $t = 0$ $q_\alpha(0) = A_\alpha,\ \dot{q}_\alpha = B_\alpha$*

**Example:**The Spherical Pendulum a particle of mass $m$ which moves under gravity is attached to a fixed point by a rod of length $a$. $\Rightarrow$ It is a particle constrained to move on a sphere of radius $a$. In terms of spherical coords $(a, \theta, \phi)$ with $\theta$ measured upwards from the downward direction the kinetic and potential energies can be written:

$$\begin{aligned} T &= \frac{1}{2} m a^2 \left( \dot{\theta}^2 + \sin^2 \theta \dot{\phi}^2 \right) \\ \mathbf{V} &= -mga \cos \theta \end{aligned}$$

So

$$\begin{aligned} & \mathsf{L} = T - \mathbf{V} \\ \Rightarrow \quad & \frac{\partial \mathsf{L}}{\partial \dot{\theta}} = m a^2 \dot{\theta}, \qquad \frac{\partial \mathsf{L}}{\partial \dot{\phi}} = m a^a \sin^2 \theta \dot{\phi} \\ \text{and} \quad & \frac{\partial \mathsf{L}}{\partial \theta} = m a^2 \sin \theta \cos \theta \dot{\phi}^2 - mga \sin \theta, \quad \frac{\partial \mathsf{L}}{\partial \phi} = 0 \end{aligned}$$

So by LEs

$$\begin{aligned} & m a^2 \ddot{\theta} - m a^2 \sin \theta \cos \theta \dot{\phi} + mga \sin \theta = 0 \\ \text{and} \quad & m a^2 \frac{\mathrm{d}}{\mathrm{d}t} \left( \sin^2 \theta \dot{\phi} \right) = 0 \\ \Rightarrow \quad & \sin^2 \theta \dot{\phi} = const \end{aligned}$$

## 1.3 Generalisations of Lagrange's Equations

### 1.3.1 Velocity Dependent Potentials

Suppose there exists a potential $U(q_i, \dot{q}_i)$ such that

$$Q_j = -\frac{\partial U}{\partial q_j} + \frac{\mathrm{d}}{\mathrm{d}t} \left( \frac{\partial U}{\partial \dot{q}_j} \right)$$

then we would be able to define a Lagrangian $\mathsf{L} = T - U$ and the form of LEs would be unaltered. $U$ will be called a *generalised* or *velocity dependent potential*
**Example:**Maxwell's Equations[2]

$$\begin{aligned} \nabla \wedge \mathbf{E} + \frac{1}{c^2} \dot{\mathbf{B}} &= 0 \\ \nabla \cdot \mathbf{D} &= 4\pi\rho \\ \nabla \wedge \mathbf{H} - \frac{1}{c^2} \dot{\mathbf{D}} &= 0 \\ \nabla \cdot \mathbf{B} &= 0 \end{aligned}$$

The force on a charge $q$ is not simply

$$\mathbf{F} = q\mathbf{E} = -\nabla \phi$$

[2]Presented here using the auxiliary fields $\mathbf{D}$ and $\mathbf{H}$ and in Gaussian units: So for those who attended Electromagnetism we have $\mathbf{B} = \mathbf{H} - 4\pi\mathbf{M}$ and $\mathbf{E} = \mathbf{D} - 4\pi\mathbf{P}$

but rather

$$\mathbf{F} = q\left(\mathbf{E} + \frac{1}{c}\mathbf{v} \wedge \mathbf{B}\right)$$

$\mathbf{E}$ is *not* the gradient of a scalar function: $\nabla \cdot \mathbf{B} = 0 \Rightarrow \exists \mathbf{A}$ st. $\mathbf{B} = \nabla \wedge \mathbf{A}$. $\mathbf{A}$ is called the *Magnetic Vector Potential*
We can write ?? as

$$\nabla \wedge \mathbf{E} + \frac{1}{c}\frac{\partial}{\partial t}(\nabla \wedge \mathbf{A}) = \nabla \wedge \left(\mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}\right) = 0$$

If we now set $\mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} = -\nabla\phi$ this becomes

$$\mathbf{F} = q\left\{-\nabla\phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} + \frac{1}{c}(\mathbf{v} \wedge [\nabla \wedge \mathbf{A}])\right\}$$

Consider:

$$\begin{aligned}(\mathbf{v}(\nabla \wedge \mathbf{A}))_x &= v_y\left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right) - v_z\left(\frac{\partial A_x}{\partial x} - \frac{\partial A_z}{\partial x}\right) \\ &= v_y\frac{\partial A_y}{\partial x} + v_z\frac{\partial A_z}{\partial x} + v_x\frac{\partial A_x}{\partial x} - v_y\frac{\partial A_x}{\partial y} - v_z\frac{\partial A_x}{\partial z} - v_x\frac{\partial A_z}{\partial x}\end{aligned}$$

Now we note that

$$\frac{\mathrm{d}A_x}{\mathrm{d}t} = \frac{\partial A_x}{\partial t} + v_x\frac{\partial A_x}{\partial x} + v_y\frac{\partial A_y}{\partial y} + v_z\frac{\partial A_z}{\partial z}$$

So that

$$\mathbf{v} \wedge (\nabla \wedge \mathbf{A}) = \frac{\partial}{\partial x}(\mathbf{v} \cdot \mathbf{A}) - \frac{\mathrm{d}A_x}{\mathrm{d}t} + \frac{\partial A_x}{\partial t}$$

$\Rightarrow$

$$\begin{aligned}F_x &= q\left(-\frac{\partial}{\partial x}\left(\phi - \frac{1}{c}\mathbf{A} \cdot \mathbf{v}\right) - \frac{1}{c}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial}{\partial v_x}(\mathbf{A} \cdot \mathbf{v})\right)\right) \\ &= -\frac{\partial u}{\partial x} + \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial u}{\partial \dot{x}}\right)\end{aligned}$$

$$U = q\phi - \frac{q}{c}\mathbf{A} \cdot \mathbf{v}$$

So we define

$$\mathsf{L} - T - U$$

Suppose now that *not* all the forces are derivable from a potential. We can retain Lagrange's Equations in the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathsf{L}}{\partial \dot{q}_j}\right) - \frac{\partial \mathsf{L}}{\partial q_j} = \tilde{Q}_j$$

Where $\mathsf{L}$ contains the potential of the conservative and velocity dependent forces and $\tilde{Q}_j$ represents those forces which cannot be derived from a potential of either type.

Suppose now that we have a frictional force, $\mathbf{F}$[3] which is proportional to velocity

$$F_x = -k_x v_x$$

Define

$$\mathcal{F}_i = \frac{1}{2}\sum\left(k_x v_{ix}^2 + k_y v_{iy}^2 + k_z v_{iz}^2\right)$$

(This is known as Rayleigh's Function)
Then $F_{xi} = -\frac{\partial \mathsf{F}}{\partial v_x}$, or $\mathbf{F}_i = -\nabla \mathcal{F}_i$, and the work done by the $i$th particle against friction is

$$\mathrm{d}W_i = -\mathbf{F}_i \cdot \mathrm{d}\mathbf{r}_i = \mathbf{k} \cdot \mathbf{v}_i^2 \,\mathrm{d}t$$

The component of the generalised force is

$$\tilde{Q}_j = \sum_i \mathbf{F}_i \cdot \frac{\partial r_i}{\partial q_j}$$

And LEs become

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathsf{L}}{\partial dotq_j}\right) - \frac{\partial \mathsf{L}}{\partial q_j} + \frac{\partial \mathcal{F}}{\partial q_j} = 0$$

### 1.3.2 Constraints

Now let us look at a system which may be rheonomic (time dependent constraints), non-conservative and non-holonomic.
Consider a system of $N$ particles with masses $m_i$ and positions $r_i$, and accelerations $\mathbf{a}_i$. Let $\mathbf{F}_i = m_i \mathbf{a}_i$ or $\sum_1^N (m_i \mathbf{a}_i - \mathbf{F}_i) = 0$. for any virtual displacement we have

$$\sum_1^N (m_i \mathbf{a}_i - \mathbf{F}_i) \cdot \delta r_i = 0$$

Define

$$\delta W = \sum_1^N \mathbf{F} \cdot \delta \mathbf{r}_i$$

Let us consider a maximal set $\{q_\alpha\}_{\alpha=1}$, where all holonomic constraints have been absorbed.
Suppose the more general (non-holonomic) constraints can be written

$$\sum_i^n A_{\beta\alpha}(q_i, t)\dot{q}_\alpha + A_\beta(q_i, t) = 0 \tag{1.13}$$

where $\beta = 1, .., m$. Ie. there are $n < N$ constraints.[4]
We can write ?? as

$$\sum_{\alpha=1}^n A_{\alpha\beta}(q, t)\,\mathrm{d}q_\alpha + A_\beta(q, t)\,\mathrm{d}t = 0$$

[3]Notation: For this section $F$ is the frictional force, $i$ is the particle number, and $x$ is the direction
[4]?? Notation here ??

But the kinetic energy of the system is $T = \frac{1}{2}\sum_{i=1}^{N} m_i \dot{r}_i \cdot \dot{r}_i$, so

$$\dot{r}_i = \sum_{\alpha=1}^{n} \frac{\partial r_i}{\partial q_\alpha}\dot{q}_\alpha + \frac{\partial r_i}{\partial t}$$

Define $S_\alpha$ by

$$S_\alpha = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{q}_\alpha}\right) - \frac{\partial T}{\partial q_\alpha}$$

(Remember cancellation of the dots and the interchange of the $\mathrm{d}$ and the $\partial$).
$\Rightarrow$

$$S_\alpha = \sum_{i=1}^{N} m_i \mathbf{a}_i \cdot \frac{\partial r_i}{\partial q_\alpha}$$

Now let $\delta q_\alpha$ satisfy

$$\sum_{\alpha=1}^{n} A_{\beta\alpha}\delta q_\alpha = 0, \quad , \beta = 1, ..., m$$

**Remark 1.** *We are dealing with virtual displacements, ie. terms in* $\mathrm{d}t$ *are lost.*

$\delta r_i$ are virtual displacements satisfying the constraints

$$\sum_{\alpha=1}^{n} S_\alpha \delta q_\alpha = \delta W$$

We may write

$$\begin{aligned} \delta W &= \sum_\alpha Q_\alpha \delta q_\alpha \\ \Rightarrow \quad & \sum_{\alpha=1}^{n} (S_\alpha - Q_\alpha)\,\delta q_\alpha = 0 \\ \not\Rightarrow \quad & S_\alpha = Q_\alpha \end{aligned}$$

because the $\delta q_\alpha$ are not independent, but are subject to

$$\sum_{\alpha=1}^{n} A_{\beta\alpha}\delta q_\alpha = 0, \quad \beta = 1, ..., m.$$

Now write $S_\alpha - Q_\alpha = B_\alpha$ and define

$$\begin{aligned} \mathsf{F} \;=\; & (B_1 - \lambda_1 A_{11} - \lambda_2 A_{21} - ... - \lambda_m A_{m1})\,\delta q_1 \\ + \; & (B_2 - \lambda_1 A_{12} - \lambda_2 A_{22} - ... - \lambda_m A_{m2})\,\delta q_2 \\ & \vdots \\ + \; & (B_n - \lambda_1 A_{1n} - \lambda_2 A_{2n} - ... - \lambda_m A_{mn})\,\delta q_n \end{aligned}$$

(The $\lambda$ are Lagrange multipliers, and as such are arbitrary.)

**Note 2.** $\mathsf{F} = 0 \;\; \forall \delta q_\alpha$ *satisfying* $\sum_{\alpha=1}^{n} B_\alpha \delta q_\alpha = 0, \;\; \sum_{\alpha=1}^{n} A_{\beta\alpha} \delta q_\alpha = 0$

We have $m$ arbitrary $\lambda$s, we chose these so that

$$\begin{aligned} B_1 &= \lambda_1 A_{11} + \lambda_2 A_{21} + ... + \lambda_m A_{m1} \\ &\vdots \\ B_m &= \lambda_1 A_{1m} + \lambda_2 A_{2m} + ... + \lambda_m A_{nm} \end{aligned}$$

Then

$$\begin{aligned} \mathsf{F} &= \left(B_{m+1} - \lambda_1 A_{1,m+1} - \lambda_2 A_{2m+1} - ... - \lambda_m A_{m,m+1}\right) \delta q_{m+1} \\ &\vdots \\ &+ \left(B_n - \lambda_1 A_{1n} - \lambda_2 A_{2n} - ... - \lambda_m A_{mn}\right) \delta q_n \end{aligned}$$

$\Rightarrow \mathsf{F} = 0$ since each column is zero by the constraints. Hence the equations of motion are:

$$\sum_{\alpha=1}^{n} A_{\beta\alpha} \dot{q}_\alpha + A_\beta = 0, \qquad \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{q}_\alpha}\right) - \frac{\partial T}{\partial q_\alpha} = Q_\alpha + \sum_{\beta}^{m} \lambda_\beta A_{\beta\alpha}$$

This will work for holonomic constraints and for many non-holonomic constraints.

**Example:** The Rolling Loop

Consider a loop rolling without slipping down an inclined pane. This is actually a holonomic constraint, but it will still serve to illustrate the principle.

The constraint is

$$r\,\mathrm{d}\theta = \,\mathrm{d}x,$$

Also it is clear that

$$V = mg(l - x)\sin\phi \tag{1.14}$$

$$T = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}mr^2\dot{\theta}^2 \tag{1.15}$$

$$\Rightarrow \quad \mathsf{L} = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}mr^2\dot{\theta}^2 - mg\left(()\, l - x\right)\sin\phi \tag{1.16}$$

One constraint $\Rightarrow$ one Lagrange Multiplier.
The constraint is of the form

$$\sum_{\alpha=1}^{n} A_{1\alpha} \dot{q}_\alpha = 0$$

with $A_{1\theta} = r, \;\; A_{1x} = -1$ So Lagrange's Equations $\Rightarrow$

$$m\ddot{x} - mg\sin\phi + \lambda = 0 \tag{1.17}$$

$$mr^2\ddot{\theta} - \lambda r = 0 \tag{1.18}$$

$$r\dot{\theta} = \dot{x} \tag{1.19}$$

where the last equation is the constraint. We have 3 equations for 3 unknowns, $\theta, x, \lambda \Rightarrow$

$$\frac{\mathrm{d}}{\mathrm{d}t}\,(\text{constraint}) \tag{1.20}$$

$$\Rightarrow \quad r\ddot{\theta} = \ddot{x} \Rightarrow m\ddot{x} = \lambda \tag{1.21}$$

$$\Rightarrow \quad \ddot{x} = \frac{g\sin\phi}{2} \text{ and } \lambda = \frac{mg\sin\phi}{2} \text{ and } \ddot{\theta} = \frac{g\sin\phi}{2r} \tag{1.22}$$

$\Rightarrow$ hoop rolls down the plane with half the acceleration it would have if it slipped down a frictionless plane. The force constraint is $\lambda = \frac{mg\sin\phi}{2}$, and notice that the force of constraint appears via the Lagrange Multiplier. **Example:**Atwood's Machine We have two masses and a frictionless massless pulley. There is only one independent coordinate, $x$, the position of the second weight determined by the constraint that the string has length $l$.

$$\begin{aligned} \mathbf{V} &= -m_1 g x - m_2 g(l-x) \\ T &= \frac{1}{2}\left(m_1+m_2\right)\dot{x}^2 \end{aligned}$$

So

$$\begin{aligned} & \mathsf{L} = T - \mathbf{V} = \frac{1}{2}\left(m_1+m_2\right)\dot{x}^2 + m_1 g x + m_2 g(l-x) \\ \Rightarrow \quad & \frac{\partial \mathsf{L}}{\partial x} = \left(m_1 - m_2\right) g \\ \text{and} \quad & \frac{\partial \mathsf{L}}{\partial \dot{x}} = \left(m_1+m_2\right)\dot{x} \\ \text{LEs} \Rightarrow \quad & \left(m_1+m_2\right)\ddot{x} = \left(m_1-m_2\right) g \\ \text{or} \quad & \ddot{x} = \left(\frac{m_1-m_2}{m_1+m_2}\right) g \end{aligned}$$

Notice

- The force of constraint (tension in the string) appears *nowhere*. We don't need to say anything about it to find the equations of motion
- We can't deduce anything about it. ie. we cannot determine the tension.

