2007 S7: Classical Mechanics

Dr J Paton

June 3, 2008

Lecture 1

1 Introduction

The course is mainly about some of the development of classical mechanics in the 19th century which

- are helpful in solving problems (most recently of satellite and rocket motion)
- give entirely new insight into the subject so have an elegance which "is their own reward"
- allow unified treatment of non-mechanical branches of physics by analogy
- are a springboard for understanding the modifications to classical theory which are required for quantization.

Before getting into these developments, which go under the name of Lagrangian & Hamiltonian formulations, I spend some time on mechanics of a system of particles in the Newtonian formalism. This will review and certainly extend somewhat material covered in the first-year course.

2 Particle Mechanics à la Newton

There is a brief but readable introduction to the foundations of mechanics in Chapter 1 of Kibble. Position and time (given for a particle by $3 + 1$ numbers relative to some frame of reference) are chosen to be basic concepts and Newton's Laws are regarded as containing definitions in addition to physical laws. N1 contains the definition of an **inertial frame** as well as the physical assertion that such frames exist. N2 & N3 contain the definitions of mass and force. These laws, supplemented by the laws of force (e.g. law of gravitation), provide for N particles a set of $3N$ second-order ordinary differential equations, which

in principle may be used to determine the subsequent motion $\mathbf{r}_i(t)$, $i = 1 \ldots N$, given initial values of positions $\mathbf{r}_i(0)$ and velocities $\dot{\mathbf{r}}_i(0)$.

To be more specific, the N particles are each characterised by a mass m_i and at a given time t a position $\mathbf{r}_i(t) = (x_i(t), y_i(t), z_i(t))$ where (x, y, z) are Cartesian co-ordinates as measured in an inertial frame.

N2 says that $m_i\ddot{\mathbf{r}}_i = \mathbf{F}_i$ (total force on ith particle). $\mathbf{F}_i$ is usually taken to be a superposition $\mathbf{F}_i = \mathbf{F}_i(\text{int.}) + \mathbf{F}_i(\text{ext.})$, where $\mathbf{F}_i(\text{ext.})$ is the force due to everything not included in the N-particle system, and $\mathbf{F}_i(\text{int.})$ is given by

$$\mathbf{F}_i(\text{int.}) = \sum_{j \neq i} \mathbf{F}_{ij} \text{ (two-body force assumption)}$$

where $\mathbf{F}_{ij}$ is the force on the ith particle due to the jth. N3 then says

$$\mathbf{F}_{ij} = -\mathbf{F}_{ji} \tag{1}$$

The $\mathbf{F}$s are assumed to be known functions of $\mathbf{r}$s and t. I now review basic results on momentum, angular momentum, and energy of our system of N particles.

2.1 (Linear) Momentum and Centre of Mass

The total (linear) momentum of our system is defined to be $\mathbf{P} = \sum_i m_i\dot{\mathbf{r}}_i$

Then $\frac{d\mathbf{P}}{dt} = \sum_i m_i\ddot{\mathbf{r}}_i = \sum_i \mathbf{F}_i(\text{ext.})$ since the internal forces cancel in pairs by 1. Define the centre of mass $\mathbf{R}$ of the system by

$$\mathbf{R} = \frac{\sum m_i\mathbf{r_i}}{\sum m_i}$$

and writing $M = \sum m_i$, we have

$$\mathbf{P} = \sum m_i\dot{\mathbf{r}}_i = M\dot{\mathbf{R}} \quad \& \quad \dot{\mathbf{P}} = M\ddot{\mathbf{R}} = \sum_i \mathbf{F}_i(\text{ext.})$$

In words, the centre of mass of the system moves like a particle subject to a force equal to the sum of the external forces acting on the system. In particular, if the net external force on the system is $\mathbf{0}$, i.e. $\sum_i \mathbf{F}_i(\text{ext.}) = \mathbf{0}$,

$$\mathbf{P} = \text{constant vector, independent of time}$$

(conservation of total linear momentum) or, equivalently,

$$\frac{d\mathbf{R}}{dt} = \text{constant vector}$$

2.2 Angular Momentum

The angular momentum of a particle about the origin is more properly in Newtonian mechanics called "moment of momentum"

$$\mathbf{r}_i \wedge \mathbf{p}_i = \mathbf{r}_i \wedge m_i\dot{\mathbf{r}}_i = m_i\mathbf{r}_i \wedge \dot{\mathbf{r}}_i$$

The total angular momentum of the system is defined as the sum of these over all particles:

$$\mathbf{L} = \sum_i m_i \mathbf{r}_i \wedge \dot{\mathbf{r}}_i$$

So

$$\begin{aligned} \frac{d\mathbf{L}}{dt} &= \sum_i m_i \left(\underbrace{\dot{\mathbf{r}}_i \wedge \dot{\mathbf{r}}_i}_{=0} + \mathbf{r}_i \wedge \ddot{\mathbf{r}}_i \right) \\ &= \sum_i \mathbf{r}_i \wedge \mathbf{F}_i \,(\text{sum of moments of forces}) \end{aligned}$$

Now

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

if $\mathbf{F}_{ij}$ is along the line joining the particles i and j, an extra assumption at the moment.

Under this plausible assumption, the rate of change of total angular momentum about the origin

$$\begin{aligned} \frac{d\mathbf{L}}{dt} &= \sum_i \mathbf{r}_i \wedge \mathbf{F}_i \,(\text{ext.}) \\ &= \text{sum of moments of external forces about the origin} \\ &= \text{total external couple about the origin acting on the system} \end{aligned}$$

i.e. the sum of all external couples[1] about the origin = rate of change of the total angular momentum about the origin.

We can equally choose a point $\mathbf{r}_0 \neq \mathbf{0}$ different from the origin with

$$\begin{aligned} \mathbf{L}' &= (\mathbf{r} - \mathbf{r}_0) \wedge m\,(\dot{\mathbf{r}} - \dot{\mathbf{r}}_0) \\ &= (\mathbf{r} - \mathbf{r}_0) \wedge m\dot{\mathbf{r}} \end{aligned}$$

if $\mathbf{r}_0$ is fixed in the inertial frame, and we would find

$$\frac{d\mathbf{L}'}{dt} = \sum (\mathbf{r}_i - \mathbf{r}_0) \wedge \mathbf{F}_i \,(\text{ext.})$$

i.e. rate of change of angular momentum about $\mathbf{r}_0$ = total external couple about $\mathbf{r}_0$.

[One of the problems asks you to show the same result when $\mathbf{r}_0$ is replaced by the centre of mass.]

In particular, if the net external couple is $\mathbf{0}$, then $\frac{d\mathbf{L}}{dt} = 0$, i.e. angular momentum = constant. (Conservation of total angular momentum.)

No further ado, since we shall get a far clearer understanding of this and the corresponding conservation law of linear momentum in the later lectures.

[1] "couple" = "moment of force"

2.3 Energy

Suppose the forces on the particles of the system are "derivable from a potential". This means that there is a scalar potential energy function $V(\mathbf{r}_1 \ldots \mathbf{r}_N, t)$ with $\mathbf{F}_i = -\nabla_i V(\mathbf{r}_1 \ldots \mathbf{r}_N, t)$. (The t-dependence might occur because the external forces $\mathbf{F}_i$ (ext.) are due to moving bodies outside the system.)

Then, defining the total **kinetic energy** of the system by $T = \frac{1}{2}\sum_i m_i \dot{\mathbf{r}}_i^2$ and the total energy by $E = T + V$ we find

$$\begin{aligned} \frac{dE}{dt} &= \sum_i m_i \dot{\mathbf{r}}_i . \ddot{\mathbf{r}}_i + \frac{\partial V}{\partial t} + \dot{\mathbf{r}}_i . \nabla_i V \\ &= \frac{\partial V}{\partial t} \quad \text{since} \quad m_i \ddot{\mathbf{r}}_i = -\nabla_i V \end{aligned}$$

So, in particular if V is independent of time, $E = \text{const.}$ - the conservation law of total energy.

[Parenthetically, we can write $\mathbf{r}_i = \mathbf{R} + \mathbf{r}_i'$ where $\mathbf{R}$ is the position of the centre of mass. Then $T = \frac{1}{2}M\dot{R}^2 + \sum_i \frac{1}{2} m_i \dot{\mathbf{r}}_i'^2$, since $\sum_i \mathbf{r}_i' = \mathbf{0}$.]

Also,

$$\mathbf{L} = M\mathbf{R} \wedge \dot{\mathbf{R}} + \sum_i (\mathbf{r}_i - \mathbf{R}) \wedge m_i \left(\dot{\mathbf{r}}_i - \dot{\mathbf{R}}_i \right)$$

i.e. both kinetic energy and angular momentum can be considered as the sum of a piece associated with motion of the centre of mass plus a piece associated with internal motion about the centre of mass.

Lecture 2

3 Rigid Bodies

A **rigid body** is an idealization of a system of particles in which the internal forces are such that the separation distances between all pairs of particles are fixed. This is often a good approximation to a real system provided the external forces are not too violent. The **general displacement** of a rigid body is a combination of a displacement and a rotation, i.e. for all particles i of the rigid body,

$$\mathbf{r}_i' = \mathcal{R}(\mathbf{r}_i + \delta) = \mathcal{R}(\mathbf{r}_i) + \delta'$$

where $\mathcal{R}$ is an operator which rotates points about the origin.