**Example:**Motion on the sphere
A particle of mass $m$ moving under gravity on a smooth sphere of radius $b$.
The constraint is $x^2+y^2+z^2=b^2$ (this is actually holonomic) $\Rightarrow x\dot{x}+y\dot{y}+z\dot{z}=0$.
We also have that

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

Define the generalised forces $\tilde{X}=0,\ \tilde{Y}=0,\ \tilde{Z}=-mg$, then the equations of motion are

$$\begin{aligned} x\dot{x}+y\dot{y}+z\dot{z} &= 0, && (1.23) \\ m\ddot{x} &= \lambda x, && (1.24) \\ m\ddot{y} &= \lambda y, && (1.25) \\ m\ddot{z} &= -mg+\lambda z && (1.26) \end{aligned}$$

### 1.3.3 Lagrange's Equations for Impulsive Forces

Consider any dynamical system which moves according to Lagrange's Equations:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{q}_\alpha}\right) - \frac{\partial T}{\partial q_\alpha} = Q_\alpha$$

Now integrating w.r.t. $t$ we have:

$$\left[\frac{\partial T}{\partial \dot{q}_\alpha}\right]_{t_1}^{t_2} - \int_{t_1}^{t_2} \frac{\partial T}{\partial q_\alpha}\,\mathrm{d}t = \int_{t_1}^{t_2} Q_\alpha\,\mathrm{d}t$$

Think Dirac delta-function

Let us assume that as $t_1 \to t_2$, $Q_\alpha \to \infty$ such that $\hat{Q}_\alpha = \lim_{t_1\to t_2}\int_{t_1}^{t_2} Q_\alpha\,\mathrm{d}t$ remains finite. Then we call the $\hat{Q}_\alpha$ the *generalised impulsive forces*.
In the infinitesimal interval $|t_1 - t_2|$ we assume the generalised coordinate don't change and the generalised velocities remain finite.Then we write

$$\Delta\left[\frac{\partial T}{\partial \dot{q}_\alpha}\right] = \hat{Q}_\alpha$$

### 1.3.4 Some Definitions

**Definition:**For a holonomic system

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{q}_i}\right) = \frac{\partial T}{\partial q_i} + \tilde{Q}_i$$

The $\frac{\partial T}{\partial q_i}$ are sometimes called *fictitious forces* due to our change of coordinates.

**Note 3.** *These are different from the fictitious forces introduced to make a non-inertial frame appear inertial*

**Definition:** The instantaneous configuration of a system can be described by the $n$ generalised coordinates $q_1, .., q_n$. This corresponds to a particular point in the Cartesian hyperspace where the $q$'s form the coordinate axes. This $n$-dimensional space is called *configuration space* or *coordinate space*.

**Note 4.** *As time goes on the system point moves in configuration space tracing out the* path of the system. *Configuration space is not necessarily the same as physical space*

**Definition:** We say a system is *monogenic* if all the forces (except the forces of constraint) are derivable from a generalised potential which may be a function of the (generalised) coordinates, the (generalised) velocities and the time. For such a system we have...

## 1.4 Hamilton's Principle

### 1.4.1 Hamilton's Principle

The motion of the system from time $t_1$ to $t_2$ is such that the integral

$$I = \int_{t_1}^{t_2} \mathsf{L}(q_i, \dot{q}_i, t)\,\mathrm{d}t,$$

has a stationary value. We write this as $\delta I = 0$.

This is where we started the course: The Variational Principle. Only now we're in *coordinate space*, and we're seeking only a stationary value, not necessarily a minimum.

**Remarks 1.** *1. We've seen that if* $\mathsf{L} = T - U$ *and we have holonomic constraints then*

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathsf{L}}{\partial \dot{q}_i}\right) - \frac{\partial \mathsf{L}}{\partial q_i} = 0$$

$\Rightarrow$ *Newton's Laws, and also Maxwell's Equations.*

*2. Recall that*

$$\delta \int f(y_i, y_i', x)\,\mathrm{d}x = 0 \iff \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y_i'}\right) - \frac{\partial f}{\partial y_i} = 0$$

ie: The variational principle $\iff$ Lagrange's Equations. (assuming the $q_i$'s are independent).
If we have non-holonomic constraints then

$$\sum_\alpha A_{\beta\alpha}\,\mathrm{d}q_\alpha + A_\beta\,\mathrm{d}t = 0, \quad \beta = 1, ..., m$$

And that using Lagrange multipliers we can write

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathsf{L}}{\partial \dot{q}_\alpha}\right) - \frac{\partial \mathsf{L}}{\partial q_\alpha} = \sum_\beta \lambda_\beta A_{\beta\alpha} \tag{1.27}$$

$$\sum_\alpha A_{\alpha\beta}\dot{q}_\alpha + A_\beta = 0 \tag{1.28}$$

ie. $n+m$ equations for $n+m$ variables, $\{q_i\}_i^n,\ \{\lambda_i\}_1^m$.

**Note 5.** *From now on we shall assume the holonomicy of the constraints*

### 1.4.2 Conservation Laws and Symmetries

If $\mathsf{L}$ is not a function of a given $q_i$ then

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathsf{L}}{\partial \dot{q}_i}\right) = 0 \Rightarrow \frac{\partial \mathsf{L}}{\partial \dot{q}_i} = A.$$

That is, $A$ is a constant of the motion, it is *conserved*. eg. Suppose $\mathsf{L} = \frac{1}{2}m\dot{r}_i^2$ then $\frac{\partial \mathsf{L}}{\partial \dot{q}_i} = \frac{\mathrm{d}\mathsf{L}}{\mathrm{d}\dot{x}_i} = m\dot{x}_i$ and because $\frac{\partial \mathsf{L}}{\partial x_i} = 0$ we have $m\dot{x} = const$ Ie Newton's First Law, that momentum is conserved. We say that momentum is a *conjuagte variable* to position.
**Definition:** It will be convenient to define *generalized momenta* by

$$p_i = \frac{\partial \mathsf{L}}{\partial \dot{q}_i}.$$

Note that $p_i$ need not be an ordinary momentum. Suppose that the Lagrangian is independent of time

$$\frac{\mathrm{d}\mathsf{L}}{\mathrm{d}t} = \frac{\partial \mathsf{L}}{\partial t} + \sum_i \dot{q}_i \frac{\partial \mathsf{L}}{\partial q_i} + \sum_i \frac{\partial \mathsf{L}}{\partial \dot{q}_i}\frac{\mathrm{d}\dot{q}_i}{\mathrm{d}t} \tag{1.29}$$

$$= \sum_i \frac{\mathrm{d}}{\mathrm{d}t}\left(\dot{q}_i\left(\frac{\partial \mathsf{L}}{\partial \dot{q}_i}\right)\right) \tag{1.30}$$

$$\Rightarrow \left(\sum_i \dot{q}_i \frac{\partial \mathsf{L}}{\partial \dot{q} i} - \mathsf{L}\right) = const \tag{1.31}$$

So we define

$$\mathsf{H} = \sum_i P_i \dot{q}_i - \mathsf{L}$$

(in Cartesian coordinates $\mathsf{H} = \sum_i m_i \dot{r}_i^2 - T - V = T + V =$ energy).
Consider a generalised coordinate $q_j$ for which a change in $\mathrm{d}q_j$ represents a translation of the entire system (eg. $q_j$ is the CoM). Now $v_i$ is claerly idependent of the origin of coordinates $\Rightarrow \frac{\partial v_i}{\partial q_j} = 0$ and hence $\frac{\partial T}{\partial q_j} = 0$. Suppose also that we have a conservative system, ie

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{q}_j}\right) = \dot{p}_j = -\frac{\partial V}{\partial q_j} = Q_j = \sum_i \frac{\partial r_i}{\partial q_j} \cdot \mathbf{F}_i$$

Now considering the effect of the infinitesimal $\mathrm{d}q_j$ (translation of the system along some axis)

$$r_i(q_j) \longrightarrow r_i(q_j + \mathrm{d}q_j)$$

$$\frac{\partial r_i}{\partial q_j} \equiv \lim_{\mathrm{d}q_j \to 0} \frac{r_i(q_j + \mathrm{d}q_j) - r_i(q_j)}{\mathrm{d}q_j} \tag{1.32}$$

$$= \frac{\mathrm{d}\hat{\mathbf{n}}}{\mathrm{d}q_j}. \tag{1.33}$$

Where $\hat{\mathbf{n}}$ is a unit vector in the direction of the translation.

$$\tilde{Q}_j = \sum_i \frac{\partial r_i}{\partial q_j} \cdot \mathbf{F}_i = \sum_i \hat{\mathbf{n}} \cdot \mathbf{F}_i = \hat{\mathbf{n}} \cdot \mathbf{F}$$

Now suppose that $q_j$ does not appear in $V$ (and hence in $\mathsf{L}$). Then

$$T = \frac{1}{2} \sum_i m_i \dot{r}_i^2$$

and

$$p_j = \sum_i m_i \dot{r}_i \cdot \frac{\partial r_i}{\partial \dot{q}_j} \tag{1.34}$$

$$= \sum_i m_i \dot{r}_i \cdot \frac{\partial r_i}{\partial \dot{q}_j} \tag{1.35}$$

$$= \sum_i m_i \cdot r_i \cdot \hat{\mathbf{n}} \tag{1.36}$$

$$= \hat{\mathbf{n}} \cdot \left(\sum_i m_i \mathbf{v}_i\right) \tag{1.37}$$

Now since $q_j$ is not in $\mathsf{L} \Rightarrow$

$$\tilde{Q}_j = 0 \Rightarrow \mathbf{F} \cdot \hat{\mathbf{n}} = 0 \Rightarrow \hat{\mathbf{n}} \cdot \sum_i m_i \mathbf{v}_i = const$$

arned however: *Never*
*e an ignorable coor-*
*e*

A variable in generalised coordinates which does *not* appear in the Lagrangian is said to by *cyclic* or *ignorable*.
We have seen that if the Lagrangian is independent of translation in a given direction $\hat{\mathbf{n}}$ there is no force in this direction and the momentum component is conserved.
Suppose $q_j$ is a cyclic coordinate, and $\mathrm{d}q_j$ corresponds to a rotation of the system about some axis. Now, just as before, we will argue that a rotation of the coordinate system cannot affect the magnitude of the velocities ie. $\frac{\partial T}{\partial q_j} = 0$. We are assuming that $q_j$ is ignorable so $\frac{\partial \mathsf{L}}{\partial q_j} = 0$, hence $\frac{\partial V}{\partial q_j} = 0$.

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{q}_j}\right) = p_j = -\frac{\partial V}{\partial q_j}$$

Now the deriivative has a different meaning.

The change $\mathrm{d}q_j$ must correspond to an infinitesimal rotation keeping the magnitude of $r_i$ fixed ie

$$|\mathbf{r}(q_j)| = |r_i(q_j + \mathrm{d}q_j)|$$

So

$$|\,\mathrm{d}r_i| = |r_i \sin\theta \, \mathrm{d}q_j| \tag{1.38}$$

$$\Rightarrow \quad \left|\frac{\partial r_i}{\partial q_j}\right| = |r_i \sin\theta| \tag{1.39}$$

Let $\hat{\mathbf{n}}$ indicate a unit vector defining the axis about which we rotate.

$$\frac{\partial r_i}{\partial q_j} = \hat{\mathbf{n}} \wedge r_i$$

(since $\mathrm{d}\mathbf{r}_i \perp r_i$ and $\hat{\mathbf{n}}$). So

$$Q_j = \sum_i \mathbf{F}_i \cdot \frac{\partial r_i}{\partial q_j} \tag{1.40}$$

$$= \sum_i \mathbf{F}_i \cdot \left(\mathbf{n} \hat{\wedge} \mathbf{r}_i\right) \tag{1.41}$$

$$= \sum_i \hat{\mathbf{n}} \cdot (r_i \wedge \mathbf{F}_i) \tag{1.42}$$

$$= \hat{\mathbf{n}} \cdot \left(\sum_i \mathbf{N}_i\right), \text{ the torque on the } i\text{th particle} \tag{1.43}$$

So $Q_j = 0 \Rightarrow \hat{\mathbf{n}} \cdot \mathsf{N} = 0$, where $\mathsf{N}$ is the total torque.. But this $\Rightarrow p_j = const = \mathbf{n} \cdot \sum_i m_i r_i \wedge \mathbf{v}_i = \mathbf{n} \cdot \mathbf{L}$
We deduce that $q_j$ (rotation about $\hat{\mathbf{n}}$) is ignorable $\Rightarrow$ zero torque $\Rightarrow$ angular momentum is conserved.

**Summary 1.** *We have reviewed IA Dynamics, and seen how Lagrange's equations are equivalent to Newton's Laws. It should be apparent, however, that they offer a more powerful approach to finding and solving the equations of motion: The equations themselves are easy to find; Conserved quantities are immediately apparent.*

# Chapter 2

# Rigid Bodies

**Synopsis:** This chapter is an in-depth application of the Lagrange formalism as developed in the previous chapter. We study rotating frames of reference, Eulerian angles, the Moment of Inertia Tensor and go on to investigate rigid body rotation, in particular the motion of a symmetrical spinning 'top'.