We shall now assume that one point of the rigid body is fixed or that positions are measured relative to the centre of mass. Either way, and dropping the subscript i, we have $\mathbf{r}' = \mathcal{R}\mathbf{r}$. Consider now the velocities of the points of the body. These are given by

$$\dot{\mathbf{r}} = \lim_{\delta t \to 0} \frac{\mathbf{r}(t + \delta t) - \mathbf{r}(t)}{\delta t} = \lim_{\delta t \to 0} \frac{\mathcal{R}(t + \delta t)\,\mathbf{r}(0) - \mathcal{R}(t)\,\mathbf{r}(0)}{\delta t}$$

where $\mathcal{R}(t)$ means the rotation which takes points initially at $\mathbf{r}(0)$ to their positions $\mathbf{r}(t)$ at time t. This may be rewritten

$$\dot{\mathbf{r}} = \lim_{\delta t \to 0} \frac{\mathcal{R}(t+\delta t) - \mathcal{R}(t)}{\delta t} \mathbf{r}(0)$$

Now use the inverse rotation $\mathbf{r}(0) = \mathcal{R}^{-1}(t)\,\mathbf{r}(t)$ to give

$$\dot{\mathbf{r}} = \lim_{\delta t \to 0} \frac{\mathcal{R}(t+\delta t)\,\mathcal{R}^{-1}(t) - \mathbb{1}}{\delta t} \mathbf{r}(0)$$

involving the combined rotation $\mathcal{R}(t+\delta t)\,\mathcal{R}^{-1}(t)$. $\mathbb{1}$ is the identity rotation, with $\mathbb{1}\mathbf{r}(0) = \mathbf{r}(0)$.

Now any rotation is characterised by an axis, i.e. a unit vector $\hat{\mathbf{n}}$ specifying the direction of the axis and an angle. So $\mathcal{R}(t+\delta t)\,\mathcal{R}^{-1}(t)$ may be written $\mathcal{R}(\hat{\mathbf{n}},\, \delta\theta)$. In terms of vectors represented by column matrices of Cartesian components

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \underline{\underline{R}}(\hat{\mathbf{n}},\, \delta\theta) \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

where the 3×3 matrix $\underline{\underline{R}}(\hat{\mathbf{n}},\, \delta\theta) = \underline{\underline{R}}$:

1. satisfies $\underline{\underline{R}} \begin{pmatrix} n_x \\ n_y \\ n_z \end{pmatrix} = \begin{pmatrix} n_x \\ n_y \\ n_z \end{pmatrix}$
2. is an orthogonal matrix, and
3. has determinant $+1$.

It must also be close to the unit matrix. Out of a hat, I claim that the matrix

$$\begin{pmatrix} 1 & -\delta\theta n_z & +\delta\theta n_y \\ +\delta\theta n_z & 1 & -\delta\theta n_x \\ -\delta\theta n_y & +\delta\theta n_x & 1 \end{pmatrix}$$

has all three properties (1), (2), (3) to first order in $\delta\theta$. So, finally,

$$\begin{aligned} \begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{pmatrix} &= \lim_{\delta t \to 0} \frac{\delta\theta}{\delta t} \begin{pmatrix} 1 & -\delta\theta n_z & +\delta\theta n_y \\ +\delta\theta n_z & 1 & -\delta\theta n_x \\ -\delta\theta n_y & +\delta\theta n_x & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \\ &= \frac{d\theta}{dt} \begin{pmatrix} -yn_z + zn_y \\ -zn_x + xn_z \\ -xn_y + yn_x \end{pmatrix} \end{aligned}$$

i.e. $\dot{\mathbf{r}} = \boldsymbol{\omega} \wedge \mathbf{r}$ where $\boldsymbol{\omega} = \dot{\theta}\mathbf{n}$ is the **instantaneous angular velocity (pseudo-)vector**, in general dependent on t.

The angular momentum of a rigid body (about fixed origin or centre of mass) is (summation over particles of body, suppressing particle labels)

$$\mathbf{L} = \sum m\mathbf{r} \wedge (\boldsymbol{\omega} \wedge \mathbf{r}) = \sum m \left(\mathbf{r}^2\boldsymbol{\omega} - \mathbf{r}\left(\mathbf{r}.\boldsymbol{\omega}\right)\right)$$

So $\mathbf{L}$ is a linear function of $\boldsymbol{\omega}$, $\mathbf{L} = \underline{\underline{I}}\boldsymbol{\omega}$ where $\underline{\underline{I}}$ has a co-ordinate-free definition. In terms of a given set of Cartesian co-ordinates it may be represented my the 3×3 symmetric matrix

$$\begin{pmatrix} \sum m\left(y^2+z^2\right) & -\sum mxy & -\sum mxz \\ -\sum mxy & \sum m\left(z^2+x^2\right) & -\sum myz \\ -\sum mxz & -\sum myz & \sum m\left(x^2+y^2\right) \end{pmatrix}$$

It is called the **moment of inertia tensor**. It may also be used to write the kinetic energy of a rigid body:

$$\begin{aligned} T = \frac{1}{2}\sum m\dot{\mathbf{r}}^2 &= \frac{1}{2}\sum m\left(\boldsymbol{\omega} \wedge \mathbf{r}\right)^2 \\ &= \frac{1}{2}\sum m\left(\mathbf{r}^2\boldsymbol{\omega}^2 - \left(\boldsymbol{\omega}.\mathbf{r}\right)\left(\mathbf{r}.\boldsymbol{\omega}\right)\right) \\ &= \frac{1}{2}\boldsymbol{\omega}.\underline{\underline{I}}\boldsymbol{\omega} \end{aligned}$$

where, using matrices, the left-hand $\boldsymbol{\omega}$ is a row and the right a column.

A (Cartesian) **tensor** may be defined as a generalisation of a vector, an object which has representation in any given Cartesian co-ordinate frame by a set of components $\begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix}$. If one changes co-ordinates by $\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \underline{\underline{R}} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, where $\underline{\underline{R}}$ is a rotation matrix, then the components of a vector **v**alter in the same way: $\begin{pmatrix} v'_x \\ v'_y \\ v'_z \end{pmatrix} = \underline{\underline{R}} \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix}$.

To show how the components of $\underline{\underline{I}}$ change, consider the relation $\mathbf{L} = \underline{\underline{I}}\boldsymbol{\omega}$. The two vectors $\mathbf{L}$, $\boldsymbol{\omega}$ change to $\mathbf{L}' = \underline{\underline{R}}\mathbf{L} = \underline{\underline{R}}\underline{\underline{I}}\boldsymbol{\omega} = \underline{\underline{R}}\underline{\underline{I}}\underline{\underline{R}}^{-1}\left(\underline{\underline{R}}\boldsymbol{\omega}\right) = \underline{\underline{I}}'\boldsymbol{\omega}'$, where $\underline{\underline{I}}' = \underline{\underline{R}}\underline{\underline{I}}\underline{\underline{R}}^{-1}$. This last is the relationship between the components of a **second-rank tensor** in two co-ordinate systems.

Now the components of $\underline{\underline{I}}$ in a given Cartesian co-ordinate system form a symmetric (3×3) matrix. Such a matrix has three independent eigenvectors which may be chosen orthogonal. Take them to be $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ with $\mathbf{v}_1 = \mathbf{v}_2 \wedge \mathbf{v}_3$ and the corresponding (real) eigenvalues to be I_1, I_2, I_3. Choosing co-ordinates with respect to axes along $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ gives

$$\underline{\underline{I}}' = \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix}$$

(Question: What is $\underline{\underline{R}}$ in terms of $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$?)

I_1, I_2, I_3 are called the **principal moments of inertia** of the body and $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ the **principal axes**. In genera, of course, when the body is moving, $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ will vary with time. If at a given time $\boldsymbol{\omega} = \omega_1\mathbf{v}_1 + \omega_2\mathbf{v}_2 + \omega_3\mathbf{v}_3$ then the kinetic energy is $T = \frac{1}{2}\left(I_1\omega_1^2 + I_2\omega_2^2 + I_3\omega_3^2\right)$ and the angular momentum is $\mathbf{L} = I_1\omega_1\mathbf{v}_1 + I_2\omega_2\mathbf{v}_2 + I_3\omega_3\mathbf{v}_3$. Note that in general $\mathbf{L}$ is **not** parallel to $\boldsymbol{\omega}$. $\mathbf{L}$ is parallel to $\boldsymbol{\omega}$ only if $\boldsymbol{\omega}$ is along an eigenvector of $\underline{\underline{I}}$.

The motion of a rigid body is entirely determined (at least, in principle) by the three equations:

$$M\ddot{\mathbf{R}} = \mathbf{F}_{\text{ext.}} \tag{2}$$

(for the motion of the centre of mass),

$$\frac{d\mathbf{L}}{dt} = \mathbf{C}_{\text{ext.}} \tag{3}$$

(for the motion about the centre of mass **or** a fixed point), where $\mathbf{C}_{\text{ext.}} = \sum_i \mathbf{r}_i \wedge \mathbf{F}_{i\,\text{ext.}}$ =Sum of moments of external forces about the centre of mass or fixed point (the external couple acting on the body), and the relationship

$$\mathbf{L} = \underline{\underline{I}}\mathbf{R}\omega \tag{4}$$

The solution to these equations is complicated by:

1. Using components with respect to fixed axes, $\underline{\underline{I}}$ depends in an unknown way on time, **or**
2. Using components with respect to the principal axes, $\mathbf{C}_{\text{ext.}}$ depends in an unknown way on time.

The general motion of a rigid body is a difficult problem and it pays to have the sophisticated tools of Lagrangian and/or Hamiltonian mechanics to treat it. Even then, it is not trivial (but is exactly solvable) in the general case $I_1 \neq I_2 \neq I_3$ **when** there is no external couple. The case of a symmetric body (two of I_1, I_2, I_3 equal) is treatable using Lagrangian or Hamiltonian methods, but not in the syllabus because of time constraints.

Example: Free motion of a rigid body

$$\frac{d\mathbf{L}}{dt} = \mathbf{0} \qquad \mathbf{L} = \underline{\underline{I}}\boldsymbol{\omega}$$

Using components along the principal axes, $\mathbf{v}_i$, $i = 1, 2, 3$,

$$\mathbf{L} = L_1\mathbf{v}_1 + L_2\mathbf{v}_2 + L_3\mathbf{v}_3 \tag{5}$$

$$\boldsymbol{\omega} = \omega_1\mathbf{v}_1 + \omega_2\mathbf{v}_2 + \omega_3\mathbf{v}_3 \tag{6}$$

$$\begin{aligned} \frac{d\mathbf{L}}{dt} &= \frac{dL_1}{dt}\mathbf{v}_1 + \frac{dL_2}{dt}\mathbf{v}_2 + \frac{dL_3}{dt}\mathbf{v}_3 \\ &\quad + L_1\boldsymbol{\omega} \wedge \mathbf{v}_1 + L_2\boldsymbol{\omega} \wedge \mathbf{v}_2 + L_3\boldsymbol{\omega} \wedge \mathbf{v}_3 \end{aligned} \tag{7}$$

(The second line is because $\frac{d\mathbf{v}_1}{dt} = \boldsymbol{\omega} \wedge \mathbf{v}_i$.)