## 2.1 Frame of Reference

### 2.1.1 Rotating Frames

Let $OXY$ represent a fixed (inertial) frame. Let $Oxy$ be similar. Then

$$\mathbf{i}\cdot\mathbf{i} = \mathbf{j}\cdot\mathbf{j} = \mathbf{I}\cdot\mathbf{I} = \mathbf{J}\cdot\mathbf{J} = 1, \text{ and } \mathbf{i}\cdot\mathbf{j} = \mathbf{I}\cdot\mathbf{J} = 0.$$

The relationship between the frames is

$$\mathbf{i} = \mathbf{I}\cos\theta + \mathbf{J}\sin\theta, \quad \mathbf{j} = -\mathbf{I}\sin\theta + \mathbf{J}\cos theta,$$

So that

$$\frac{\mathrm{d}\mathbf{i}}{\mathrm{d}t} = (-\mathbf{I}\sin\theta + \mathbf{J}\cos\theta)\,\dot\theta \quad = \mathbf{j}\dot\theta \tag{2.1}$$

$$\frac{\mathrm{d}\mathbf{j}}{\mathrm{d}t} = (-\mathbf{I}\cos\theta - \mathbf{J}\sin\theta)\,\dot\theta \quad = -\mathbf{i}\dot\theta \tag{2.2}$$

So some general vector $\mathbf{r} = X\mathbf{I} + Y\mathbf{J} = x\mathbf{i} + y\mathbf{j} \Rightarrow$

$$\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = \dot x\mathbf{i} + \dot y\mathbf{j} + x\dot{\mathbf{i}} + y\dot{\mathbf{j}} \tag{2.3}$$

$$= \dot x\mathbf{i} + \dot y\mathbf{j} + x\dot\theta\mathbf{j} - y\dot\theta\mathbf{i} \tag{2.4}$$

Now we define $\boldsymbol{\omega} = \dot\theta\mathbf{k}$, so that

$$\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = \frac{\partial\mathbf{r}}{\partial t} + \boldsymbol{\omega}$$

Where the first term is from $\dot x\mathbf{i} + \dot y\mathbf{j}$ and the second from $x\omega\mathbf{j} - y\omega\mathbf{i}$. Now this is clearly true of any vector so

$$\frac{\mathrm{d}}{\mathrm{d}t} = \left(\frac{\partial}{\partial t} + \boldsymbol{\omega}\wedge\right) \tag{2.5}$$

For example $\mathbf{F} = m\dot{\mathbf{v}}$, where m is constant so (??) $\Rightarrow$

$$\mathbf{F} = m\left(\frac{\partial \mathbf{v}}{\partial t} + \boldsymbol{\omega} \wedge \mathbf{v}\right) \tag{2.6}$$

$$= m\left(\frac{\partial}{\partial t} + \boldsymbol{\omega}\wedge\right)(\dot{\mathbf{r}} + \boldsymbol{\omega} \wedge \mathbf{r}) \tag{2.7}$$

$$= m\frac{\partial^2 \mathbf{r}}{\partial t^2} + 2\boldsymbol{\omega} \wedge \frac{\partial \mathbf{r}}{\partial t} + \frac{\mathrm{d}\boldsymbol{\omega}}{\mathrm{d}t} \wedge \mathbf{r} + \boldsymbol{\omega} \wedge (\boldsymbol{\omega} \wedge \mathbf{r}) \tag{2.8}$$

So

$$m\frac{\partial^2 \mathbf{r}}{\partial t^2} = \mathbf{F} - 2\boldsymbol{\omega} \wedge \frac{\partial \mathbf{r}}{\partial t} - \frac{\mathrm{d}\boldsymbol{\omega}}{\mathrm{d}t} \wedge \mathbf{r} - \boldsymbol{\omega} \wedge (\boldsymbol{\omega} \wedge \mathbf{r})$$

And you see we've shown that the *Coriolis Force* is $2m\mathbf{v} \wedge \dot{\mathbf{r}}$ and the *Centrifugal Force* is $m\boldsymbol{\omega} \wedge (\boldsymbol{\omega} \wedge \mathbf{r})$.
We interpret this as saying that if we wish to pretend a non-inertial frame is inertial, we must invent 'fictitious forces'.

### 2.1.2 Transforming Between Frames

A rigid body with $N$ particles can have at most $3N$ degrees of freedom. This number will be reduced by the constraint $r_{ij}^2 = c_{ij}^2$, fixed. You need at most 6 coordinates: To establish the position of one particle in the body we need three coordinates. Call this particle 1. Now to fix the position of particle 2 we need only two coordinates (it must lie on a sphere centred on particle 1 and with radius $c_{ij}$). If we take a third particle, 3, we can only rotate the axis joining particles 1 and 2, this is the final degree of freedom. All other particles are uniquely fixed.
The relation between 2 Cartesian frames can be written

$$(x_1, x_2, x_3) \to (x_1', x_2', x_3'); \quad \mathbf{x} \mapsto \mathbf{x}' = A\mathbf{x}.$$

Then, for a transformation preserving length

$$\mathbf{x} \cdot \mathbf{x} = A\mathbf{x} \cdot A\mathbf{x} \tag{2.9}$$

$$\Rightarrow \quad \sum_k a_{jk} a_{ik} = \delta_{ij} \tag{2.10}$$

$$\text{ie.} \quad AA^{\mathrm{T}} = A^{\mathrm{T}} A = I \tag{2.11}$$

$$\Rightarrow \quad \det A = \pm 1, \tag{2.12}$$

Let us assume that $\{x_i'\} = \{x_i\}$ at time $t = 0$.

**Remark 2.** $\det A = -1$ *cannot be achieved by* any rigid *change of coordinate axis. We will assume* $\det A = +1$ *always.*

We can transform from a given Cartesian frame to another with the same origin by at most three rotations.
Suppose we start with $X, Y, Z$. Then rotate through $\phi$ counter-clockwise about the $Z$ axis.

Then rotate c.clockwise through $\theta$ about the $\xi$ axis:

Finally rotate c.clockwise through $\psi$ about $\rho'$
This defines our new axes $x', y', z'$ Let us formalise this in matricies

$$B = \begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{2.13}$$

$$C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix} \tag{2.14}$$

$$D = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{2.15}$$

Then the entire rotation has the form $A = BCD$, or

$$A = \begin{pmatrix} \cos\psi\cos\phi - \cos\theta\sin\phi\sin\psi & \cos\psi\sin\phi + \cos\theta\cos\phi\sin\psi & \sin\psi\sin\theta \\ -\sin\psi\cos\phi - \cos\theta\sin\phi\cos\psi & -\sin\psi\, sin\phi + \cos\theta\cos\phi\cos\psi & \cos\psi\cos\theta \\ \sin\theta\sin\phi & -\sin\theta\cos\phi & \cos\theta \end{pmatrix}$$

### 2.1.3 Euler's Theorem

The most general displacement of a rigid body with one fixed point is a rotation about some axes.
At any instant the orientation of such a body can be specified by an orthogonal transformation $A(t)$, for simplicity we assume $A(0) = I$. $A(t)$ will be a cts. function of time. The transformation will be a rotation if

1. The transformation leaves one direction unchanged. (the axis about which it rotates).
2. The magnitude of vectors are left unchanged

Note: (2) follows from the orthogonality condition.

**Proposition 1.** *The real orthogonal transformation specifying the physical motion of a system with one fixed point always has eigenvalue* $+1$.

Proof

$$(A - \lambda I)\,\mathbf{r} = 0 \text{ has a solution iff} \tag{2.16}$$

$$\det(A - \lambda I) = 0 \tag{2.17}$$

$$\Rightarrow \quad (A - I)\,A^{\mathrm{T}} = \left(I - A^{\mathrm{T}}\right) \tag{2.18}$$

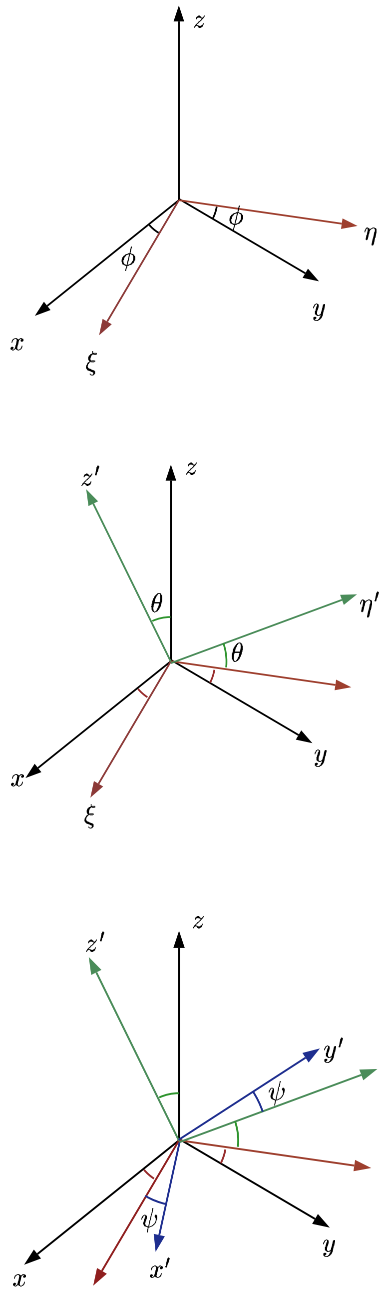


Figure 2.1: Eulerian Angles: First rotate about the $z$-axis, then about the (new) $\xi$-axis, and finally about the $z'$-axis.

So $\det(A-I)-\det A^{\mathrm{T}} = \det(I-A^{\mathrm{T}})$. But $(I-A)^{\mathrm{T}} = I-A^{\mathrm{T}}$, hence $\det(A-I) = \det(I-A)$ using $\det A^{\mathrm{T}} = \det A$. □
For any $n \times n$ matrix we have that $\det(-B) = (-1)^n \det(B)$, so $\det(A-I) = \det(I-A) = -\det(A-I) \Rightarrow \det(A-I) = 0$. Hence the eigenvalues is $\lambda = +1$. □.
Now we can transform $A$ st.

$$XAX^{-1} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

where the $\lambda_i$ are the eigenvalues of A. This $\Rightarrow \det A = \lambda_1\lambda_2\lambda_3$ and $\lambda_i = 1$ for some i=$1, 2, 3$. Suppose that $\lambda_3 = 1$ then $\lambda_1 = \lambda_2^*$ and $|\lambda_1| = |\lambda_2| = 1$, since A is a rotation. Then there are three cases:

1. $\lambda_1 = \lambda_2\lambda_3 \Rightarrow A = I$, a 'rotation by $2\pi$'.
2. $\lambda_1 = \lambda_2 = -\lambda_3 \Rightarrow$ rotation through $\pi$.
3. $\lambda_1 = e^{i\phi},\ \lambda_2 = e^{-i\phi}$ then

$$\begin{pmatrix} e^{-\phi} & 0 & 0 \\ 0 & e^{-i\phi} & 0 \\ 0 & 0 & 1 \end{pmatrix} \overset{\text{similar}}{\longrightarrow} \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

ie. rotation about the $z$-axis in the new frame. This proves Euler's Theorem.

**Note 6.** *1. For*

$$A = \begin{pmatrix} e^{i\phi} & 0 & 0 \\ 0 & e^{-i\phi} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

*we have $mathrmtr(A) = 1_2\cos\phi$, and remember that the trace is the same for similar matricies.*

2. *The sense of direction of the rotation is* not yet *well defined, since if $\lambda$ is an eigenvalue, so too is $-\lambda$. Ie. if $\mathbf{x}$ is an eigenvector the $A\mathbf{x} = \lambda\mathbf{x}$ then $\mathbf{x}$ is also an eigenvector with the same eigenvalue.*

We assign $\phi$ with $A$ and $-\phi$ with $\bar{A}(= A^{-1})$, and use a *right hand rule*.
Similarly we have *Chasles Theorem*: The most general displacement of a rigid body is a translation plus a rotation.
This suggests that the 6 coordinates needed could well be the 3 cartesian coordinates to fix the body in space and then the 3 Eulerian angles.

## 2.2 The Moment of Inertia

We know that the total kinetic energy of the system can be written

$$= \frac{mv^2}{2} + T'(\theta, \phi, \psi)$$

Ie. the sum of translational and rotational energies. The total angular momentum about

We assume here that $\dot{\theta}$ etc. are not independent of $\theta, \phi\psi$.

some point $O$ is

$$L = \mathbf{R} \wedge M\mathbf{v} + \sum_i \mathbf{r}'_i \wedge \mathbf{P}'_i$$

(again the ang. mom. of the body concentrated at CoM plus the ang. mom. about the CoM.) The essence of rigid body motion is that all the particles that make up the body move and rotate together. When a rigid body moves with one point stationary then the total ang. mom. about that point is

$$L = \sum_i m_i \left( r_i \wedge \mathbf{v}_i \right)$$

with $r_i$ and $\mathbf{v}_i$ given wrt the fixed point.
Since $r_i$ is fixed relative to the body the velocity $\mathbf{v}_i$ wrt. the space arises solely from the rotation

$$\mathbf{v}_i = \frac{\partial r_i}{\partial t} + \boldsymbol{\omega} \wedge r_i$$

And

$$\begin{aligned} L &= \sum_i m_i \left( r_i \wedge (\boldsymbol{\omega} \wedge r_i) \right) \\ &= \sum_i m_i \left( \boldsymbol{\omega} r_i^2 - r_i \left( r_i \cdot \boldsymbol{\omega} \right) \right) \\ \text{ie.} \begin{pmatrix} L_x \\ L_y \\ L_z \end{pmatrix} &= \begin{pmatrix} \sum_i m_i(r_i^2 - x_i^2) & -\sum_i m_i x_i y_i & -\sum_i m_i x_i z_i \\ -\sum_i m_i x_i y_i & \sum_i m_i(r_i^2 - y_i^2) & -\sum_i m_i z_i y_i \\ -\sum_i m_i x_i z_i & -\sum_i m_i z_i y_i & \sum_i m_i(r_i^2 - z_i^2) \end{pmatrix} \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix} \end{aligned}$$

The ang. mom. vector is related to the ang. mom. by the linear transformation

$$\mathbf{L} = \begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix} \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_x \end{pmatrix}$$

The diagonal terms are called the *moments of inertia* while the off-diagonal terms are called the *products of inertia*. In the case of a continuous mass distribution we would replace the sums by integrals in the obvious way

$$I_{xy} = -\int_V \rho(r) xy \, \mathrm{d}\tau$$

Notation: Sometime we will make use of the notation $(x_1, x_2, x_3) = (x, y, z)$. In this notation

$$I_{ij} = \int_V \rho(r)(r^2 \delta_{ij} - x_i x_j) \, \mathrm{d}\tau$$

So then

$$\mathbf{L} = \mathbb{I} \, \boldsymbol{\omega}$$

where $\mathbb{I}$ is the *moment of inertia tensor*. A second rank tensor.