From 6,

$$\begin{aligned} \boldsymbol{\omega} \wedge \mathbf{v}_1 &= (\omega_1 \mathbf{v}_1 + \omega_2 \mathbf{v}_2 + \omega_3 \mathbf{v}_3) \wedge \mathbf{v}_1 \\ &= -\omega_2 \mathbf{v}_3 + \omega_3 \mathbf{v}_2 \end{aligned}$$

since $\mathbf{v}_i$ are orthogonal unit vectors.

Similarly,

$$\begin{aligned} \boldsymbol{\omega} \wedge \mathbf{v}_2 &= -\omega_3 \mathbf{v}_1 + \omega_1 \mathbf{v}_3 \\ \boldsymbol{\omega} \wedge \mathbf{v}_3 &= -\omega_1 \mathbf{v}_2 + \omega_2 \mathbf{v}_1 \end{aligned}$$

So second line of 7 is given by

$$\begin{aligned} \begin{bmatrix} L_1 (-\omega_2 \mathbf{v}_3 + \omega_3 \mathbf{v}_2) \\ +L_2 (-\omega_3 \mathbf{v}_1 + \omega_1 \mathbf{v}_3) \\ +L_3 (-\omega_1 \mathbf{v}_2 + \omega_2 \mathbf{v}_1) \end{bmatrix} &= \begin{bmatrix} I_1 \omega_1 (-\omega_2 \mathbf{v}_3 + \omega_3 \mathbf{v}_2) \\ +I_2 \omega_2 (-\omega_3 \mathbf{v}_1 + \omega_1 \mathbf{v}_3) \\ +I_3 \omega_3 (-\omega_1 \mathbf{v}_2 + \omega_2 \mathbf{v}_1) \end{bmatrix} \\ \text{(using } L_i = I_i \omega_i, \, i = 1, \, 2, \, 3) &= \begin{bmatrix} (-I_2 \omega_2 \omega_3 + I_3 \omega_3 \omega_2) \, \mathbf{v}_1 \\ +(I_1 \omega_1 \omega_3 - I_3 \omega_1 \omega_3) \, \mathbf{v}_2 \\ +(-\omega_2 \omega_1 I_1 + \omega_1 \omega_2 I_2) \, \mathbf{v}_3 \end{bmatrix} \end{aligned}$$

So the full RHS of 7 is

$$\left\{ I_1 \frac{d\omega_1}{dt} + (I_3 - I_2) \, \omega_2 \omega_3 \right\} \mathbf{v}_1 + \left\{ I_2 \frac{d\omega_2}{dt} + (I_1 - I_3) \, \omega_3 \omega_1 \right\} \mathbf{v}_2 + \left\{ I_3 \frac{d\omega_3}{dt} + (I_2 - I_1) \, \omega_1 \omega_2 \right\}$$

But

$$\text{LHS} = \frac{d\mathbf{L}}{dt} = \mathbf{0}$$

if there is **no** external couple, giving:

$$\boxed{\begin{array}{l} I_1 \frac{d\omega_1}{dt} + \omega_2 \omega_3 (-I_2 + I_3) = 0 \\ I_2 \frac{d\omega_2}{dt} + \omega_3 \omega_1 (-I_3 + I_1) = 0 \\ I_3 \frac{d\omega_3}{dt} + \omega_1 \omega_2 (-I_1 + I_2) = 0 \end{array}}$$

These are called **Euler's (Rigid Body) Equations**.

Solutions:

1. $\boldsymbol{\omega}$ along a principal axis $\mathbf{v}_i$, $\boldsymbol{\omega} = \omega_i \mathbf{v}_i$ - Easy! $\mathbf{L} = \text{const.} = I_i \boldsymbol{\omega}$
2. $\boldsymbol{\omega}$ close to being along a principal axis.
   Let $\omega_1 \gg \omega_2, \, \omega_3$. Then $\frac{d\omega_1}{dt} \approx 0$ to first order, so $\omega_1 \approx \text{const.}$

$$\begin{aligned} \frac{d\omega_2}{dt} &= +\omega_1 \frac{I_3 - I_1}{I_2} \omega_3 \\ \frac{d\omega_3}{dt} &= -\omega_1 \frac{I_2 - I_1}{I_3} \omega_2 \end{aligned}$$

from which

$$\frac{d^2}{dt^2}\begin{pmatrix}\omega_2\\ \omega_3\end{pmatrix} = -\omega_1^2\frac{(I_3-I_1)(I_2-I_1)}{I_2I_3}\begin{pmatrix}\omega_2\\ \omega_3\end{pmatrix}$$

This will give stable oscillations of ω_2, ω_3 with angular frequency $\omega' = \omega_1\sqrt{\frac{(I_3-I_1)(I_2-I_1)}{I_2I_3}}$ provided I_1 is either the largest or smallest of I_1, I_2, I_3 and instability otherwise, a result "easily" demonstrated with a matchbox. It is not difficult to show that $\begin{pmatrix}\omega_1\\ \omega_2\\ \omega_3\end{pmatrix}$ executes an elliptical motion (circular if $I_2 = I_3$) about $\begin{pmatrix}\omega_1\\ 0\\ 0\end{pmatrix}$. This is a description of the motion in the **body frame**, i.e. with respect to axes along the principal axes of the body. If the body is the Earth, this is a description appropriate to an observer on the Earth with ω_1 corresponding to the period of 1 day, and ω_1 along the principal axis close to the north-south axis. See, for example, Kibble or Goldstein for the application of this result to the Earth, which is not a perfectly rigid body. Nevertheless treating it as such with $I_2 = I_3$ and slightly less than I_1(oblate spheroid) gives $\omega' = \frac{\Delta I}{I}\omega_1$ with ω_1 corresponding to the period of 1 day. $\frac{\Delta I}{I}$ is 0.00327, giving ω'corresponding to a period of $\sim$306 days.

[**Addendum**: An **irregular** wobble called the Chandler wobble of the Earth's axis of rotation is known, having fist been observed by Mr Chandler in the 1880s. It is quite small, giving a variation in the positions of the poles of order a few metres, but Fourier analysis yields a peak at a somewhat longer period of $\sim$400 days and a lot of structure. The lengthening of the period is attributed to the lack of rigidity of the Earth. The question of the cause of the structure is a subject of present geophysics research. The latest proposal, based on detailed mathematical modelling, is that it is caused mainly in variations in pressure at the bottom of oceans.]

[**Note** that this effect has **nothing** to do with something called the precession of the equinoxes which has a period of about 26,000 years - explained in terms of **moments of forces acting on the Earth** due to the Sun and Moon.]

Lecture 3

4 Generalised Co-ordinates and Lagrangian Mechanics

Sometimes the $3N$ Cartesian co-ordinates required to apply Newtonian procedures are inconvenient.

4.1 Example: Simple Pendulum

For example, for a simple pendulum moving in a plane, there is only one independent variable, θ. To do it by Newton, use Cartesian co-ordinates related to θ by

$$\begin{aligned} x &= \ell \sin\theta \quad \text{(horizontal)} \\ y &= \ell \cos\theta \quad \text{(vertical down from support)} \end{aligned}$$

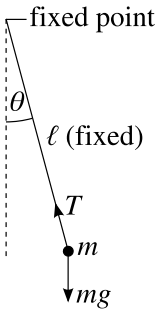


Figure 1: The simple pendulum

$$\begin{aligned} \therefore \dot{x} &= \ell \cos\theta \dot{\theta} \\ \dot{y} &= -\ell \sin\theta \dot{\theta} \end{aligned}$$

$$\begin{aligned} \therefore \ddot{x} &= -\ell \sin\theta \theta^2 + \dot{\ell} \cos\theta \ddot{\theta} \\ \ddot{y} &= -\ell \cos\theta \dot{\theta}^2 - \ell \sin\theta \ddot{\theta} \end{aligned}$$

From Newton's second law:

$$m\ell \left(-\sin\theta \dot{\theta}^2 + \cos\theta \ddot{\theta} \right) = -T \sin\theta \tag{8}$$

$$m\ell \left(-\cos\theta \dot{\theta}^2 - \sin\theta \ddot{\theta} \right) = mg - T \cos\theta \tag{9}$$

$\Big[(8) \times \cos\theta - (9) \times \sin\theta \Big] \div m\ell$ gives

$$\boxed{\ddot{\theta} = -\frac{g}{\ell} \sin\theta}$$

The Lagrangian method allows one to go very much more quickly to the boxed equation. For more complicated examples it is a major simplification. First some terminology:

For any system, a set of **generalised co-ordinates** $\{q_i\}$ is a set of variables that together completely specify the instantaneous positions of all parts of the system. They may be **constrained** as $\{x, y\}$ of the simple pendulum example which satisfy the constraint $x^2 + y^2 = \ell^2$, or the $3N$ co-ordinates of the particle forming a rigid body for which all $|\mathbf{r}_i - \mathbf{r}_j|$ are fixed.

The time derivatives $\dot{q}_i$ are called **generalised velocities**.