**Remark 3.** *Sometimes one makes use of 'dyads'. A dyad,* $\mathbf{ab} \equiv \mathbf{a}^{\mathrm{T}}\mathbf{b} \not\equiv \mathbf{ba}$ *is the* outer product *of two vectors,* $\mathbf{a}$ *and* $\mathbf{b}$*: It is a 2nd rank tensor. Don't confuse it with the* inner product, $\mathbf{a} \cdot \mathbf{b} \equiv \mathbf{ab}^{\mathrm{T}}$

In this language

$$\mathbb{I} = \sum_i m_i \left( r_i \delta_{ij} - r_i \mathbf{r}_j \right)$$

$$\mathbf{L} = \mathbb{I} \cdot \boldsymbol{\omega} = \sum m_i \left( r_i \boldsymbol{\omega} - r_i (r_i \boldsymbol{\omega}) \right)$$

$$\begin{aligned} T &= \sum_i \frac{1}{2} m_i v_i^2 && (2.19) \\ &= \sum \frac{1}{2} m_i \mathbf{v}_i \left( \boldsymbol{\omega} \wedge r_i \right) && (2.20) \\ &= \sum_i \frac{\boldsymbol{\omega}}{2} \cdot m_i (r_i \wedge \mathbf{v}_i) && (2.21) \\ &= \boldsymbol{\omega} \cdot \mathbf{L} && (2.22) \\ &= \frac{1}{2} \boldsymbol{\omega} \cdot \mathbb{I}\, \boldsymbol{\omega} && (2.23) \end{aligned}$$

Let $n$ be a unit vector in the direction of rotation ie.

$$\boldsymbol{\omega} = \omega \mathbf{n} \Rightarrow T = \frac{\omega^2}{2} \mathbf{n}^{\mathrm{T}} \mathbb{I} \mathbf{n} = \frac{1}{2} I \omega^2$$

Then we say that $I$ is the *moment of inertia* about the axis $\mathbf{n}$ of rotation. Now let us consider the vector $r_i \wedge \mathbf{n}$. It's magnitude will be the perpendicular from the axis of the rotation. So

$$\begin{aligned} I = \frac{2T}{\omega^2} &= \sum \frac{m_i}{\omega^2} \left( \mathbf{v}_i \cdot \mathbf{v}_i \right) && (2.24) \\ &= \sum \frac{m_i}{\omega^2} \left( \boldsymbol{\omega} \wedge r_i \right) \cdot \left( \boldsymbol{\omega} \wedge r_i \right) && (2.25) \\ &= \sum m_i \left( \mathbf{n} \wedge r_i \right) \cdot \left( \mathbf{n} \wedge r_i \right) && (2.26) \end{aligned}$$

$I$ the moment of inertia about an axis is the sum over all the particles in the body, of the product of the masses times their perpendicular distance from the axis.

Let the vector from the origin, $O$, to the CoM be $\mathbf{R}$. Let the radius vector from $O$ and $\mathbf{R}$ be $r_i$ and $r_i'$ respectively. Then the moment of inertia about an axis $\mathbf{a}$ is

$$\begin{aligned} I_{\mathbf{a}} &= \sum m_i \left( r_i \wedge \mathbf{n} \right)^2 && (2.27) \\ &= \sum m_i \left( (r_i' + \mathbf{R}) \wedge \mathbf{n} \right)^2 && (2.28) \\ &= \left( \sum m_i \right) \left( \mathbf{R} \wedge \mathbf{n} \right)^2 + \sum m_i \left( r_i' \wedge \mathbf{n} \right)^2 + \sum 2 m_i \left( \mathbf{R} \wedge \mathbf{n} \right) \cdot \left( r_i' \wedge \mathbf{n} \right) && (2.29) \end{aligned}$$

If we write $M = \sum m_i$, the total mass of the system, and $\sum m_i r_i' = 0$, by the definition of the CoM we have that

$$I_{\mathbf{a}} = I_{\mathbf{b}} + M \left( \mathbf{R} \wedge \mathbf{n} \right)^2$$

This is called the *parallel axis theorem*. It states that the moment of inertia about a given axis is the same as the MoI about a parallel axis going through the CoM + the MoI of the CoM wrt. the original axis.

| Body | Dimension(s) | Axis | Moment of Inertia |
|---|---|---|---|
| Solid Sphere | radius $a$ | diameter | $\frac{2}{5}Ma^2$ |
| Hollow Sphere | radius $a$ | diameter | $\frac{2}{3}Ma^2$ |
| Solid Cylinder | radius $a$, length $l$ | symmetry axis | $\frac{1}{2}Ma^2$ |
| | | through centre, $\perp$ symmetry axis | $\frac{1}{4}Ma^2 + \frac{1}{12}Ml^2$ |
| Hollow Cylinder | radius $a$, length $l$ | symmetry axis | $Ma^2$ |
| Solid Cuboid | sides $2a, 2b, 2c$ | through centre, $\parallel$ side of length $2c$ | $\frac{1}{3}M\left(a^2+b^2\right)$ |
| Solid Cone | radius $a$, height $l$ | symmetry axis | $\frac{3}{10}Ma^2$ |

Table 2.1: Some common moments of inertia

### 2.2.1 Properties of the Moment of Inertia Tensor

The Moment of Inertia Tensor has the following properties:

1. It is symmetric, $I_{xy} = I_{yx}$.
2. All its values are real $\Rightarrow$ real eigenvalues.
3. Together these imply that it is self-adjoint.

Lemma: All eigenvalues of $\mathbb{I}$ are real and it's eigenvectors are mutually orthogonal.
We know that $\mathbb{I}$ can be put in diagonal form. The axes corresponding to this diagonal form are known as the *principle axes* and the diagonal elements $I_1$, $I_2$, $I_3$ (ie. the eigenvalues of the tensor) are the *principle moments of inertia*. They satisfy

Notation: $I_i$ is an eigenvalue and $I$ is the identity matrix

$$\det(\mathbb{I} - I_i I) = 0$$

Now $I_{xx} = \sum m_i \left( y_i^2 + z_i^2 \right) \geq 0 \Rightarrow I_1,\, I_2,\, I_3 \geq 0$. Now consider an inertial frame $B$ whose origin is at a fixed point of a rigid body (on a system of space axes $S$ with the origin at the centre). For an axis fixed in the body

$$\left( \frac{\mathrm{d}\mathbf{L}}{\mathrm{d}t} \right)_S = \left( \frac{\partial \mathbf{L}}{\partial t} \right)_B + \boldsymbol{\omega} \wedge \mathbf{L} = \mathbf{N}$$

ie. $N_i = \frac{\partial L_i}{\partial t} + sum_{jk} \varepsilon_{ijk} \omega_j L_k$. The angular momentum components are $L_i = I_i \omega_i$, the principle moments of inertia are time independent:

$$I_i \frac{\mathrm{d}\omega_i}{\mathrm{d}t} + \sum_{jk} \varepsilon_{ijk} \omega_j \omega_k I_k = N_i$$

These are *Euler's Equations*. In full

$$N_1 = I_1 \dot{\omega}_1 - \omega_2 \omega_3 (I_2 - I_3) \quad (2.30)$$

$$N_2 = I_2 \dot{\omega}_2 - \omega_3 \omega_1 (I_3 - I_1) \quad (2.31)$$

$$N_3 = I_3 \dot{\omega}_3 - \omega_1 \omega_2 (I_1 - I_2) \quad (2.32)$$

## 2.3 Spinning Tops

### 2.3.1 Deriving the Lagrangian

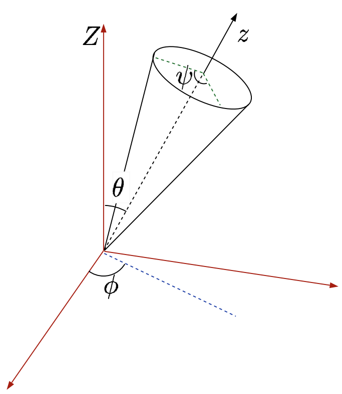


Figure 2.2: Eulerian Angles as applied to the spinning top

Consider a symmetric top with one point fixed. We will used a body fixed set $(x, y, z)$. One of the principle axes will be the $z$-axis, as fixed in the body.

Since one point is fixed - the Eulerian angles are all we need to describe the body. $\theta$ gives the inclination of the $z$-axis about the vertical. $\phi$ measures the azimuth at the top of the vertical and $\psi$ measures the rotation angle of the top about it's own $z$-axis. The general infinitesimal rotation associated with $\boldsymbol{\omega}$ can be considered the result of 3 rotations:

1. $\dot{\phi} = \omega_\phi$ about $Z$ (space frame)

2. $\dot{\theta} = \omega_\theta$ about $\xi'$

3. $\dot{\psi} = \omega_\psi$ about $z$ (body frame)

Now $\boldsymbol{\omega}$ is parallel to the space-fixed axis $Z \Rightarrow$ to put it in terms of the body frame we need to apply the orthogonal transformation $A = BCD$, with

$$(\omega_\phi)_x = \dot{\phi}\sin\theta\sin\psi \tag{2.33}$$
$$(\omega_\phi)_y = \dot{\phi}\sin\theta\cos\psi \tag{2.34}$$
$$(\omega_\phi)_z = \dot{\phi}\cos\theta \tag{2.35}$$

Now the direction of $\boldsymbol{\omega}_\theta$ coincides with the $\xi'$ axis. So the components of $\boldsymbol{\omega}_\theta$ wrt. the body fixed axes is given by applying B:

$$(\omega_\theta)_x = \dot{\theta}\cos\psi \tag{2.36}$$
$$(\omega_\theta)_y = -\dot{\theta}\sin\psi \tag{2.37}$$
$$(\omega_\theta)_z = 0 \tag{2.38}$$

No transformation is needed for $\omega_\psi$ since it is already about $z$.
Adding the components of $\boldsymbol{\omega} = \omega_\phi + \omega_\psi + \omega_\theta$ we get

$$\boldsymbol{\omega} = \begin{pmatrix} \dot{\theta}\sin\theta\sin\psi + \dot{\theta}\cos\psi \\ \dot{\phi}\sin\theta\cos\psi + \dot{\theta}\sin\psi \\ \dot{\phi}\cos\theta + \dot{\psi} \end{pmatrix}$$

Hence the body is symmetric.

$$T = \frac{1}{2}I_1\left(\omega_x^2 + \omega_y^2\right) + \frac{1}{2}I_3\omega_z^2 \tag{2.39}$$
$$= \frac{1}{2}I_1\left(\dot{\theta}^2 + \dot{\phi}^2\sin^2\theta\right) + \frac{1}{2}I_3\left(\dot{\psi} + \dot{\phi}\cos\theta\right)^2 \tag{2.40}$$
$$V = -\sum_i m_i\mathbf{r}_i\cdot\mathbf{g} = -M\mathbf{R}\cdot\mathbf{g} \tag{2.41}$$
$$= Mgl\cos\theta \tag{2.42}$$

where $l$ is the distance from the fixed point to the CoM, and the angles are Eulerian Angles. Hence

$$\mathsf{L} = \frac{1}{2}I_1\left(\dot{\theta}^2 + \dot{\phi}^2\sin^2\theta\right) + \frac{1}{2}I - 3\left(\dot{\psi} + \dot{\phi}\cos\theta\right)^2 + Mgl\cos\theta.$$

The $\phi$ and $\psi$ are cyclic. Hence $p_\psi = \frac{\partial \mathsf{L}}{\partial\dot{\psi}} = I_3\left(\dot{\psi} + \dot{\phi}\cos\theta\right) = const. = I_3\omega_3$. And $p_\phi = \frac{\partial \mathsf{L}}{\partial\dot{\phi}} = \left(I_1\sin^2\theta + I_3\cos^2\theta\right)\dot{p}hi + I_3\dot{\psi}\cos\theta = const. = I_1 b$, and we define $I_1 a = I_3\omega_3$.
So the two constraints of the motion $p_\psi$ and $p_\phi$ can be expressed in terms of $a$ and $b$.

### 2.3.2 Conserved Quantities

The total energy is given by

$$E = T + V = \frac{1}{2}I1\left(\dot{\theta}^2 + \dot{\phi}^2\sin^2\theta\right) + \frac{I_3}{2}\omega_3^2 + Mgl\cos\theta$$

Now

$$I_3\dot{\psi} = I_1 a - I_3\dot{\phi}\cos\theta \tag{2.43}$$

If we substitute in for $p_\phi$ we get

$$I - 1\dot{\phi}\sin^2\theta + I_1 a\cos\theta = I_1 b \tag{2.44}$$

Then equations (??) and (??) $\Rightarrow$

$$\dot{\psi} = \frac{b - a\cos\theta}{\sin^2\theta} \tag{2.45}$$

and

$$\dot{\psi} = \frac{I_1 a}{I_3} - \frac{\cos\theta\,(b - acos\theta)}{\sin^2\theta} \tag{2.46}$$

Now $\omega_3 = \frac{I_1 a}{I_3}$ is a constant of the motion. It is (sometimes denoted $n$ and) called the *spin*.
Define $E' = E - \frac{I_3}{\omega_3}2$, another constant. We can write

$$E' = \frac{I_1\dot{\theta}^2}{2} + \underbrace{\frac{I_1\,(b - a\cos\theta)^2}{2\sin^2\theta} + Mgl\cos\theta}_{\tilde{V}(\theta)}$$

Or

$$E' = \frac{I_1\dot{\theta}^2}{2} + \tilde{V}(\theta)$$

This looks like a one dimensional problem, with an effective potential $\tilde{V}(\theta)$. Making the change of variable $u = \cos\theta$ we have:

$$E'(1 - u^2) = \frac{I_2}{2}\dot{u}^2 + \frac{I_1}{2}\,(b - au)^2 + Mglu\left(1 - u^2\right)$$

And letting $\alpha = \frac{2E'}{I_1}$ and $\beta = \frac{2Mgl}{I_1}$ then

$$\dot{u}^2 = (1 - u^2)(\alpha - \beta u) - (b - au)^2$$

Hence

$$\epsilon = \int_{u_1(\epsilon)}^{u_1(\epsilon)} \frac{\mathrm{d}u}{\sqrt{(1 - u^2)(\alpha - \beta u) - (b - au)^2}}$$

Unfortunately this integral is elliptic, and the solutions for $\theta$, $\phi$, $\psi$ are in terms of elliptic integrals.

If we look at the equations of motion, as derived from Lagrange's Equations we have that

$$I_1\ddot{\theta} - I_1\dot{\phi}^2 \sin\theta\cos\theta + I_3 \sin\theta\left(\dot{\phi} + \dot{\phi}\cos\theta\right) - Mgl\sin\theta = 0 \quad (2.47)$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(I_1\dot{\phi}\sin^2\theta + I_3\cos\theta\left(\dot{\psi} + \dot{\phi}\cos\theta\right)\right) = 0 \quad (2.48)$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(I_3\left(\dot{\psi} + \dot{\phi}\cos\theta\right)\right) = 0 \quad (2.49)$$

This last reveals

$$\dot{\psi} + \dot{\phi}\cos\theta = const. = \omega_3 = \text{'spin'}$$

And so we can write

$$I_1\ddot{\theta} - I_1\dot{\phi}^2 \sin^2\theta\cos\theta + I_3\omega_3\dot{\phi}\sin\theta - Mgl\sin\theta = 0 \quad (2.50)$$

$$\Rightarrow I_1\dot{\theta}^2 I_1\dot{\phi}^2\sin^2\theta + 2Mgl\cos\theta = const = F \quad (2.51)$$

This is the conservation of energy. And finally the middle equation gives

$$I_1\dot{\phi}\sin^2\theta + I_3\omega_3\cos\theta = const. = D$$

The motion due to the change in $\theta$ is called *nutation*, and the motion due to change in $\phi$ is called precession.