For a system, for which there are n **un**constrained generalised co-ordinates, with a potential energy V, the Lagrangian method consists of the following steps:

1. express V in terms of the $\{q_i\}$ (and possibly t)
2. express the kinetic energy T in terms of $\{\dot{q}_i\}$ and $\{q_i\}$

3. form $L = T - V$ (**not L** = angular momentum), which is again a function of the $\{q_i\}$, $\{\dot{q}_i\}$, and t. $L(\{q_i\}, \{\dot{q}_i\})$ is the **Lagrangian**

4. then the **Lagrange equations** are

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = \frac{\partial L}{\partial q_i} \qquad i = 1, 2, \ldots, n$$

For the simple pendulum example, taking θ as the generalised co-ordinate (unconstrained)

1. $V = -mg\cos\theta$
2. $T = \frac{1}{2}m\ell^2\dot{\theta}^2$
3. $L = T - V = \frac{1}{2}m\ell^2\dot{\theta}^2 + mg\cos\theta$
4. $\frac{d}{dt}\left(m\ell^2\dot{\theta}\right) = -mg\sin\theta$
   or
$$\ddot{\theta} = -\frac{g}{\ell}\sin\theta$$

For more complicated problems, expressing T and V in terms of generalised co-ordinates and velocities (2 scalar functions) is a lot easier than expressing Cartesian components of acceleration in terms of independent variables and eliminating **forces of constraint** (tension T in the simple pendulum example). So this is a substantial advantage. A possible disadvantage is if one wishes to know the force of constraint. For that one needs to introduce **Lagrange multipliers**, which I hope to come back to later.

N.B. I have not yet shown the equivalence of the Lagrange method to that of Newton. See later.

4.2 Example: Spherical Pendulum

This is like the simple pendulum, but now no longer confined to a plane. Choose generalised co-ordinates to be (θ, ϕ), spherical polar angles.

$$\left.\begin{array}{rcl} T &=& \frac{1}{2}m\ell^2\left(\dot{\theta}^2 + \sin^2\theta\dot{\phi}^2\right) \\ V &=& -mg\ell\cos\theta \end{array}\right\} \Rightarrow L = \frac{1}{2}m\ell^2\left(\dot{\theta}^2 + \sin^2\theta\dot{\phi}^2\right) + mg\ell\cos\theta$$

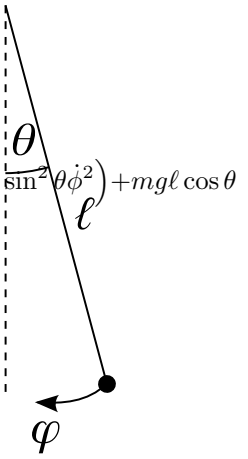


Figure 2: The spherical pendulum

$$\begin{aligned}
\frac{\partial L}{\partial \dot{\theta}} &= m\ell^2\dot{\theta} \\
\frac{\partial L}{\partial \theta} &= m\ell^2\left(\sin\theta\cos\theta\dot{\phi}^2\right) - mg\ell\sin\theta \\
\frac{\partial L}{\partial \dot{\phi}} &= m\ell^2\sin^2\theta\dot{\phi}^2 \\
\frac{\partial L}{\partial \phi} &= 0
\end{aligned}$$

So the two Lagrange equations are

$$\begin{aligned}
\ddot{\theta} &= \sin\theta\cos\theta\dot{\phi}^2 - \frac{g}{\ell}\sin\theta \qquad (10) \\
\frac{d}{dt}\left(\sin^2\theta\dot{\phi}\right) &= 0 \\
\text{or } \sin^2\theta\dot{\phi} &= \text{const. } k
\end{aligned}$$

The second Lagrange equation is worthy of a remark. Its simplicity is a consequence of L being **independent of the generalised co-ordinate ϕ**. As a consequence, $\frac{\partial L}{\partial \dot{\phi}}$, called the **generalised momentum conjugate to ϕ**, p_ϕ, is a **constant of the motion**.

$p_\phi = k$ may be used to eliminate $\dot{\phi}$ from the θ equation, making it an equation just in the one variable θ.

In this example there is a further constant of the motion, the total energy, which allows the θ equation to be reduced to first order. To see this, multiply (10) by $\dot{\theta}$ and eliminate $\dot{\phi} = \frac{k}{\sin^2\theta}$ to give

$$\dot{\theta}\ddot{\theta} = \left(\frac{\cos\theta}{\sin^3\theta}k^2 - \frac{g}{\ell}\sin\theta\right)\dot{\theta}$$

and integrate both sides with respect to t to give

$$\frac{1}{2}\dot{\theta}^2 + \frac{1}{2}\frac{k^2}{\sin^2\theta} - \frac{g}{\ell}\cos\theta = \text{const.}$$

which may be checked to be KE+PE÷factor $m\ell^2$. There is a standard procedure for obtaining a conserved quantity, usually but not invariably the total energy. We shall come back to it shortly.

4.3 Equivalence of Lagrangian and Newtonian Formulations

So far I have just plucked Lagrange's equations out of a hat; I now need to justify them. This is done in two steps:

1. Show that if the generalised co-ordinates are taken to be simply the $3N$ Cartesian co-ordinates $(x_i,\, y_i,\, z_i)$ $i = 1, \ldots,\, N$, then Lagrange's equations are just N2.
   This is straightforward. Taking

$$L = \sum_i \frac{1}{2} m_i \left(\dot{x}_i^2 + \dot{y}_i^2 + \dot{z}_i^2\right) - V\left(x_1,\, y_1,\, z_1, \ldots\right)$$

   then

$$\frac{\partial L}{\partial \dot{x}_i} = m_i \dot{x}_i \quad \text{and} \quad \frac{\partial L}{\partial x_i} = -\frac{\partial V}{\partial x_i}$$

   The Lagrange equation in x_i is just

$$m_i \ddot{x}_i = -\frac{\partial V}{\partial x_i}$$

   which is just the x_i component of N2.

2. Show that if the Lagrange equations are valid for one set of generalised co-ordinates then they are valid for any other.

Combining (1) and (2) would prove that Newton's Laws imply Lagrange's equations for any set of generalised co-ordinates. There are two ways of proving (2), which goes under the fancy title of **covariance of Lagrange's equations under arbitrary point (or contact) transformations**. This means the following:

Let $\{q_i\}$ and $\{q_i'\}$ be two sets of generalised co-ordinates. The Lagrangian $T - V$ may be expressed either in terms of $\{q_i\}$ and $\{\dot{q}_i\}$ or in terms of $\{q_i'\}$ and $\{\dot{q}_i'\}$.It is the same physical quantity but in general a different function. So we have

$$T - V = L\left(\{q_i\},\, \{\dot{q}_i\}\right) = L'\left(\{q_i'\},\, \{\dot{q}_i'\}\right)$$

Then (2) says

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = \frac{\partial L}{\partial q_i} \quad (\forall i) \Rightarrow \frac{d}{dt}\left(\frac{\partial L'}{\partial \dot{q}_i'}\right) = \frac{\partial L'}{\partial q_i'} \quad (\forall i)$$

A straightforward but tedious way of proving (2) is to use the chain rule for partial derivatives. Since q's and q''s are both a specification of the mechanical system's position at a given time, there must be a relation between the two set of generalised co-ordinates $q_i' = f_i\left(q_1 \ldots q_n,\, t\right)$. So for example

$$\frac{\partial L}{\partial q_i} = \sum_j \left(\frac{\partial L'}{\partial q_j'}\right)\left(\frac{\partial f_j}{\partial q_i}\right) + \ldots$$

where the second term comes from the fact that $\dot{q}_i'$

$$\dot{q}_i' = \frac{df_i}{dt} = \frac{\partial f_i}{dt} + \sum_j \frac{\partial f_i}{\partial q_j}\dot{q}_j$$

depends on both generalised velocities **and** co-ordinates. So

$$\cdots = \sum_j \frac{\partial L'}{\partial \dot{q}'_j} \frac{\partial \dot{q}'_j}{\partial q_i} = \sum_j \frac{\partial L'}{\partial \dot{q}'_j} \left(\frac{\partial}{\partial q_i} \frac{\partial f_j}{\partial t} + \sum_k \frac{\partial^2 f_j}{\partial q_i \partial q_k} \dot{q}_k \right)$$

We do not pursue this method here, as the other method based on Hamilton's principle (Principle of Stationary Action) is straightforward and gives considerably more insight.

Lecture 4

5 Principle of Stationary Action (Hamilton's Principle)

This concerns motion of a mechanical system in **configuration space**, an n-dimensional space the points $\mathbf{q}$ of which correspond to a complete specification of the positions of all elements of the system at a given time. Co-ordinates in configuration space are simply a complete set of generalised co-ordinates $(q_1, q_2, \ldots q_n)$. The Lagrangian of the system depends on position and velocity in configuration space and has a co-ordinate independent definition (for the systems we have considered so far, $L = T - V$) with a specific functional form $L(q_1 \ldots q_n, \dot{q}_1 \ldots \dot{q}_n, t)$ for any specific set of co-ordinates in configuration space. The principle of stationary action concerns the **action integral** $\int L\,dt$ **along paths in configuration space**.

Suppose we consider a specific path $\mathbf{q}(t)$ in configuration space starting at $\mathbf{q}_0$ at $t = t_0$ and ending at time $t = t_1$ at $\mathbf{q}_1$. Representing each $\mathbf{q}$ by only one dimension instead of the full n dimensions gives the highly simplified picture shown in figure 3. Different paths $\mathbf{q}(t)$ in configuration space will lead to different values of the action integral $\int_{t_0}^{t_1} L\,dt$. $\int_{t_0}^{t_1} L\,dt$ is called a path-dependent integral or **path integral** since its value depends on the path $\mathbf{q}(t)$ along which it is evaluated. It is also called a **functional integral** since, for a given set of generalised co-ordinates $q_1(t) \ldots q_n(t)$ its value

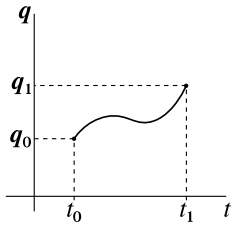


Figure 3: A simplified configuration space

$$\int dt\, L(q_1(t), q_2(t), \ldots, q_n(t), \dot{q}_1(t), \dot{q}_2(t), \ldots, \dot{q}_n(t), t)$$

depends on not just a set of variables as an ordinary (single or multiple) integral, but on a set of functions. The principle of stationary action (Hamilton's principle) requires us to compare its value for the **actual path** of the

system in configuration space with its value for neighbouring paths. All paths must begin at the same point $\mathbf{q}_0 = \mathbf{q}(t_0)$ and end at the same point $\mathbf{q}_1 = \mathbf{q}(t_1)$, but may deviate in between by a "small amount". Using a set of generalised co-ordinates $(q_1(t), q_2(t), \ldots, q_n(t))$ specifies the actual path and $(q_1(t) + \delta q_1(t), q_2(t) + \delta q_2(t), \ldots, q_n(t) + \delta q_n(t))$ the neighbouring path, with $\delta q_i(t_0) = 0 = \delta q_i(t_1)$, $i = 1, 2, \ldots, n$. We have, therefore, for the deviated path

$$\int_{t_0}^{t_1} L\,dt = \int_{t_0}^{t_1} dt\, L\left[\begin{array}{c} q_1(t) + \delta q_1(t), q_2(t) + \delta q_2(t), \ldots, q_n(t) + \delta q_n(t), \\ \dot{q}_1(t) + \delta\dot{q}_1(t), \dot{q}_2(t) + \delta\dot{q}_2(t), \ldots, \dot{q}_n(t) + \delta\dot{q}_n(t), t \end{array}\right]$$

So, (action integral for deviated path) − (action integral for the actual path) to first order in the deviation $\delta q_1(t), \delta q_2(t), \ldots, \delta q_n(t)$ is, using the chain rule,

$$\begin{aligned} &\int_{t_0}^{t_1} dt \sum_i \left(\frac{\partial L}{\partial q_i}\delta q_i + \frac{\partial L}{\partial \dot{q}_i}\delta\dot{q}_i\right) \\ &\quad = \int_{t_0}^{t_1} dt \sum_i \left\{\frac{\partial L}{\partial q_i}\delta q_i + \frac{d}{dt}\left[\left(\frac{\partial L}{\partial \dot{q}_i}\right)\delta q_i\right] - \delta q_i \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right)\right\} \end{aligned}$$

Now the terms $\frac{d}{dt}[\ldots]$ may be integrated $\int dt$ to give $\left[\frac{\partial L}{\partial \dot{q}_i}\delta q_i\right]_{t_0}^{t_1}$, which vanishes since $\delta q_i(t_0) = 0 = \delta q_i(t_1)$. So the change in the action integral is, to first order

$$\int dt\, \delta q_i \left[\frac{\partial L}{\partial q_i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right)\right]$$

This vanishes for arbitrary $\delta q_i(t)$ (i.e. no change in the action integral to first order, i.e. action stationary) if and only if

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

for all t and $i = 1 \ldots n$. These are Lagrange's equations.

The conclusion therefore is that the principle of stationary action, a co-ordinate independent statement, is entirely equivalent to Lagrange's equations in terms of any chosen set of generalised co-ordinates. I can deduce therefore that Lagrange's equations, if they are true for one set of generalised co-ordinates, are true for any other set. Since we know (1) (from 3) that using Cartesian co-ordinates of particles, the equations of Newton's Second Law are the same as Lagrange's equations, we have now demonstrated the validity of Lagrange's equations for **any** set of generalised co-ordinates.

[The derivation of the equivalence of the action principle to Lagrange's equations is a particular example of what is called the "calculus of variations". The value of a functional integral being stationary to first order is equivalent to a set of differential equations (whether maximum or minimum depends on second-order terms). The books (and Binney notes) give other examples, but they are not in the syllabus for this course.]

The statement that a functional integral is stationary is called a "**variational principle**". The differential equations equivalent to the variational principle are called the **Euler-Lagrange equations**.

Lecture 5

6 Symmetries, Invariance, and Conserved Quantities in the Lagrangian Formalism

What is meant by **symmetry**? Something stays the same or is **invariant** when a specific change is made. For example, I have a rectangular flower bed in front of my sitting room window containing two identical roses R symmetrically placed and two identical lilies L, as shown in figure 4.

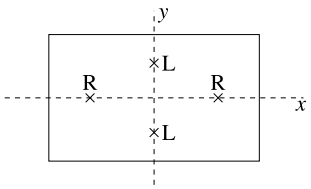


Figure 4: My flower bed.

There are two ways of looking at it:

- **Passive** point of view: Rotate co-ordinates $x \to -x$, $y \to -y$ (i.e. by 180°). The rose and lily configuration is the same in the new co-ordinate system as it was in the old.
- **Active** point of view: Physically move the two roses and two lilies, roots and all, to their new positions. The new configuration of roses and lilies is the same as the old.

We shall be dealing with symmetries of a dynamical system under infinitesimal changes. They may still be viewed passively (change in co-ordinate description) or actively (system altered to a different system). Either way, we shall find that the symmetry implies that a certain quantity is conserved (constant of motion).

6.1 Case 1: L independent of t

[$L\left(\{q_i\}, \{\dot{q}_i\}, t\right)$ in fact having no explicit dependence on t. i.e. $\frac{\partial L}{\partial t} = 0$]

Let $t \to t' = t + \delta t$ (**infinitesimal time translation**)

$$\delta L = \sum_i \left(\frac{\partial L}{\partial q_i} \delta q_i + \frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i \right) + \left(\frac{\partial L}{\partial t} \delta t \right)$$

where the last term is zero by assumption and

$$\begin{aligned} \delta q_i &= \dot{q}_i \delta t \\ \delta \dot{q}_i &= \frac{d}{dt} \delta q_i t \\ \delta L &= \frac{dL}{dt} \delta \end{aligned} \tag{11}$$

and

$$\frac{\partial L}{\partial \dot{q}_i} \frac{d}{dt} \delta q_i = -\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) \delta q_i + \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \delta q_i \right)$$

So,

$$\delta L = \sum_i \delta q_i \left[\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right] + \frac{d}{dt} \sum_i \left(\frac{\partial L}{\partial \dot{q}_i} \delta q_i \right)$$

Substituting in (11) gives

$$\delta L = \frac{dL}{dt} \delta t = \sum_i \delta q_i \left[\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right] + \frac{d}{dt} \sum_i \left(\frac{\partial L}{\partial \dot{q}_i} \dot{q}_i \delta t \right)$$

Now **for the actual motion of the system**, $[\quad]_i = 0$ by Lagrange, so the coefficient of δt

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

or

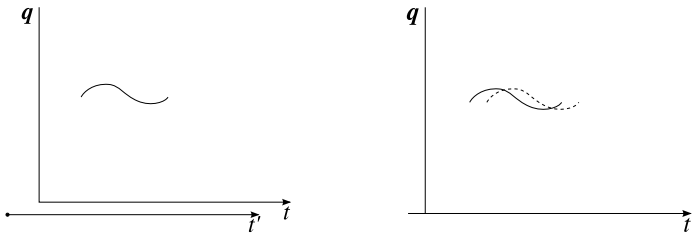
$$h = \sum_i \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i - L = \text{const.}$$

As a consequence of L having no explicit dependence on t, $h\left(\{q_i\}, \{\dot{q}_i\}\right)$ is a constant of the motion. Now h is at the moment a function of q's and $\dot{q}$'s. If the n equations $p_i = \frac{\partial L}{\partial \dot{q}_i}$ are solved for $\dot{q}$'s instead p's and the resulting formulae substituted into $h\left(\{q_i\}, \{\dot{q}_i\}\right)$ then the resulting function $H\left(\{q_i\}, \{p_i\}\right)$ is called the **Hamiltonian function**, the basis of the Hamiltonian formulation of mechanics. p_i is called the **generalised momentum conjugate to** q_i, and $H\left(\{q_i\}, \{p_i\}, t\right)$ (in general it may depend on t) is the basis for the $2n$ first-order equations called Hamilton's equations.

Example

Very often $L = T - V$ where V (as we have assumed) depends only on q_i's and T is quadratic in $\dot{q}_i$. Then $\sum_i \frac{\partial L}{\partial q_i} \dot{q}_i = 2T$ and $h = 2T - (T - V) = T + V$, which may be viewed as a deeper understanding of energy conservation. It is a consequence of **time translation invariance**, meaning:

- Passive point of view: a motion of the system with changed time co-ordinates looks like a motion with the original time co-ordinates.
- Active point of view: A motion starting at a later time will exactly reproduce the motion starting at the earlier time.



(a) Passive point of view (redefinition of time co-ordinate by $\delta t = t' - t$)

(b) Active point of view; dotted curve $\equiv$ identical curve but performed δt later

Figure 5: The passive and active points of view in this example.

6.2 Case 2: L independent of a specific generalised co-ordinate, q_0

(Ignorable or cyclic co-ordinate)

In this case we can say that L is symmetric of invariant under $q_0 \to q_0 +$ const.. As mentioned previously, Lagrange's equation for q_0 immediately gives $\frac{d}{dt}p_0 = 0$ where p_0 is the generalised momentum conjugate to q_0, i.e. $p_0 = \frac{\partial L}{\partial \dot{q}_0}$. So p_0 is a constant of the motion, as a consequence of the invariance of L under $q_0 \to q_0 +$ const.. We had an example of this in the case of the spherical pendulum.

6.3 Case 3: Emmy Noether's theorem

If L is invariant under the change of all q_i's $q_i \to q_i + \epsilon f_i(q_1 \dots q_n)$ where ϵ is an infinitesimal parameter, then

$$\begin{aligned} 0 = \frac{\delta L}{\epsilon} &= \sum_i \left(\frac{\partial L}{\partial q_i} f_i + \frac{\partial L}{\partial \dot{q}_i} \frac{df_i}{dt} \right) \\ &= \sum_i \left\{ \frac{\partial L}{\partial \dot{q}_i} f_i - f_i \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) + \frac{d}{dt} \left[f_i \frac{\partial L}{\partial \dot{q}_i} \right] \right\} \end{aligned}$$

For an actual motion, the first two terms cancel by Lagrange's equation, so we get the result

$$\frac{d}{dt} \left(\sum_i f_i \frac{\partial L}{\partial \dot{q}_i} \right) = 0 \quad \text{or} \quad \sum_i f_i p_i = \text{const.}$$

where $p_i = \frac{\partial L}{\partial \dot{q}_i}$, the generalised momentum conjugate to q_i.