### 2.3.3 Steady Motion

Steady motion has $\theta = \alpha = const. \Rightarrow \dot{\phi} = const. = \Omega$, say. Now provided we have $\theta \neq 0$ we have:

$$I_1\Omega^2\cos\alpha - I_3\omega_3\Omega + mgl = 0$$

$\Rightarrow$ a pair of real distinct precessional angular velocities $\Omega_1$ and $\Omega_2$, provided $\omega_3^2 > 4I_1 Mgl\cos\alpha$. Ie. for sufficiently fast spin, $\omega$, about the axis of symmetry the top can perform steady motion, with $\theta = \alpha$, with 2 possible precessional velocities, $\omega_1$ and $\omega_2$.

### 2.3.4 Stability Investigation

Let $\gamma(\theta) = -I_1\dot{\phi}^2\sin\theta\cos\theta + I_3\omega_3\dot{\phi}\sin\theta - Mgl\sin\theta$ and the stability condition $\gamma(\alpha) = 0$. Now putting $\theta = \alpha + \epsilon$ where $\epsilon$ is small, we have

$$I_1\epsilon'' + \epsilon\gamma'(\epsilon) = 0 \sim \gamma(\alpha+\epsilon) \approx \gamma'(\alpha)$$

If we can now show that $\gamma'(\epsilon) > 0$ then this equation will reduce to SHM: The steady point is stable. Now we know that

$$\dot{\phi} = \frac{(D - I_3\omega_3\cos\theta)}{I_1\sin^2\theta} \quad (2.52)$$

$$\Rightarrow I_1\gamma(\theta)\sin^3\theta = -\cos\theta\left(D - I_3\omega_3\cos\theta\right)^2 + I_3\omega_3\sin^2\theta\left(D - I_3\omega_3\cos\theta\right) - I_1 Mgl\sin\theta \quad (2.53)$$

Now differentiating both sides and putting $\gamma(\alpha) = 0$, we have

$$\begin{aligned} I_1\gamma'(\alpha)\sin^3\alpha &= \sin\alpha\,(D - I_3\omega_3\cos\alpha) && (2.54)\\ &\quad -2I_3\omega_3\cos\alpha\sin\alpha\,(D - I_3\omega_3\cos\alpha) && (2.55)\\ &\quad +2I_3\omega_3\cos\alpha\sin\alpha\,(D - I_3\omega_3\cos\alpha) && (2.56)\\ &\quad +I_3^2\omega_3^2\sin^2\alpha - 4I_3Mgl\sin^2\alpha\cos\alpha && (2.57) \end{aligned}$$

So that

$$\begin{aligned} D - I_3\omega_3\cos\alpha &= I_1\Omega\sin^2\alpha && (2.58)\\ I_3\omega_3 &= I_1\Omega\cos\alpha + \frac{Mgl}{\Omega} && (2.59) \end{aligned}$$

Thus

$$I_1\gamma'(\alpha) = I_1^2\Omega^2 - 2I_1Mglcos\alpha + \left(\frac{Mgl}{\Omega}\right)^2 \quad (2.60)$$

$$\Rightarrow \quad \epsilon'' + \frac{\gamma'(\alpha)\epsilon}{I_{1?????}} = 0 \quad (2.61)$$

So that

$$\frac{\gamma'(\alpha)}{I_1} = \left\{\Omega - \left(\frac{Mgl}{I_1\Omega}\right)\right\}^2 + \frac{2Mgl\,(1 - cos\alpha)}{I_1}$$

Therefore $\gamma'(\alpha) > 0\,\forall\alpha \neq 0$, and so the motion is SHM about the stable point.

# Chapter 3

# Hamilton's Equations, & Onwards to Abstraction

**Synopsis:** The alternative Hamiltonian formalism is developed. Momenta and Coordinates become indistinguishable. Canonical transformations/Poisson brackets allow us to reformulate the Hamiltonian: By finding a 'good' set of coordinates solving the equations of motion becomes trivial (Hamilton-Jacobi theory). Conserved quantities and symmetries are related. Liouville's Theorem provides an important link to Statistical/Continuum mechanics.

**Notation 1.** *Until now we have used $P_i$, or $\mathbf{P}$ for momenta - whether or not they were generalised momenta. However, central to the Hamiltonian method of doing things is the concept that coordinates and momenta are viewed equally: Henceforth we shall write $p_i$ for the momenta of a system with coordinates $q_i$. (The transformed coordinates shall then be written as $P_i$ and $Q_i$.)*

## 3.1 Hamilton's Equations

### 3.1.1 An alternative approach

So far we have formulated everything in terms of the $n$ independent variables $q_i$. We have, in effect, treated $\dot{q}_i$ as distinct variables, independent of the $q_i$ of which they are the time derivative. For example we have used $\frac{\partial \mathsf{L}}{\partial q_i}$ to mean the derivative wrt. $q_i$ with $q_j \neq q_i$ and $\dot{q}_j$ held constant, and the symbol $\frac{\partial \mathsf{L}}{\partial \dot{q}_i}$ has been taken to mean the derivative wrt. $\dot{q}_i$ with all $\dot{q}_j$ and all $q_j$ held constant.
We could work in a space define by $q_i, \dot{q}_i, t$ but it will introduce a greater symmetry if we work with the *conjugate momenta*

$$p_j = \frac{\partial \mathsf{L}}{\partial \dot{q}_j}.$$

The generalised momenta $p_j$ are said to be *conjugate* to the $q_j$ and the quantities are said to be *canonical* variables. **Definition:** The *Hamiltonian* of the system is defined by

So, in the Lagrangian formulation when a variable is absent from $\mathsf{L}$, its conjugate variable is conserved.

$$\mathsf{H}(q_i, p_i, t) = \sum_i \dot{q}_i p_i - L(q_i, \dot{q}_i, t)$$

Then

$$\mathrm{dH} = \frac{\partial \mathrm{H}}{\partial q_i}\,\mathrm{d}q_i + \frac{\partial \mathrm{H}}{\partial p_i}\,\mathrm{d}p_i + \frac{\partial \mathrm{H}}{\partial t} \tag{3.1}$$

$$= \sum_i (\dot{q}_i\,\mathrm{d}p_i + p_i\,\mathrm{d}\dot{q}_i) - \frac{\partial \mathrm{L}}{\partial \dot{q}_i}\,\mathrm{d}q_i - \frac{\partial \mathrm{L}}{\partial q_i}\,\mathrm{d}q_i - \frac{\partial \mathrm{L}}{\partial t} \tag{3.2}$$

But $\frac{\mathrm{d}}{\mathrm{d}t}p_i = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathrm{L}}{\partial \dot{q}_i}\right) = \frac{\partial \mathrm{L}}{\partial q_i}$ so that $\mathrm{dH} = \sum_i \dot{q}_i\,\mathrm{d}p_i - \dot{p}_i\,\mathrm{d}q_i - \frac{\partial \mathrm{L}}{\partial t}\,\mathrm{d}t$ so equating variables we have

$$\boxed{\dot{q}_i = \frac{\mathrm{dH}}{\mathrm{d}p_i}, \qquad \dot{p}_i = -\frac{\partial \mathrm{H}}{\partial q_i}, \qquad -\frac{\partial \mathrm{L}}{\partial t} = \frac{\partial \mathrm{H}}{\partial t}}$$

These are the canonical equations of Hamilton. They are a set of coupled partial differential equations.

**Note 7.** *Features:*

- *Hamilton's Equations are first order. Lagrange's Equations were second order.*
- *Hamilton's Equations are in $2n$ variables $q_i$ and $p_i$, Lagranges equations were in the $n$ constraints $q_i$. We now need to determine $2n$ constants.*
- $$\frac{\partial \mathrm{H}}{\partial t} = 0 \Rightarrow \frac{\mathrm{dH}}{\mathrm{d}t} = \sum_i \frac{\partial \mathrm{H}}{\partial q_i}\frac{\mathrm{d}q_i}{\mathrm{d}t} + \frac{\partial \mathrm{H}}{\partial p_i}\frac{\mathrm{d}p_i}{\mathrm{d}t} \Rightarrow \sum_i \frac{\partial \mathrm{H}}{\partial q_i}\frac{\partial \mathrm{H}}{\partial p_i} - \frac{\mathrm{dH}}{\mathrm{d}q_i}\frac{\partial \mathrm{H}}{\partial p_i} = 0$$

  $\Rightarrow \mathrm{H}$ *is a constant of the motion.*

**Example 1.** *One Dimensional Motion*
*Consider one dimensional motion and suppose the existence of a potential $V(x)$ st $\mathbf{F} = -\frac{\mathrm{d}}{\mathrm{d}x}V(x)$. Then*

$$\mathrm{L} = \frac{1}{2}mv^2 - V(x) \tag{3.3}$$

$$\mathrm{H} = mv^2 - \frac{1}{2}mv^2 + V(x) \tag{3.4}$$

$$= T + V = \textit{Total Energy} \tag{3.5}$$

Notation: here $q$ is the charge, not a generalised coordinate

**Example 2.** *Electromagnetic Field*
*Consider a small (non-relativistic) particle in an EM field:*

$$\mathrm{L} = T - U = \frac{1}{2}mv^2 - q\phi + \frac{q}{c}\mathbf{A}\cdot\mathbf{v} \tag{3.6}$$

$$= \frac{1}{2}\sum_i m\dot{x}_i\dot{x}_i + \frac{q}{c}A_i\dot{x}_i - q\phi, \tag{3.7}$$

*in Cartesian coords. The* generalised momenta *are given by*

$$p_i = m\dot{x}_i + \frac{q}{c}A_i$$

*Then*

$$\mathrm{H} = \sum_i \frac{1}{2m}\left(p_i - \frac{q}{c}A_i\right)^2 + q\phi \tag{3.8}$$

$$\Rightarrow \mathrm{H} = \frac{1}{2m}\left(\mathbf{p} - \frac{q}{c}\mathbf{A}\right)^2 + q\phi \tag{3.9}$$

### 3.1.2 Cyclic Coordinates and Conservation Theorems

Observe

$$\dot{p}_j = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathsf{L}}{\partial \dot{q}_j}\right) = -\frac{\partial \mathsf{H}}{\partial q_j}.$$

*So if a coordinate is absent from the Hamiltonian the corresponding (conjugate) generalised momentum is conserved.*
If the generalised momentum $p_j$ is absent from the Hamiltonian then

$$\frac{\partial \mathsf{H}}{\partial p_j} = 0 \Rightarrow \dot{q}_j = 0$$

So $q_j$ is conserved.

### 3.1.3 The principle of Least Action

In coordinate space we have

$$\delta \int_{t_1}^{t_2} \mathsf{L}(q_i, \dot{q}_i, t)\,\mathrm{d}t = 0,$$

so we should be able to write

$$\delta \int_{t_1}^{T_2} \left(\sum_i p_i \dot{q}_i - \mathsf{H}(p_i, q_i, t)\right)\,\mathrm{d}t = 0.$$

But let us stop here and think: This equation is implicitly in phase space. We have to think about what we mean by an 'independent variation': Since we derived Lagrange's Equation (in $n$ variables) be assuming $\delta q_i = 0$, and by slight of hand we now have $2n$ variables $(q_i, p_i)$, and so by writing the above we are assuming $\delta q_i = 0$ *and* $\delta p_i = 0$. In the $\Delta$-variation:

- The varied path over which the integral is evaluated may end at different times from the 'correct path'
- There may be a variation in the coordinates at the end points.

Consider the family of varied paths defined by

$$q_i(t, \alpha) = \underbrace{q_i(t, 0)}_{\text{'true path'}} + \alpha \eta_i(t)$$

So $\alpha$ is an infinitesimal parameter which goes to zero for the 'exact' path. The $\eta_i$ *do not necessarily* have to vanish at the end points, $t_1$ and $t_2$.
All that is required of the $\eta_i$ is that they are continuously differentiable. We are interested in finding the $\Delta$-variation on the action integral, ie:

$$\Delta \int_{t_1}^{t_2} \mathsf{L}\,\mathrm{d}t$$

which we define by

$$\Delta \int_{t_1}^{t_2} \mathsf{L}\,\mathrm{d}t = \int_{t_1+\Delta t_1}^{t_2+\Delta t_2} \mathsf{L}(\alpha)\,\mathrm{d}t - \int_{t_1}^{t_2} \mathsf{L}(0)\,\mathrm{d}t,$$

where $\mathsf{L}(\alpha)$ means the integral is evaluated along the path $\alpha = \alpha$, and $\mathsf{L}(0)$ means the integral is evaluated along the path $\alpha = 0$, it the physical path. The variation is clearly composed of two parts:

1. The part arising from the change in the limit of integration, which to first order infinitesimals is

$$\mathsf{L}(t_2)\Delta t_2 - \mathsf{L}(t_1)\Delta t_1$$

2. The part which comes from the change in the integrand along the varied path[1]

$$\int_{t_1}^{t_2} \delta\mathsf{L}\,\mathrm{d}t$$

So

$$\Delta \int_{t_1}^{t_2} \mathsf{L}\,\mathrm{d}t = \mathsf{L}(t_2)\Delta t_2 - \mathsf{L}(t_1)\Delta t_1 + \int_{t_1}^{t_2} \delta\mathsf{L}\,\mathrm{d}t$$

Looking at the third term

$$\int_{t_1}^{t_2} \delta\mathsf{L}\,\mathrm{d}t = \sum_i \int_{t_1}^{t_2} \left\{ \frac{\partial \mathsf{L}}{\partial q_i} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathsf{L}}{\partial \dot{q}_i}\right)\right\} \delta q_i\,\mathrm{d}t + \left[\frac{\partial \mathsf{L}}{\partial \dot{q}_i}\delta q_i\right]_{t_1}^{t_2}$$

................ Now

$$\delta \int_{t_1}^{t_2} f(q_i, p_i, \dot{q}_i, \dot{p}_i, t)\,\mathrm{d}t = 0$$

[1]Take care: $\delta q_1(t_1)$ and $\delta q_2(t_2) \neq 0$ necessarily

implies (via the Euler-Lagrange Equations) that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial f}{\partial \dot{q}_i}\right) - \frac{\partial f}{\partial q_i} = 0 \tag{3.10}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial f}{\partial \dot{p}_i}\right) - \frac{\partial f}{\partial p_i} = 0 \tag{3.11}$$

So if we identify $f(q_i, p_i, \dot{q}_i, \dot{p}_i, t) = \sum_i p_i \dot{q}_i - \mathsf{H}(q_i, p_i, t)$ we have

$$\dot{p}_j = -\frac{\partial \mathsf{H}}{\partial q_j} \tag{3.12}$$

$$\dot{q}_j = \frac{\partial \mathsf{H}}{\partial p_j} \tag{3.13}$$

Ie. We have recovered Hamilton's Equations.

**Remarks 2.** *The two Variational principles*

1. *We have two forms of Hamilton's principle:*
   - *In* Coordinate Space $\delta \int_{t_1}^{t_2} \mathsf{L}(q_i, \dot{q}_i, t)\,\mathrm{d}t = 0$*, and we require only that* $\delta q_i = 0$
   - *In* phase Space $\delta \int_{t_1}^{t_2} \sum p_i \dot{q}_i - \mathsf{H}(q_i, p_i, t)\,\mathrm{d}t = 0$*, and we require that* $\delta q_i = \delta p_i = 0$*. Here we treat* $q_i$*,* $p_i$ *as independent variables.*

   Both principles give us Hamilton's Equations

2. $\sum p_i \dot{q}_i - \mathsf{H}(q_i, p_i, t)$ *is independent of* $\dot{p}_i$ *and in our derivation we need the* $\delta$*-terms at the endpoints to vanish, so that we can dispose of the surface terms when we integrate by parts. (Because our integrand is independent of* $\dot{p}_i$ *we don't actually make use of the condition* $\delta p_i = 0$*.*
   At no stage *in the variational derivation do we make use of our original defining equation*

   $$p_i = \frac{\partial \mathsf{L}}{\partial \dot{q}_i}.$$

   *That is to say that neither of the coordinates* $q_i$*,* $p_i$ *is more fundamental.*

3. *Suppose* $F(q, p, t)$ *is an arbitrary twice diffable function of* $p$*,* $q$ *Then if we add* $\frac{\mathrm{d}F}{\mathrm{d}t}$ *to the integrand* $\sum p_i \dot{q}_i - \mathsf{H} + \frac{\mathrm{d}F}{\mathrm{d}t}$ *then the variational principle is unaltered.*

We can apply Lagrange's Equations

$$\int_{t_1}^{t_2} \delta \mathsf{L}\,\mathrm{d}t = \sum \frac{\partial \mathsf{L}}{\partial \dot{q}_i}\delta q_i \Big|_{t_1}^{t_2}$$

The $\delta q_i$ refers to the variation in $q_i$ at the original end points:

$$\Delta \int_{t_1}^{t_2} \mathsf{L}\,\mathrm{d}t = \sum \left(\mathsf{L}\Delta t + p_i \delta q_i\right)\Big|_{t_1}^{t_2}$$

Now

$$\Delta q_i(t_2) = q_i(t_2 + \Delta t_2, \alpha) - q_1(t_2, 0) \tag{3.14}$$
$$= q_i(t_2 + \Delta t_2, 0) - q_i(t_2, 0) + \alpha\eta_i(t_2 + \Delta t_2) \tag{3.15}$$

So to first order in $\alpha$ and $\Delta t_2$ we have

$$\Delta q_i(t_2) = \dot{q}_i(t_2)\Delta t_2 + \delta q_i(t_2)$$

Hence

$$\Delta \int_{t_1}^{t_2} \mathsf{L}\,\mathrm{d}t = \sum (\mathsf{L}\Delta t - p_i\dot{q}_i\Delta t + p_i\Delta q_i)\Big|_{t_1}^{t_2} \tag{3.16}$$
$$= \sum p_i\Delta q_i - \mathsf{H}\Delta t\Big|_{t_1}^{t_2} \tag{3.17}$$

We now make the following assumptions:

1. The only systems we consider are those such that $\frac{\partial \mathsf{L}}{\partial t} = \frac{\partial \mathsf{H}}{\partial t} = 0$.
2. The variation is st. $\mathsf{H}$ is conserved on the varied path.

3. The varied paths are st. $\Delta q_i = 0$ at the end points.

**Remark 4.** *The varied path might even describe the* same *path in* configuration space *as the* actual *path, the difference is in the* speed *the system point transverses the curve.*

Now given the above qualifications we have

$$\Delta \int_{t_1}^{t_2} \mathsf{L}\,\mathrm{d}t = -\mathsf{H}(\Delta t_2 - \Delta t_1)$$

But under the same cosiderations

$$\int_{t_1}^{t_2} \mathsf{L}\,\mathrm{d}t = \sum \int_{t_1}^{t_2} pi\dot{q}_i\,\mathrm{d}t - \mathsf{H}(t_2 - t_1)$$

and

$$\Delta \int_{t_1}^{t_2} \underbrace{\sum p_i\dot{q}_i}_{\text{`action'}}\,\mathrm{d}t = 0$$

**Remark 5.** *In older books the quantity $\sum p_i\dot{q}_i$ is called the action. For us, however the action is* $\mathsf{L}$.

## 3.2 Canonical Transformations

### 3.2.1 Canonical Transformations

*Canonical Transformations* (henceforth *CT*s) are those that leave the Hamiltonian structure of the system invariant. Suppose we started with a Hamiltonian $\mathsf{H}(q, p, t)$ satisfying Hamilton's Equations:

$$\dot{q}_i = \frac{\mathrm{d}\mathsf{H}}{\mathrm{d}p_i}, \quad \dot{p}_i = -\frac{\partial \mathsf{H}}{\partial q_i} \iff \delta \int (p\dot{q} - \mathsf{H})\,\mathrm{d}t = 0$$

exclude the triv- case of the 'scale formation' where $\lambda_i P_i$, $q_i = \lambda_i Q_i$ for scalars $\lambda_i$.

Suppose there exist functions $P(p, q, t)$, $Q(p, q, t)$ and an associated Hamiltonian, $\mathsf{K}(P, Q, t)$, not necessarily the transformation of $\mathsf{H}$ into the new coordinates, st.

$$\frac{\partial K}{\partial P} = \dot{Q}, \qquad \frac{\partial K}{\partial Q} = -\dot{P} \iff \delta \int (P\dot{Q} - \mathsf{K})\,\mathrm{d}t = 0$$

We ask, how are the new coordinates related to the old ones? Consider the function $F$ defined such that $P\,\mathrm{d}Q - \mathsf{K}\,\mathrm{d}t + \mathrm{d}F = p\,\mathrm{d}q - \mathsf{H}\,\mathrm{d}t$. Ie. we want

$$\mathrm{d}F = p\,\mathrm{d}q - P\,\mathrm{d}Q + (\mathsf{K} - \mathsf{H})\ \mathrm{d}t$$

Let us take the particular case where $F = F_1(q, Q, t)$ then we have that

$$\mathrm{d}F_1 = \frac{\partial F_1}{\partial q}\,\mathrm{d}q + \frac{\partial F_1}{\partial Q}\,\mathrm{d}Q + \frac{\partial F_1}{\partial t}\,\mathrm{d}t \tag{3.18}$$

$$\Rightarrow \quad \frac{\partial F_1}{\partial q} = p, \qquad \frac{\partial F_1}{\partial Q} = -P, \qquad \frac{\partial F_1}{\partial t} = \mathsf{K} - \mathsf{H} \tag{3.19}$$

So

$$\delta \int_{t_1}^{t_2} (p\dot{q} - \mathsf{H})\ \mathrm{d}t = \delta \int \left(P\dot{Q} - \mathsf{K}\right)\ \mathrm{d}t + \delta F_1|_{t_1}^{t_2}$$

And so the transformation is Canonical.

$F_1$ acts as a bridge between the old and then new coordinates. Half the variables are from the old and half from the new. For example suppose $F_1 = qQ$ then $p = Q$, $P = -q$ and $\mathsf{K} = \mathsf{H}$.

Now there is no need for $F$, the *generating function*, to be a function of $q$, $Q$ and $t$. for example consider $F_2 = F_2(q, P, t)$ and define $F$ by

$$F = F_2(q, P, t) - QP$$

Then

$$p\dot{q} - \mathsf{H} \;=\; P\dot{Q} - \mathsf{K}\frac{\mathrm{d}F}{\mathrm{d}t} \tag{3.20}$$

$$\;=\; -\dot{P}Q - \mathsf{K} + \frac{\mathrm{d}F_2(q, P\,t)}{\mathrm{d}t} \tag{3.21}$$

But

$$\frac{\mathrm{d}}{\mathrm{d}t}F_2(q, P, t) = \frac{\partial F_2}{\partial q}\,\mathrm{d}q + \frac{\partial F_2}{\partial P}\,\mathrm{d}P + \frac{\partial F_2}{\partial t}\,\mathrm{d}t \tag{3.22}$$

$$\Rightarrow \quad p = \frac{\partial F_2}{\partial q}, \qquad Q = \frac{\partial F_2}{\partial P}, \qquad \mathsf{K} = \mathsf{H} + \frac{\partial F_2}{\partial t} \tag{3.23}$$

We can also devise generating functions which are mixed in the sense that they depend only on $p, Q, t$ or $p, P, t$ Define $F = qp + F_3(p, Q, t)$ then similarly from above we will have

$$q = -\frac{\partial F_3}{\partial p}, \qquad P = -\frac{\partial F_3}{\partial Q}, \qquad \mathsf{K} = \mathsf{H} + \frac{\partial F_3}{\partial t}$$

And, finally, defining $F = qp - QP + F_4(p, P, t)$ we get

$$q = -\frac{\partial F_4}{\partial p}, \qquad Q = \frac{\partial F_4}{\partial P}, \qquad \mathsf{K} = \mathsf{H} + \frac{\partial F_4}{\partial t}$$

These four generating functions are all we need to describe a specific change.

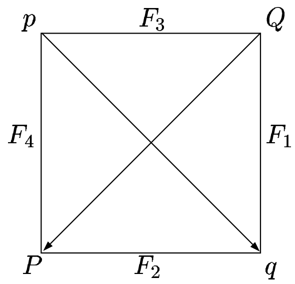


Figure 3.1: The functional dependence of the generating functions $F_1$ to $F_4$.

### 3.2.2 Generalisation to Higher Dimensions

Now we can generalise this to $n$-dimensions:

$$\text{Take} \quad F_1(q_i, Q_i, t), \qquad p_i = \frac{\partial F}{\partial q_i}, \qquad P_i = -\frac{\partial F}{\partial Q_i} \tag{3.24}$$

$$\text{then} \quad F = F_2(q_i, P_i, t) - \sum Q_i P_i \Rightarrow p_i = \frac{\partial F_2}{\partial q_i}, \qquad Q_i = \frac{\partial F_2}{\partial P_i} \tag{3.25}$$

$$\text{etc.} \tag{3.26}$$

In the general case,the generating function does *not* have to conform to *one* of the four general types for *all* degrees of freedom.

**Example 3.** *Let us consider the Harmonic Oscillator*

$$\mathsf{H} = \frac{p^2}{2m} + \frac{kq^2}{2} = \frac{1}{2}m\left(p^2 + m^2\omega^2 q^2\right), \ \textit{with}\ \omega^2 = \frac{k}{m}$$

*Let us take*

$$p = f(P)\cos Q \tag{3.27}$$

$$q = \frac{F(P)\sin Q}{m\omega} \tag{3.28}$$

*We require*

$$\mathsf{H} = \mathsf{K} = \frac{F^2(P)}{2m}\left(\cos^2 Q + \sin^2 Q\right) = \frac{F^2(P)}{2m}$$

*So this choice has made $Q$ cyclic.*
*We want to find the function $f$ as yet unspecified, such that the transformation is canonical. To do this we observe*

$$\frac{p}{q} = m\omega \cot Q \tag{3.29}$$

$$\Rightarrow \quad p = m\omega q \cot Q \tag{3.30}$$

*Now suppose there exists a function of the type $F_1(q, Q)$ then*

$$p = \frac{\partial F_1}{\partial q}(q, Q) = m\omega q \cot Q$$

*The simplest solution to this is*

$$F_1 = \frac{m\omega q^2}{2}\cot Q$$

*And then we must have that*

$$P = -\frac{\partial F_1}{\partial Q} = \frac{m\omega q^2}{2\sin^2 Q} \Rightarrow q = \sqrt{\frac{2P}{m\omega}}\sin Q$$

*but* $p = F(P)\cos Q \Rightarrow f(P) = \sqrt{2m\omega P}$. *And we have that*

$$\mathsf{H} = \mathsf{K} = \omega P$$

$\mathsf{K}$ *is cyclic in* $Q \Rightarrow \dot{P} = 0 \Rightarrow P = const$. *This in turn* $\Rightarrow \mathsf{K} = const = E$. *And hence* $P = \frac{E}{\omega}$.
*Then*

$$\dot{Q} = \frac{\partial \mathsf{K}}{\partial P} = \omega \Rightarrow Q = \omega t + \alpha$$

*Where alpha is a constant determined from the initial conditions. We can now invert the transformation to obtain*

$$q = \sqrt{\frac{2E}{m\omega^2}}\sin(\omega t + \alpha)$$

In going from $P = \frac{m\omega q^2}{2\sin^2 Q}$ to $q$ we have used the positive square root. We could have used the negative root: The only difference would have been a trivial difference of $\pi$ in the phase:*Transformations need not be single valued*

### 3.2.3 Poisson Brackets

Let $f(p, q, t)$ be any function of $q$, $p$. Then

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial q}\frac{\mathrm{d}q}{\mathrm{d}t} + \frac{\partial f}{\partial p}\frac{\mathrm{d}p}{\mathrm{d}t} + \frac{\partial f}{\partial t}$$

But $\dot{q} = \frac{\partial \mathsf{H}}{\partial p}$ and $\dot{p} = -\frac{\partial \mathsf{H}}{\partial q}$ so we can write:

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial q}\frac{\partial \mathsf{H}}{\partial p} - \frac{\partial f}{\partial p}\frac{\partial \mathsf{H}}{\partial q} + \frac{\partial f}{\partial t} \quad (3.31)$$

$$= [f, \mathsf{H}] + \frac{\partial f}{\partial t} \quad (3.32)$$

where $[f, \mathsf{H}] \equiv \frac{\partial f}{\partial q}\frac{\partial \mathsf{H}}{\partial p} - \frac{\partial f}{\partial p}\frac{\partial \mathsf{H}}{\partial q}$. This is a *Poisson Bracket*.

**Definition:** If $f = f(q_i, p_i, t)$ and $g = g(q_i, p_i, t)$ with $1 \le i \le n$, then we define the *Poisson Bracket* by

$$[f, g]_{q,p} = \sum_i \frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i}$$

It has the following properties:

- $[f,f]_{q,p} = 0$
- (antisymmetry) $[f,g]_{q,p} = -[g,f]_{q,p}$
- (linearity) $[af+bg,h]_{q,p} = a[f,h]_{q,p} + b[g,h]_{q,p}$.
- $[p_i, \mathsf{H}]_{q,p} = \dot{p}_i$ and $[q_i, \mathsf{H}]_{q,p} = \dot{q}_i$.

  *Proof.* To see this we use $\dot{q}_i = \frac{\partial \mathsf{H}}{\partial p_i}$ and $\dot{p}_i = -\frac{\partial \mathsf{H}}{\partial q_i}$ so that

$$[p_i, \mathsf{H}]_{q,p} = \sum_j \frac{\partial p_i}{\partial q_j}\frac{\partial \mathsf{H}}{\partial p_j} - \frac{\partial p_i}{\partial p_j}\frac{\partial \mathsf{H}}{\partial q_j} = \dot{p}_i \tag{3.33}$$

$$\text{and} \quad [q_i, \mathsf{H}]_{q,p} = \sum_j \frac{\partial q_i}{\partial q_j}\frac{\partial \mathsf{H}}{\partial p_j} - \frac{\partial q_i}{\partial p_j}\frac{\partial \mathsf{H}}{\partial q_j} = \dot{q}_i \tag{3.34}$$

  similarly. □

- $[q_i, q_j]_{q,p} = [p_i, p_j]_{q,p} = 0$.