Example

$\{q_i\} = (x, y, z)$ are the Cartesian co-ordinates of a particle. If L is invariant under rotation about, say, the z-axis, i.e.

$$\begin{aligned} x' &= x\cos\theta + y\sin\theta \\ y' &= -x\sin\theta + y\cos\theta \\ z' &= z \end{aligned}$$

or, for θ infinitesimal $= \epsilon$

$$\begin{array}{cc} \delta x = \epsilon y & f_x = y \\ \delta y = -\epsilon x & f_y = -x \\ \delta z = 0 & \end{array}$$

so, according to the Noether theorem, $yp_x - xp_y = \text{const.}$. This is in fact $-L_z$, where L_z is the angular momentum component about the z-axis.

It is easy to generalise to a system of N particles:

1. If L is invariant under $\mathbf{r}_i \to \mathbf{r}_i + \boldsymbol{\delta}$ (same $\boldsymbol{\delta}$ for all i) where $\mathbf{r}_i = (x_i, y_i, z_i)$, then
$$\sum_i \mathbf{p}_i = \text{const.}$$
i.e. Total momentum conservation is a consequence of spatial homogeneity.

2. If L is invariant under $\mathbf{r}_i \to \mathbf{r}_i + \boldsymbol{\omega} \wedge \mathbf{r}_i$ (same infinitesimal $\boldsymbol{\omega}$ vector - infinitesimal rotation), then
$$\sum_i \mathbf{r}_i \wedge \mathbf{p}_i = \text{const. vector}$$
i.e. total angular momentum conservation is a consequence of spatial isotropy.

Lecture 6

7 Linearised Lagrange equations near equilibrium and small oscillations

Let L be independent of t and of the form $L = T - V$ where $T = \frac{1}{2}\sum t_{ij}(q_1 \dots q_n)\,\dot{q}_i\dot{q}_j$ and, as usual, $V = V(q_1 \dots q_n)$.

Lagrange's equations are

$$\frac{d}{dt}\left[\sum_j t_{ij}(q_1 \dots q_n)\,\dot{q}_j\right] = -\frac{\partial V}{\partial q_i} + \sum_j \frac{\partial t_{kj}}{\partial q_i}\dot{q}_k\dot{q}_j$$

At equililbrium, $q_i = \text{const.}\,\forall i$, $\dot{q}_i = 0$, so $\frac{\partial V}{\partial q_i} = 0\,\forall i$, so there is a stationary point of $V(q_1 \ldots q_n)$, as expected. For each equilibrium point, we can define

$$q_i\,(\text{new}) = q_i\,(\text{old}) - q_i\,(\text{equilib.})\,; \qquad \dot{q}_i\,(\text{new}) = \dot{q}_i\,(\text{old})$$

so that we can expand in powers of the new q_i and $\dot{q}_i$

$$\begin{aligned} V &= V(0\ldots0) + \frac{1}{2}\sum_{ij} v_{ij} q_i q_j \\ T &= \frac{1}{2}\sum_{ij} t_{ij}\dot{q}_i\dot{q}_j \end{aligned}$$

where the coefficients v_{ij}, t_{ij} are constants and may be chosen so that $v_{ij} = v_{ji}$ and $t_{ij} = t_{ji}$; i.e. L is expressed in terms of two $n \times n$ constant symmetric matrices $\underline{\underline{v}} = (v_{ij})$ and $\underline{\underline{t}} = (t_{ij})$.

The constant $V(0\ldots0)$ can be dropped as it does not affect the equations.

In terms of the matrices, the equations are equivalent to the matrix equation

$$\underline{\underline{t}}\ddot{\mathbf{q}} = \underline{\underline{v}}\mathbf{q}$$

where $\mathbf{q}$ is the n-component column matrix of generalised co-ordinates.

Looking for particular solutions of the form $\mathbf{q} = \mathbf{a}e^{i\omega t}$ we get

$$-\omega^2 \underline{\underline{t}}\mathbf{a} = \underline{\underline{v}}\mathbf{a} \tag{12}$$

a (generalised) eigenvalue equation.

It is not difficult to convince oneself that ω^2 and the elements of $\mathbf{a}$ are real. For $\mathbf{a}$ complex, multiply both sides of (12) by the row matrix $\mathbf{a}^H$, the Hermitian conjugate of $\mathbf{a}$. Then $\mathbf{a}^H\underline{\underline{t}}\mathbf{a}$ and $\mathbf{a}^H\underline{\underline{v}}\mathbf{a}$ are both real. So ω^2 is real, so $\mathfrak{Re}\mathbf{a}$ and $\mathfrak{Im}\mathbf{a}$ both satisfy (12) separately, so $\mathbf{a}$ is real if non-degenerate, and may be chosen real otherwise. From (12), $\det\left(\omega^2\underline{\underline{t}} + \underline{\underline{v}}\right) = 0$ and in general therefore there are n roots for ω^2 (which may be degenerate). So, as in the usual eigenvalue problem there are n eigenvalues ω_i^2 and corresponding eigenvectors $\mathbf{a}_i$. ω_i^2 may be positive, negative, or zero. If all ω_i^2 are positive we will have **small oscillations** about the equilibrium. If any of the ω_i^2 is negative, $e^{\pm i\omega_i t}$ will contain exploding terms, we have unstable equilibrium and the quadratic approximation is not valid in general. If any $\omega_i = 0$, then the potential energy has a flat direction near the equilibrium point (see example below). One would need to investigate high orders in $q_i$to see whether the equilibrium point is indeed stable.

The different normal mode eigenvectors $\mathbf{a}_i$, $\mathbf{a}_j$ satisfy

$$-\omega_i^2\underline{\underline{t}}\mathbf{a}_i = \underline{\underline{v}}\mathbf{a}_i \tag{13}$$

$$-\omega_j^2\underline{\underline{t}}\mathbf{a}_j = \underline{\underline{v}}\mathbf{a}_j \tag{14}$$

Multiplying (13) by the transpose of $\mathbf{a}_j$ on the left and the transpose of (14) by $\mathbf{a}_i$ on the right, we find the orthogonality property

$$\mathbf{a}_j^T\underline{\underline{t}}\mathbf{a}_i = 0, \qquad \mathbf{a}_j^T\underline{\underline{v}}\mathbf{a}_i = 0 \qquad i \neq j$$

The **normal co-ordinates**, Q_i, are new generalised co-ordinates, each a linear combination of the q_i so that

$$\mathbf{q} = \sum Q_i \mathbf{a}_i$$

$$\mathbf{a}_j^T \underline{\underline{v}} \mathbf{q} = \sum_i Q_i \mathbf{a}_j^T \underline{\underline{v}} \mathbf{a}_i = Q_j$$

using the orthogonality property (i.e. $\mathbf{a}_j^T \underline{\underline{v}} \mathbf{a}_i = \delta_{ij}$, $Q_i(t) = Q_i e^{i\omega t}$), if the $\mathbf{a}_i$ are normalised to $\mathbf{a}_i^T \underline{\underline{v}} \mathbf{a}_i = 1$.

In terms of the normal co-ordinates,

$$L = \sum_i \left(-Q_i^2 + \dot{Q}_i^2 \right)$$

so in the quadratic approximation, zero frequency corresponds to L independent of the corresponding normal co-ordinate, so that the corresponding p_Q (generalised momentum conjugate to this normal co-ordinate) is a constant of the motion. This could be a consequence of the quadratic approximation to L, but it could be there is a physical reason for the symmetry to be exact.

7.1 Example: Free vibrations of a linear BAB molecule

We model this system with two springs as shown in figure 6 with natural length b, with masses of B, A being m and M respectively. We consider vibrations along the line of the molecule only. Then

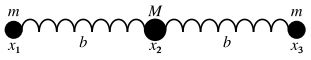


Figure 6: A BAB molecule.

$$\begin{aligned} V &= \frac{k}{2}(x_2 - x_1 - b)^2 + \frac{k}{2}(x_3 - x_2 - b)^2 \\ T &= \frac{1}{2}m\left(\dot{x}_1^2 + \dot{x}_3^2\right) + \frac{1}{2}M\dot{x}_2^2 \end{aligned}$$

Let q_i be the co-ordinates with respect to equilibrium positions x_{01}, x_{02}, x_{03}, i.e.

$$q_i = x_i - x_{0i}$$

$$b = x_{02} - x_{01} = x_{03} - x_{02}$$

Then

$$\begin{aligned} V &= \frac{k}{2}(q_2 - q_1)^2 + \frac{k}{2}(q_3 - q_2)^2 \\ &= \frac{k}{2}\left[q_1^2 + 2q_2^2 + q_3^2 - 2q_1q_2 - 2q_2q_3\right] \end{aligned}$$

so

$$V = \frac{1}{2}\begin{pmatrix} q_1 & q_2 & q_3 \end{pmatrix} \begin{pmatrix} k & -k & 0 \\ -k & 2k & -k \\ 0 & -k & k \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix}$$

and

$$\begin{aligned} T &= \frac{1}{2}m\left(\dot{q}_1^2+\dot{q}_3^2\right)+\frac{1}{2}M\dot{q}_2^2 \\ &= \frac{1}{2}\begin{pmatrix} \dot{q}_1 & \dot{q}_2 & \dot{q}_3 \end{pmatrix}\begin{pmatrix} m & 0 & 0 \\ 0 & M & 0 \\ 0 & 0 & m \end{pmatrix}\begin{pmatrix} \dot{q}_1 \\ \dot{q}_2 \\ \dot{q}_3 \end{pmatrix} \end{aligned}$$

The determinant equation for the normal mode frequencies is

$$\left|\underline{\underline{v}}-\omega^2\underline{\underline{t}}\right|=0$$

or

$$\begin{vmatrix} k-\omega^2 m & -k & 0 \\ -k & 2k-\omega^2 M & -k \\ 0 & -k & k-\omega^2 m \end{vmatrix}=0$$

Multiplying out by the top row:

$$\left(k-\omega^2 m\right)\left[\left(2k-\omega^2 M\right)\left(k-\omega^2 m\right)-k^2\right]+k\left[-k\left(k-\omega^2 m\right)\right]=0$$