  *Proof.* Observe

$$[q_i, q_j]_{q,p} = \sum_k \frac{\partial q_i}{\partial q_k}\frac{\partial q_j}{\partial p_k} - \frac{\partial q_i}{\partial p_k}\frac{\partial q_j}{\partial q_k} \tag{3.35}$$

$$= \sum \delta_{ij} - \delta_{ij} = 0 \tag{3.36}$$

  □

- $[q_i, p_j]_{q,p} = \delta_{ij}$

  *Proof.*

$$[q_i, p_j]_{q,p} = \sum_k \frac{\partial q_i}{\partial q_k}\frac{\partial p_j}{\partial p_k} - \frac{\partial q_i}{\partial p_k}\frac{\partial p_j}{\partial q_k} \tag{3.37}$$

$$= \sum \delta_{ik}\delta_{jk} \tag{3.38}$$

$$= \delta_{ij} = -[p_j, q_i] \tag{3.39}$$

  □

- If $\frac{\partial \mathsf{H}}{\partial t} = 0$ the Hamiltonian is conserved.

  *Proof.* Since $\frac{\mathrm{dH}}{\mathrm{d}t} = [\mathsf{H},\mathsf{H}] + \frac{\partial \mathsf{H}}{\partial t}$, then if $\frac{\partial \mathsf{H}}{\partial t} = 0 \Rightarrow \frac{\mathrm{dH}}{\mathrm{d}t} = 0$ and the Hamiltonian is conserved. □

- (The Jacobi Identity): $[f,[g,h]] + [g,[h,f]] + [h,[f,g]] = 0$.

  *Proof.* Left to example sheet □

Antisymmetry, Linearity and the Jacobi property define what is called a *Lie Algebra*, wherein the Poisson Bracket is the 'product'. Other Lie algebras include the vector product and the matrix commutator. The QM correspondence principle says that: $[f,g] \rightarrow \frac{1}{i\hbar}(fg - gf)$ and this only works because both

## 3.3 The Sympletic Condition and CTs

### 3.3.1 The Special Case

We are looking for conditions that make a given transformation canonical. Let us begin by considering some restricted canonical transformations: Those where time does not play an explicit part. In terms of the generating function $F$ we have that $\frac{\partial F}{\partial t} = 0$ so that $\mathsf{K} = \mathsf{H}$. Then

$$Q = Q(q,p), \qquad P = P(q,p)$$

So we have

$$\dot{Q}_i = \sum_j \frac{\partial Q_i}{\partial q_j}\dot{q}_j + \frac{\partial Q_i}{\partial p_j}\dot{p}_j \tag{3.40}$$

$$= \sum_j \frac{\partial Q_i}{\partial q_j}\frac{\partial zH}{\partial p_j} - \frac{\partial Q_i}{\partial p_j}\frac{\partial \mathsf{H}}{\partial q_j} \tag{3.41}$$

On the other hand we may invert the transformation to get $q_j = q_j(Q,P)$ and $p_j = p_j(Q,P)$ as

$$\frac{\partial \mathsf{H}}{\partial P_i} = \sum_j \frac{\partial \mathsf{H}}{\partial p_j}\frac{\partial p_j}{\partial P_i} + \frac{\partial \mathsf{H}}{\partial q_j}\frac{\partial q_j}{\partial P_i}$$

Then the transformation is canonical *only if*

$$\dot{Q}_i = \frac{\partial \mathsf{H}}{\partial P_i} \tag{3.42}$$

$$\Rightarrow \quad \underbrace{\left(\frac{\partial Q_i}{\partial q_j}\right)_{q,p}}_{Q \text{ as a function of } q,\, p} = \underbrace{\left(\frac{\partial p_i}{\partial P_j}\right)_{Q,P}}_{p \text{ as a function of } Q,\, P} \tag{3.43}$$

$$\text{and} \quad \left(\frac{\partial Q_i}{\partial p_j}\right)_{q,p} = -\left(\frac{\partial q_i}{\partial P_j}\right)_{Q,P} \tag{3.44}$$

In the same way, by considering $\dot{P}_i$ we find that

$$\left(\frac{\partial P_i}{\partial q_j}\right)_{q,p} = -\left(\frac{\partial p_j}{\partial Q_i}\right)_{Q,P}, \qquad \left(\frac{\partial P_i}{\partial p_j}\right)_{q,p} = \left(\frac{\partial Q_j}{\partial Q_i}\right)_{Q,P}$$

Let us further restrict our attention to a 2 dimensional phase space $(q,p)$ and consider the transformation to $Q, P$, then define

$$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad M = \begin{pmatrix} \frac{\partial Q}{\partial q} & \frac{\partial Q}{\partial p} \\ \frac{\partial P}{\partial q} & \frac{\partial P}{\partial p} \end{pmatrix} = \begin{pmatrix} \partial_q Q & \partial_p Q \\ \partial_q P & \partial_p P \end{pmatrix}$$

Consider

$$MJM^{\mathrm{T}} = \begin{pmatrix} \frac{\partial Q}{\partial q} & \frac{\partial Q}{\partial p} \\ \frac{\partial P}{\partial q} & \frac{\partial P}{\partial p} \end{pmatrix}\begin{pmatrix} \partial_p Q & \partial_p P \\ -\partial_q Q & -\partial_q P \end{pmatrix} \tag{3.45}$$

$$= \begin{pmatrix} 0 & [Q,P]_{q,p} \\ [P,Q]_{q,p} & 0 \end{pmatrix} \tag{3.46}$$

But consider

$$[Q,P]_{q,p} = \frac{\partial Q}{\partial q}\frac{\partial P}{\partial p} - \frac{\partial P}{\partial q}\frac{\partial Q}{\partial p},$$

and in particular for our transformation $\partial_q Q = \partial_P p$ etc. So

$$[Q,P]_{q,p} = \partial_P p \partial_Q q - \partial_P p \partial_Q q = [q,p]_{Q,P}$$

Now *if* it were that $[u,v]_{q,p} = [u,v]_{Q,P}$, then it would follow that the condition for the restricted transformation to be canonical would be that

$$MJM^{\mathrm{T}} = J.$$

We will now show that this is indeed the condition that the transformation is canonical: We will show that a *CT leaves Poisson brackets invariant*. So it is unnecessary to write $[f,g]_{q,p}$, since the quantity is the same no matter what canonical coordinates we choose.

**Notation 2.** *We shall write* $M_{ij} = \frac{\partial \xi_i}{\partial \eta_j}$ *where*

$$\boldsymbol{\eta} = \begin{pmatrix} q_1 \\ \vdots \\ q_n \\ P_1 \\ \vdots \\ P_n \end{pmatrix}, \qquad \boldsymbol{\xi} = \begin{pmatrix} Q_1 \\ \vdots \\ Q_n \\ P_1 \\ \vdots \\ P_n \end{pmatrix}, \qquad J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$$

*Ie. both are column vectors in a* $2n$*-dimensional space, and where* $I$ *is the* $n \times n$ *identity matrix. This might seem a bit odd, but fundamentally, the* $(q_i, p_i)$ *are our coordinates: Why* should *we treat them differently?*

**theorem 1.** *A transformation* $q,p \longrightarrow Q,P$ *is canonical iff*

$$M^{\mathrm{T}}JM = MTM^{\mathrm{T}} = J.$$

*This is called the* Sympletic Condition*.*

**Remark 6.** *We prove the result first in the case when we have a restricted CT, ie. no explicit time dependence, and then go on to prove it for the general case*

It might be useful, at each step, to write out the matricies and vectors explicitly.

Let us prove the special case:

*Proof.* First, it follows from Hamilton's Equations that

$$\dot{\boldsymbol{\eta}} = J\frac{\partial \mathsf{H}}{\partial \boldsymbol{\eta}}$$

The elements of $\boldsymbol{\xi}$ are the $Q_i$ and $P_i$, and these are functions of the $q_i$ and $p_i$, ie. of $\boldsymbol{\eta}$ so

$$\dot{\xi}_i = \sum_{j=1}^{2n} \frac{\partial \xi_i}{\partial \eta_j}\dot{\eta}_j \tag{3.47}$$

$$\Rightarrow \quad \dot{\boldsymbol{\xi}} = M\dot{\boldsymbol{\eta}} \tag{3.48}$$

Now, by the inverse transformation H can be considered as a function of $\xi$ and $\eta$, so

$$\frac{\partial \mathsf{H}}{\partial \eta_i} = \sum_j \frac{\partial \mathsf{H}}{\partial \xi_j} \frac{\partial \xi_j}{\partial \eta_i}$$

Or in vector/matrix notation

$$\frac{\partial \mathsf{H}}{\partial \boldsymbol{\eta}} = M \frac{\partial \mathsf{H}}{\partial \boldsymbol{\xi}}$$

Now

$$\dot{\boldsymbol{\xi}} = M\dot{\boldsymbol{\eta}} = MJ\frac{\partial \mathsf{H}}{\partial \boldsymbol{\eta}} = MJM^{\mathrm{T}}\frac{\partial \mathsf{H}}{\partial \boldsymbol{\xi}}$$

But

$$\dot{\boldsymbol{\xi}} = J\frac{\partial \mathsf{H}}{\partial \boldsymbol{\xi}}$$

from Hamilton's Equations. Hence

$$MJM^{\mathrm{T}} = J$$

By noting that we could just as well have gone from $Q, P$ to $p, q$ we must also have

$$M^{\mathrm{T}}JM = J.$$

□

This proves the special case: That a restricted CT will be canonical if the symplectic condition holds. The reverse condition holds, and fortunately the proof works in reverse too.

**Remark 7.** *In our example we saw that*

$$M^{\mathrm{T}}JM = \begin{pmatrix} 0 & [Q,P]_{q,p} \\ [Q,P]_{q,p} & 0 \end{pmatrix}$$

*We have just proved, however, that*

$$M^{\mathrm{T}}JM = J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

*And hence* $[Q,P]_{q,p} = -(P,Q)_{q,p} = 1$*, but it was a trivial property of the Poisson Bracket that* $[Q,P]_{Q,P} = 1 = -[Q,P]$ *so we clearly have that*

$$[Q,P]_{Q,P} = [Q,Q]_{q,p} = [P,P]_{q,p} = [P,P]_{Q,P} = 0$$

*This also holds in the general case, so that*

$$[Q_i,P_j]_{q,p} = [Q_i,P_j]_{Q,P} = \delta_{ij} \tag{3.49}$$
$$[Q_i,Q_j] = [P_i,P_j] = 0 \tag{3.50}$$

$[Q,P]_{q,p} = 1$ is precisely the same as the statement that the Jacobian of the CT is 1, which in turn implies that $\int \mathrm{d}q\,\mathrm{d}p$ is an invariant of the transformation.??? This is one of the Poincarre integral invariants

### 3.3.2 The General Case

We have yet to show that $MJM^{\mathrm{T}} = JM^{\mathrm{T}}JM$ is a necessary and sufficient condition for a CT in the general case of an arbitrary CT. Ie one with time dependence.
IN order to do this it will be necessary to introduce the important idea of an *infinitesimal contact/canonical transformation* (or an *ICT* for short).

**Lemma 1.** *The generating function $F_2 = \sum q_i P_i$ generates the identity transformation.*

*Proof.*

$$\frac{\partial F_2}{\partial q_i} = p_i \Rightarrow p_i = P_i, \text{ and } \frac{\partial F_2}{\partial P_i} = q_i \Rightarrow q_i = Q_i$$

□

In fact CTs form a group: The identity is canonical, the inverse of a CT is a CT, two successive CTs is a CT, and the product is associative. We only require that the CTs are analytic functions of cts parameters in order to have a *Lie Group*.

Now let us consider the generating function

$$F = \sum q_i P_i + \varepsilon G(q_i, P_i, t)$$

where $G$ is an arbitrary diffable function and $\varepsilon$ is infinitesimally small. Then

$$Q_i = \frac{\partial F}{\partial P_i} = q_i + \varepsilon \frac{\partial G}{\partial P_i} \tag{3.51}$$

$$p_i = \frac{\partial F}{\partial q_i} = P_i + \varepsilon \frac{\partial G}{\partial q_i} \tag{3.52}$$

Now take

$$\delta p_i = P_i - p_i = -\varepsilon \frac{\partial G}{\partial q_i} \tag{3.53}$$

$$\delta q_i = Q_i - q_i = \varepsilon \frac{\partial G}{\partial P_i} \tag{3.54}$$

Now $G(q_i, P_i(q,p)) = G(q_i, p_i + \varepsilon f(q,p)) \Rightarrow$ (to first order in $\varepsilon$)

$$\delta q_i = \varepsilon \frac{\partial G}{\partial p_i} \tag{3.55}$$

$$\delta p_i = -\varepsilon \frac{\partial G}{\partial q_i} \tag{3.56}$$

$$\Rightarrow \delta \boldsymbol{\eta} = \varepsilon J \frac{\partial G}{\partial \boldsymbol{\eta}} \tag{3.57}$$

Now $M = \frac{\partial \boldsymbol{\xi}}{\partial \boldsymbol{\eta}} = I + \frac{\partial}{\partial \boldsymbol{\eta}} \delta \boldsymbol{\eta} = I + \varepsilon J \frac{\partial^2 G}{\partial \boldsymbol{\eta}^2}$ So the second derivative is a square matrix: $\left(\frac{\partial^2 G}{\partial \boldsymbol{\eta}^2}\right)_{ij} = \frac{\partial^2 G}{\partial \eta_i \partial \eta_j}$. Now $M = I - \varepsilon \left(J \frac{\partial^2 G}{\partial \boldsymbol{\eta}^2}\right)^{\mathrm{T}}$, but $\frac{\partial^2 G}{\partial \boldsymbol{\eta}^2}$ is symmetric, and J is antisymmetric, $\Rightarrow$

$$M^{\mathrm{T}} = I - \varepsilon \frac{\partial^2 G}{\partial \boldsymbol{\eta}^2} J.$$

And hence

$$M^{\mathrm{T}}JM = \left(I + \varepsilon J\frac{\partial^2 G}{\partial \boldsymbol{\eta}^2}\right) J \left(I - \varepsilon \frac{\partial^2 G}{\partial \boldsymbol{\eta}^2} J\right) \tag{3.58}$$

$$= J + \varepsilon J \frac{\partial^2 G}{\partial \boldsymbol{\eta}^2} J - \varepsilon \frac{\partial^2 G}{\partial \boldsymbol{\eta}^2} J + \mathcal{O}(\varepsilon^2) \tag{3.59}$$

Thus to first order

$$M^{\mathrm{T}}JM = J$$

$\Rightarrow$ for any ICT the symplectic condition holds.
Now consider the CT $\xi = \xi(\boldsymbol{\eta}, t)$, this evolves ctsly as time increases from some initial value. Let $G = \mathsf{H}(q, p, t)$. Then

$$\delta q_i = \mathrm{d}t \frac{\partial \mathsf{H}}{\partial p_i} = \dot{q}_i \, \mathrm{d}t = \mathrm{d}q_i \tag{3.60}$$

$$\delta p_i = -\mathrm{d}t \frac{\partial \mathsf{H}}{\partial q_i} = \dot{p}_i \, \mathrm{d}t = \mathrm{d}p_i \tag{3.61}$$

Thus the Hamiltonian acts as the generator of an ICT which corresponds to the evolution in time of the system $\longrightarrow$ symplectic condition holds.
The continuous evolution of the transformation $\xi(\boldsymbol{\eta}, t)$ from $\xi(\boldsymbol{\eta}, t_0)$ to $\xi(\boldsymbol{\eta}, t)$ can be built up as a succession of ICTs in steps of $\mathrm{d}T$, so if $\eta(t_0) \rightarrow \xi(t_0)$ is canonical as $\xi(t_0) \rightarrow \xi(t)$ is canonical we must have that $\eta(t_0) \rightarrow \xi(t)$ is canonical.
It can be shown that the product of two successive CTs is a CT (question 1, problem sheet 4). So the symplectic condition holds in general. $\square$ In the course of our argument we have seen that

$$[P_i, Q_j] = [p_i, q_j] = -\delta ij,$$

where the Poisson bracket is evaluated wrt. the canonical set. It can be shown that if $u$, $v$ are arbitrarily diffable functions of $q$, $p$ then

$$[u, v]_{q,p} = [u, v]_{Q,P}$$

Ie. *all Poisson brackets are invariant*. (question 2, problem sheet 4).