Factorise out $\left(k-\omega^2 m\right)$

$$\begin{aligned} \left(k-\omega^2 m\right)\left[2k^2-\omega^2 k\left(M+2m\right)+\omega^4 Mm-k^2-k^2\right] &= 0 \\ \omega^2\left(k-\omega^2 m\right)\left(\omega^2 Mm-k\left(M+2m\right)\right) &= 0 \end{aligned}$$

So the three normal mode frequencies are

$$\omega_1=0,\ \omega_2=\sqrt{\frac{k}{m}},\ \omega_3=\sqrt{\frac{k}{m}\left(1+\frac{2m}{M}\right)}$$

The corresponding normal mode vectors are (unnormalised)

$$\mathbf{a}_1=\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},\ \mathbf{a}_2=\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix},\ \mathbf{a}_3=\begin{pmatrix} 1 \\ -\frac{2m}{M} \\ 1 \end{pmatrix}$$

Check the maths before proceeding. Is $\mathbf{a}_1$ orthogonal to $\mathbf{a}_3$?

$$\mathbf{a}_1\underline{\underline{t}}\mathbf{a}_3=\begin{pmatrix} 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} m & 0 & 0 \\ 0 & M & 0 \\ 0 & 0 & m \end{pmatrix}\begin{pmatrix} 1 \\ -\frac{2m}{M} \\ 1 \end{pmatrix}=0$$

$$\begin{aligned} \mathbf{a}_1\underline{\underline{v}}\mathbf{a}_3 &= \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ -\frac{2m}{M} \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1+\frac{2m}{M} \\ -2-\frac{4m}{M} \\ 1+\frac{2m}{M} \end{pmatrix}=0 \end{aligned}$$

What is the meaning of the zero frequency? In this mode, all q's are equal. The corresponding normal co-ordinate is

$$\begin{pmatrix} q_1 & q_2 & q_3 \end{pmatrix} \underline{t} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \propto \frac{mq_1 + Mq_2 + mq_3}{2m + M}$$

i.e. the centre of mass co-ordinate which must move with constant velocity, since there is no external force.

In this case, the zero frequency is simply a consequence of the translational symmetry of the system, and we could have reduced to two generalised co-ordinates by taking the centre of mass as the origin.

The mode 2 has M stationary and m moving symmetrically about it, not surprisingly with frequency $\sqrt{\frac{k}{m}}$. Mode 3 has the centre of mass at the origin, with 1 and 3 moving opposite to 2.

In less transparent examples, if a zero frequency occurs it is worth checking whether it is a consequence of an exact symmetry of the Lagrangian as in this example or whether it is a symmetry of the quadratic approximation to the Lagrangian which is independent of the corresponding normal co-ordinate and the full Lagrangian contains cubic and higher powers of the normal co-ordinate etc.

Lecture 7

8 Two Important Examples

8.1 Lagrangian for a particle in a rotating Cartesian co-ordinate system

Let (x, y, z) be co-ordinates in an inertial frame with $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$, where the unit vectors $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ are fixed. So $\frac{d\mathbf{r}}{dt} = \dot{x}\mathbf{i} + \dot{y}\mathbf{j} + \dot{z}\mathbf{k}$. Consider now $\mathbf{r}$ referred to a rotating frame; $\mathbf{i}'$, $\mathbf{j}'$, $\mathbf{k}'$ are now rotating unit vectors. We have

$$\frac{d\mathbf{i}'}{dt} = \boldsymbol{\omega} \wedge \mathbf{i}', \qquad \frac{d\mathbf{j}'}{dt} = \boldsymbol{\omega} \wedge \mathbf{j}', \qquad \frac{d\mathbf{k}'}{dt} = \boldsymbol{\omega} \wedge \mathbf{k}'$$

where $\boldsymbol{\omega}$ is the angular velocity. If x', y', z' are defined by $\mathbf{r} = x'\mathbf{i}' + y'\mathbf{j}' + z'\mathbf{k}'$ (co-ordinates with respect to rotating frame) then

$$\begin{aligned} \frac{d\mathbf{r}}{dt} &= \frac{d}{dt}\left(x'\mathbf{i}' + y'\mathbf{j}' + z'\mathbf{k}'\right) \\ &= \frac{dx'}{dt}\mathbf{i}' + \frac{dy'}{dt}\mathbf{j}' + \frac{dz'}{dt}\mathbf{k}' + x'\frac{d\mathbf{i}'}{dt} + y'\frac{d\mathbf{j}'}{dt} + z'\frac{d\mathbf{k}'}{dt} \\ \text{usually written} &= \dot{\mathbf{r}}' + \boldsymbol{\omega} \wedge \mathbf{r}' \end{aligned}$$

where the first term takes into account the variations of the co-ordinates (x', y', z') and the second the fact that the axes are rotating. So we have for a particle of mass m

$$\begin{aligned} L &= T - V \\ &= \frac{1}{2} m \left(\dot{\mathbf{r}}' + \boldsymbol{\omega} \wedge \mathbf{r}' \right)^2 - V \end{aligned}$$

where V is expressed as a function of the rotating co-ordinates (x', y', z') and quite likely time, even though V could be independent of t in a fixed co-ordinate frame.

Turning the handle to get the equations of motion (assuming $\boldsymbol{\omega}$ is fixed, otherwise there are more terms)

$$\begin{aligned} \frac{\partial L}{\partial \dot{x}'} &= m \left(\dot{x}' + \omega_y z' - \omega_z y' \right) \\ \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}'} \right) &= m \left(\ddot{x}' + \omega_y \dot{z}' - \omega_z \dot{y}' \right) \\ &= x\text{-component of } m \left(\ddot{\mathbf{r}}' + \boldsymbol{\omega} \wedge \dot{\mathbf{r}}' \right) \\ \frac{\partial L}{\partial x'} &= x\text{-component of } \left\{ -\nabla' V - m \boldsymbol{\omega} \wedge \dot{\mathbf{r}}' + \nabla' \frac{1}{2} m \left(\boldsymbol{\omega} \wedge \mathbf{r}' \right) \right\} \end{aligned}$$

So, altogether we have for all three components

$$m \ddot{\mathbf{r}}' = -\nabla' V - \nabla' \left[-\frac{1}{2} m \left(\boldsymbol{\omega} \wedge \mathbf{r}' \right)^2 \right] - 2 m \omega \wedge \dot{\mathbf{r}}'$$

The extra terms arising out of the rotation of the co-ordinate system are "fictitious forces", or "pseudoforces". The $[\ldots]$ is like the effect of an extra potential energy, the "centrifugal barrier potential", and the fictitious force is the centrifugal force. The last term $-2m\boldsymbol{\omega} \wedge \dot{\mathbf{r}}'$ is called the Coriolis force. See e.g. Kibble for a discussion of the effect of the Coriolis force in meteorology.

8.2 Lagrangian for a particle in an electromagnetic field

The force on a charged particle in an electromagnetic field is $Q \left(\mathbf{E} + \mathbf{v} \wedge \mathbf{B} \right)$ (the Lorentz force). To find a Lagrangian which gives the force equation

$$m \ddot{\mathbf{r}} = Q \left(\mathbf{E} + \dot{\mathbf{r}} \wedge \mathbf{B} \right) \tag{15}$$

we must first define the scalar and vector potentials ϕ, $\mathbf{A}$ (both, in general, functions of $\mathbf{r}$, t) by

$$\begin{aligned} \mathbf{E} &= -\nabla \phi - \frac{\partial \mathbf{A}}{\partial t} \\ \mathbf{B} &= \nabla \wedge \mathbf{A} \end{aligned}$$

[These satisfy two of Maxwell's equations $\nabla . \mathbf{B} = 0$ and $\nabla \wedge \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$ automatically; the other two Maxwell equations involve sources and imply equations to be satisfied by $\mathbf{A}$ and ϕ.]

Consider the following Lagrangian, pulled out of a hat, and we shall show that Lagrange's equations for it are exactly the Lorentz force equation (15).

$$L = \frac{1}{2}m\dot{\mathbf{r}}^2 + Q\left(\mathbf{A}.\dot{\mathbf{r}} - \phi\right)$$

$$\frac{\partial L}{\partial \dot{x}} = m\dot{x} + QA_x$$

$$\therefore \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) = m\ddot{x} + Q\frac{dA_x}{dt}$$

Now A_x is a function of in general x, y, z, t, so

$$\frac{d}{dt}A_x = \frac{\partial A_x}{\partial t} + \dot{\mathbf{r}}.\nabla A_x$$

by the chain rule.

$$\therefore \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) = m\ddot{x} + Q\left[\frac{\partial A_x}{\partial t} + \dot{x}\frac{\partial A_x}{\partial x} + \dot{y}\frac{\partial A_x}{\partial y} + \dot{z}\frac{\partial A_x}{\partial z}\right] \tag{16}$$

$$\frac{\partial L}{\partial x} = Q\left[\dot{x}\frac{\partial A_x}{\partial x} + \dot{y}\frac{\partial A_y}{\partial x} + \dot{z}\frac{\partial A_z}{\partial x} - \frac{\partial \phi}{\partial x}\right] \tag{17}$$

Equating (16)=(17), by Lagrange's x-equation, gives

$$\begin{aligned} m\ddot{x} &= Q\left[-\frac{\partial \phi}{\partial x} + \dot{y}\left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right) + \dot{z}\left(\frac{\partial A_z}{\partial x} - \frac{\partial A_x}{\partial z}\right) - \frac{\partial A_x}{\partial t}\right] \\ &= Q\left(E_x + \dot{y}B_z - \dot{z}B_y\right) \\ &= x\text{-component of } Q\left(\mathbf{E} + \dot{\mathbf{r}} \wedge \mathbf{B}\right) \end{aligned}$$

So this Lagrangian does indeed yield the correct equations of motion for a particle in an electromagnetic field. Note that it is **not** of the form $T-V$ where V depends on $\mathbf{r}$, t only. It may be written $T-U$ where $U = Q\left[\phi\left(\mathbf{r},\, t\right) - \mathbf{A}\left(\mathbf{r},\, t\right).\dot{\mathbf{r}}\right]$ (sometimes called a velocity-dependent potential).