## 3.4 More on ICTs

### 3.4.1 The Hamiltonian as the generator of an ICT

As we've seen, an ICT is a special case of a transformation that is a cts function of a parameter. If the parameter is small enough to be treated as a function of a first order infinitesimal then the transformation between canonical variables differ only in infinitesimals, ie.

$$\boldsymbol{\xi} = \boldsymbol{\eta} + \delta\boldsymbol{\eta}$$

with the change being given in terms of the generator G through the equation

$$\delta\boldsymbol{\eta} = \varepsilon J \frac{\partial G(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}}.$$

Now

$$[\boldsymbol{\eta}, u] = \left(\frac{\partial \eta_i}{\partial \boldsymbol{\eta}}\right)^{\mathrm{T}} J \frac{\partial u}{\partial \boldsymbol{\eta}} \tag{3.62}$$

$$= J \frac{\partial u}{\partial \boldsymbol{\eta}} \tag{3.63}$$

$$\Rightarrow \quad \delta \boldsymbol{\eta} = \varepsilon \, [\boldsymbol{\eta}, G] \tag{3.64}$$

Now consider an ICT in $t$ whose generator *is the Hamiltonian*:

$$\delta \boldsymbol{\eta} = \mathrm{d}t \, [\boldsymbol{\eta}, \mathsf{H}] \tag{3.65}$$

$$\dot{\boldsymbol{\eta}} = [\boldsymbol{\eta}, \mathsf{H}] \tag{3.66}$$

If the motion of the system in a time interval $\mathrm{d}t$ can be described by an ICT generated by the Hamiltonian $\longrightarrow$ The motion of the system from $t_0$ to $t$ can be generated by a single contact transformation equivalent to an infinite sequence of infinitesimals, all generated by the Hamiltonian.
We can view the Hamiltonian as the generator of an ICT (and consequently a CT) which describes the motion of the system with time.
A solution of the problem of finding the canonical transformation which relates coordinates and momenta at time $t = 0$ to their value at $t = t$ is equivalent to solving the physical problem. There are two views:

- **The passive View** We regard the transformation from $\boldsymbol{\xi} \longrightarrow \boldsymbol{\eta}$ as mapping from one phase space to a new phase space. So

$$u(p, q) \longrightarrow U(P, Q),$$

ie. the functional form will change, but not the value.

- **The Active View** We regard the transformation as a mapping within one phase space - a *point transformation*. This time the functional form remains the same, but the value changes. For example, the evolution of the system in time, as generated by the Hamiltonian.

This is the same active view that is used to define the *Lie Derivative* of a tensor field

If we are working in the active sense then we can talk about the change in the function $u$ under a CT (cf. the passive view, which has $u(p, q) = U(P, Q)$.)

### 3.4.2 Symmetry and Conserved Quantities

Suppose we have an infinitesimal transformation generated by $G(q, p, t)$, so that

$$u(\boldsymbol{\eta} + \delta\boldsymbol{\eta}) - u(\boldsymbol{\eta}) = \frac{\partial \mathbf{u}}{\partial \boldsymbol{\eta}} \delta\boldsymbol{\eta} = \varepsilon \frac{\partial u}{\partial \boldsymbol{\eta}} J \frac{\partial G}{\partial \boldsymbol{\eta}}$$

then since $\boldsymbol{\eta} = (q_1, ..., q_n, p_1, ..., p_n)^{\mathrm{T}}$ the above becomes

$$u(\boldsymbol{\eta} + \delta\boldsymbol{\eta}) - u(\boldsymbol{\eta}) = \sum \frac{\partial u}{\partial q_i} \delta q_i + \frac{\partial u}{\partial p_i} \delta p_i \tag{3.67}$$

$$= \varepsilon \sum \left( \frac{\partial u}{\partial q_i} \frac{\partial G}{\partial p_i} - \frac{\partial u}{\partial p_i} \frac{\partial u}{\partial q_i} \right) \tag{3.68}$$

$$= \varepsilon \, [u, G] \tag{3.69}$$

So assuming that $G$ is not an explicit function of time, we can ask 'does $\mathsf{K} = \mathsf{H}$?' Well

$$\delta\mathsf{H} = \varepsilon\,[\mathsf{H}, G]$$

But $\frac{\mathrm{d}G}{\mathrm{d}t} = [\mathsf{H}, G] + \underbrace{\frac{\partial G}{\partial t}}_{=0}$, so

$$\boxed{\frac{\mathrm{d}G}{\mathrm{d}t} = 0 \iff [\mathsf{H}, G] = 0}$$

Compare this to Heisenberg's formulation of QM.

*The symmetry properties of the system are equivalent to the conservation laws.*
The statement now includes *all* constants, not just the conjugate momenta to cyclic variables.

**Example 4.** *Suppose $q_i$ is cyclic. Ie. the Hamiltonian is independent of $q_i$, and will be clearly invariant under an ICT which involves $q_i$ alone. The equations of transformation would be*

$$\delta q_j = \varepsilon\delta_{ij} \tag{3.70}$$
$$\delta p_j = 0 \tag{3.71}$$

*And then $G = p_i$, so*

$$\Delta\mathsf{H} - \varepsilon\,[\mathsf{H}, p_j] = 0 \Rightarrow \frac{\mathrm{d}p_j}{\mathrm{d}t} = 0$$

## 3.5 The Hamilton-Jacobi Equation

### 3.5.1 'Nice' coordinates

Having done all this abstract theory, we can now reap some benefits. We have two approaches to solving problems:

1. If the Hamiltonian is conserved then we can transform to a new set of canonical coordinates, *all of which are cyclic*. Then the integration of the new set becomes trivial.

2. We can seek a CT from $(q, p)$ at $t = t$ to those at $t = 0$. Under such a transformation the equations linking $(q, p)$ with the new $(q_0, p_0)$ are the solutions to the problem.

Let us consider the first approach:
We can automatically require that our new variables are constant in time if the transformed Hamiltonian is zero, for then from Hamilton's Equations

$$\frac{\partial\mathsf{H}}{\partial P_i} = 0 \Rightarrow \dot{Q}_i = 0 \Rightarrow Q_i = const$$

and similarly for the $P_i$. Now we know that the transformed Hamiltonian $\mathsf{K}$ is related to the old Hamiltonian $\mathsf{H}$ by the equation

$$\mathsf{K} = \mathsf{H} + \frac{\partial F}{\partial t},$$

where $F$ is our generating function. So if we have the new Hamiltonian being zero, we must satisfy

$$\mathsf{H} + \frac{\partial F}{\partial t} = 0$$

It proves convenient to take $F = F_2(q, P, t)$. For then we have, using $p_i = \frac{\partial F_2}{\partial q_i}$,

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

This is the *Hamilton-Jacobi Equation*. It is a pde for the desired generating function in the $(n+1)$ variables $q_1, ..., q_n, t$.
Suppose there exists a solution

$$S = S(q_1, ..q_n, \underbrace{\alpha_1, ..., \alpha_{n+1}}_{\text{constants}}, t).$$

This constitutes a complete solution of the differential equation, and is called *Hamilton's principle Function*. One constant is redundant: $S$ itself only appears in the Hamilton-Jacobi Equation via it's derivatives wrt. $q_i$ and $t$, so wlog. we can add a constant to $S$ and the H-J equation will still hold.
To this end was absorb $\alpha_{n+1}$ into $S$ and look for a solution

$$S = S(q_1, ..q_n, \alpha_1, ..., \alpha_n, t).$$

Once we have the $n$ constants $\{\alpha_i\}_{i=1}^n$ the solution will be complete.
We are at liberty to take the $n$ constants to be the new (constant) momenta $P_i$, that is we can set

$$P_i = \alpha_i.$$

Now recall, for $F_2$-type transformations

$$p_i = \frac{\partial F_2}{\partial q_i}, \qquad Q_i = \frac{\partial F_2}{\partial P_i},$$

but this implies

$$p_i = \frac{\partial S}{\partial q_i}\left(q_i, \alpha_i, t\right),$$

We can evaluate the constants of motion ito. our initial conditions at $t = 0$. Ie. we find the values of $q_i$ simply by calculating $\frac{\partial S}{\partial \alpha_i}$ at $t = 0$. And $Q_i$ is given by

$$Q_i = \beta_i = \frac{\partial S}{\partial \alpha_i}\left(q, \alpha, t\right)$$

?????

**Example 5.** *Consider the Harmonic Oscillator Hamiltonian with unit mass*

$$\mathsf{H} = \frac{1}{2}\left(p^2 + \omega^2 q^2\right) = E, \quad \textit{with } \omega = \sqrt{k}$$

*If we set $p = \frac{\partial S}{\partial q}$ (thus assuming an $F_2$-type transformation) we obtain the H-J equation*

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

*We notice that $S$ depends on time only in the last term of the H-J equation, so let's try $S(q,\alpha,t) = W(q,\alpha) - \alpha t$. Then*

$$\frac{1}{2}\left[\left(\frac{\partial W}{\partial q}\right)^2 + \omega^2 q^2\right] = \underbrace{\alpha}_{-\frac{\partial S}{\partial t}}$$

*And so the H-J equation implies that $\mathsf{H} = \alpha$, and we naturally associate $\alpha$ with the energy.*
*Now*

$$W = \sqrt{2\alpha}\int \mathrm{d}q\sqrt{1 - \frac{\omega^2 q^2}{2\alpha}} \tag{3.72}$$

$$S = \sqrt{2\alpha}\int \mathrm{d}q\sqrt{1 - \frac{\omega^2 q^2}{2\alpha}} - \alpha t \tag{3.73}$$

*But*

$$\beta = \frac{\partial S}{\partial \alpha} = (2\alpha)^{-1/2}\int \frac{\mathrm{d}q}{\sqrt{1 - \frac{\omega^2 q^2}{2\alpha}}} - t, \tag{3.74}$$

$$q = \sqrt{\frac{2\alpha}{\omega^2}}\sin\omega\,(t+\beta) \tag{3.75}$$

$$p = \frac{\partial S}{\partial q} = \frac{\partial W}{\partial q} = \sqrt{2\alpha}\cos\omega\,(t+\beta) = \dot{q} \tag{3.76}$$

*The initial conditions at $t = 0$ are given by $(q_0, p_0)$. If we square the equations for $q$ and $p$ we get*

$$2\alpha = p^2 + \omega^2 q^2 = p_0^2, \omega^2 q_0^2,$$

*and the other usual trick is*

$$\frac{\omega q_0}{p_0} = \tan\omega\beta$$

*And we naturally identify $\beta$ with the phase angle of the oscillator. Doing a bit more algebra:*

$$S = \sqrt{2\alpha}\int \mathrm{d}q\sqrt{1 - \frac{\omega^2 q^2}{2\alpha}} - \alpha t \tag{3.77}$$

$$= 2\alpha\int \cos^2\omega\,(t+\beta)\;\mathrm{d}t - \alpha t \tag{3.78}$$

$$= 2\alpha\int \cos^2\omega\,(t+\beta) - \frac{1}{2}\,\mathrm{d}t \tag{3.79}$$

*And the Lagrangian of the problem is*

$$\mathsf{L} = \frac{1}{2}\left(p^2 - \omega^2 q^2\right) \tag{3.80}$$

$$= 2\alpha\left(\cos^2\omega\,(t+\beta) - \frac{1}{2}\right) \tag{3.81}$$

*I.e, in this case we have*

$$S = \int \mathsf{L}\,\mathrm{d}t$$

### 3.5.2 The principle Function and the Lagrangian

The above example furnished us with an interesting relation between the principle function $S$ and the Lagrangian $\mathsf{L}$. Does this hold in general?
We have $p_i = \frac{\partial S}{\partial q_i}$, and also $Q_i = \beta_i = \frac{\partial S}{\partial \alpha_i}(q, \alpha, t)$. This last equation can be inverted to write $q$ as

$$q_j = q_j\left(\alpha, \beta, t\right).$$

Then after differentiation in the first equation we can substitute $q_j$ to obtain

$$p_i = p_i(\alpha, \beta, t)$$

And so

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \sum \frac{\partial S}{\partial q_i}\dot{q}_i + \frac{\partial S}{\partial t} = \sum p_i \dot{q}_i - \mathsf{H} = \mathsf{L}.$$

We can write this general result as

$$S = \int \mathsf{L}\,\mathrm{d}t + const.$$

# Chapter 4

# Integrable Systems

**Synopsis:** This chapter applies some of the results of the Hamiltonian formalism, as developed in Chapter 3, as well as investigating the important area of Adiabatic Invariants: Completely Integrable systems.

## 4.1 Integrable Systems

What is an integrable system?

## 4.2 Action-Angle Variables

Let us consider a system with one degree of freedom and assume it is conservative, so

$$\mathsf{H}(q, P) = \alpha_1.$$

Suppose we are interested in emphperiodic systems: There are two things we could mean by this:

1. The path of the system describes a closed loop in phase space. For example vibration. This will happen when both $p$ and $q$ are periodic function of time, and have the same frequency.

2. The path of the system is periodic in phase space, for example a pendulum going all the way round it's fixed point. Then we will have tha t$p(q) = p(q + q_0)$

**Definition:** We define the *angle variable* $J$, by

$$J = \oint p \, \mathrm{d}q,$$

where the contour integral is taken around one peroid of the system, whether it is of type (1) or (2).

## 4.3 Adiabatic Invariants