8.2.1 Example: Charged particle in a constant uniform B-field

Check that, if $\mathbf{B}$ is independent of $\mathbf{r}$, t

$$\mathbf{A} = \frac{1}{2}\mathbf{B} \wedge \mathbf{r} \Rightarrow \mathbf{B} = \nabla \wedge \mathbf{A}$$

so

$$\begin{aligned} L &= \frac{1}{2}m\dot{\mathbf{r}}^2 + Q\dot{\mathbf{r}}.\frac{1}{2}\left(\mathbf{B} \wedge \mathbf{r}\right) \\ &= \frac{1}{2}m\dot{\mathbf{r}}^2 + \frac{Q}{2}\left(\mathbf{r} \wedge \dot{\mathbf{r}}\right).\mathbf{B} \end{aligned}$$

The second term may be written $\frac{Q}{2m}\boldsymbol{\ell}.\mathbf{B}$ where $\boldsymbol{\ell}$ is the orbital angular momentum ($\frac{Q}{2m}\boldsymbol{\ell}$ is the orbital magnetic moment). For $\mathbf{B}$ in the z-direction,

$$L = \frac{1}{2}m\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right) + \frac{QB}{2}\left(x\dot{y} - y\dot{x}\right)$$

giving the Lagrange equations

$$\begin{aligned} \frac{dv_x}{dt} &= \frac{QB}{m}v_y \\ \frac{dv_y}{dt} &= -\frac{QB}{m}v_x \\ m\frac{dv_z}{dt} &= 0 \end{aligned}$$

You have probably seen these equations before. The general solution is the corkscrew (or helix) motion

$$\begin{aligned} x &= \frac{1}{\omega}\left(a\sin\omega t - b\cos\omega t\right) + x_0 \\ y &= \frac{1}{\omega}\left(a\cos\omega t + b\sin\omega t\right) + y_0 \\ z &= v_z t + z_0 \end{aligned}$$

where $\omega = \omega_c = \frac{QB}{m}$, the cyclotron frequency.

8.2.2 Example: Charged particle subject to central electrostatic potential and in addition acted on by constant B-field

(Classical model of the normal Zeeman effect)

For $\mathbf{B}$ in the z-direction:

$$\begin{aligned} L &= \frac{1}{2}m\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right) - Q\Phi\sqrt{x^2+y^2+z^2} + \frac{1}{2}QB\left(x\dot{y} - y\dot{x}\right) \\ &= \frac{1}{2}m\left[\left(x - \frac{QB}{2m}y\right)^2 + \left(\dot{y} + \frac{QB}{2m}x\right)^2 + \dot{z}^2\right] - \frac{1}{2}m\omega_L^2\left(x^2+y^2\right) - Q\Phi \end{aligned}$$

where $\omega_L = \frac{QB}{2m} = \frac{1}{2}\omega_C$ is the Larmor frequency.

Now change co-ordinates to those in a frame rotating about the z-axis at ω_L (in the same direction as the cyclotron frequency)

$$\begin{aligned} \dot{x}' &= \dot{x} - \omega_L y \\ \dot{y}' &= \dot{y} + \omega_L x \\ \dot{z}' &= \dot{z} \\ x^2 + y^2 &= x'^2 + y'^2 \\ z' &= z \end{aligned}$$

$$\therefore L = \frac{1}{2}m\left(\dot{x}'^2 + \dot{y}'^2 + \dot{z}^2\right) - \frac{1}{2}m\omega_L^2\left(x'^2 + y'^2\right) - Q\Phi\sqrt{x'^2 + y'^2 + z'^2}$$

Note that this Lagrangian is second-order in B due to the ω_L^2 term.

Larmor's Theorem: To first order in B, the effect of a constant **B**-field is to superimpose on the motion a uniform rotation with angular velocity ω_L about **B**

For example,

1. $B = 0$: vector **L** fixed, not in z-direction $\Rightarrow B \neq 0$: vector **L** precesses about z-axis with angular frequency ω_L (vector model of atom)

2. $B = 0$: angular momentum vector **L** fixed, in z-direction; circular motion, say, with radius ρ, frequency ω,$\omega^2\rho$ =electrostatic force $\Rightarrow B \neq 0$ circular motion, frequency $\omega \pm \omega_L$ (normal Zeeman effect), with the $\pm$ entering as ω may be clockwise or anticlockwise.

Lecture 8

9 Lagrange Multipliers and Constraints

Recall from calculus that the stationary values of a function $f(x_1, x_2, \ldots, x_n)$ subject to the constraint $g(x_1, x_2, \ldots, x_n) = 0$ are given by the solutions of

$$\begin{aligned} \frac{\partial f}{\partial x_i} &= \lambda\frac{\partial g}{\partial x_i} \qquad x = 1 \ldots n \\ g(x_1 \ldots x_n) &= 0 \end{aligned} \tag{18}$$

where λ is an unknown constant to be determined; i.e. $n+1$ equations to determine $x_1 \ldots x_n$, λ.λ is called the Lagrange multiplier. This can be generalised to several constraint equations. The RHS of (18) now has a sum of terms with a Lagrange multiplier for each constraint.

The following generalises this technique to the stationary values of the action integral

$$\begin{aligned} A &= \int L\,dt \\ \delta A &= \int \sum_i \delta q_i \left(\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i}\right) dt \end{aligned}$$

When the δq_i were entirely independent, we could conclude that each $(\ldots) = 0$, Lagrange's equations. If the q's are subject to a constraint, this deduction

can no longer be made. We consider only **holonomic constraints** of the form $g(q_1 \dots q,\, t) = 0$, from which it follows that the only allowed δq_i satisfy

$$\sum_i \delta q_i \frac{\partial g}{\partial q_i} = 0$$

Consider now

$$\delta \int (L - \lambda(t)\, g)\, dt = \int \sum_i \delta q_i \left[\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} - \lambda \frac{\partial g}{\partial q_i} \right]$$

The δq_i are not independent but subject to $\sum_i \frac{\partial g}{\partial q_i} \delta q_i = 0$ at each t.

Consider $\delta q_1 \dots \delta q_{n-1}$ independent and δq_n given then by the constraint equation

$$\delta q_n = -\frac{\sum_{\ell=1}^{n-1} \frac{\partial g}{\partial q_i} \delta q_i}{\frac{\partial g}{\partial q_n}}$$

Choose $\lambda(t)$ such that

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_n}\right) - \frac{\partial L}{\partial q_n} - \lambda \frac{\partial g}{\partial q_n} = 0 \tag{19}$$

for each t.

Then the remaining $\delta q_1 \dots q_{n-1}$ are independent and so their coefficients

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

for $i = 1 \dots n-1$. (19) shows the equation is true also for q_n.

9.1 Example: The Simple Pendulum Again

Suppose the length of the pendulum, ℓ, is not constant but depends on time $\ell(t)$; for example, because of heat expansion. We can treat this problem by treating it in terms of one generalised co-ordinate, θ. The corresponding Lagrangian, L_1, is

$$L_1 = \frac{1}{2} m \ell^2 \dot{\theta}^2 + mg\ell \cos\theta$$

giving the Lagrange equation

$$\frac{d}{dt}\left(m\ell^2 \dot{\theta}\right) = -mg\ell \sin\theta \tag{20}$$

where now ℓ is the given function of t.

Alternatively, we can treat it as a problem with two generalised co-ordinates, r, θ subject to the constraint $r = \ell(t)$

$$L_2 = \frac{1}{2} m \left(r^2 \dot{\theta}^2 + \dot{r}^2\right) + mgr \cos\theta$$

We now have the constraint $r = \ell(t)$, i.e. $g(r, \theta, t) = r - \ell(t)$ in the previous notation.

Now introduce with the Lagrange multiplier, λ ($\frac{\partial g}{\partial r} = 1$, $\frac{\partial g}{\partial \theta} = 0$)

$$\theta \text{ equation:} \quad \frac{d}{dt}\left(mr^2\dot{\theta}\right) = -mgr\sin\theta \tag{21}$$

$$r \text{ equation:} \quad \frac{d}{dt}(m\dot{r}) = mr\dot{\theta}^2 + mg\cos\theta + \lambda \tag{22}$$

$$\text{constraint equation:} \quad r = \ell(t) \tag{23}$$

Equation (21) with (23) is equivalent to (20).

The first term on the right-hand side of (22) is the centrifugal pseudoforce, the second the radial component of gravity. There is a third radial force on m, the tension force. Thus the physical interpretation of the Lagrange multiplier λ is that it is the $(-)$ tension force. ($(-)$ because it acts radially inwards). The Lagrange multiplier is the **force of constraint** (in the radial direction). Going back to the old simple pendulum problem with $\ell = \text{const.}$, (22) becomes $m\ell\dot{\theta}^2 + mg\cos\theta + \lambda = 0$, and θ may be determined by (21) with $r = \ell$. Hence λ may be found.

For small amplitude oscillations, $\lambda = -mg$. Even this trivial example shows that the Lagrange multiplier method may be used to find the **generalised force of constraint**. In

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = \frac{\partial L}{\partial q_i} + \lambda\frac{\partial g}{\partial q_i}$$

$\frac{\partial L}{\partial q_i}$ is called the ith generalised force. $\lambda\frac{\partial g}{\partial q_i}$ is the extra generalised force component corresponding to the constraint.

9.2 Example

A particle otherwise free ($L = T = \frac{1}{2}m\left(\dot{x}^2 + \dot{y}^2\right)$) is constrained to move on the circle $x^2 + y^2 = \ell^2$, i.e.

$$g(x, y) = \frac{1}{2}\left(x^2 + y^{2-\ell^2}\right) = 0$$

is the constraint

$$\begin{aligned} \frac{\partial g}{\partial x} &= x \\ \frac{\partial g}{\partial y} &= y \end{aligned}$$

$$L - \lambda g = \frac{1}{2}m\left(\dot{x}^2 + \dot{y}^2\right) - \lambda\left(x^2 + y^2 + \ell^2\right)$$

Lagrange's equations are then

$$\begin{aligned} m\ddot{x} &= -\lambda x \quad (24)\\ m\ddot{y} &= -\lambda y \quad (25)\\ \therefore m\dot{x}\ddot{x} &= -\lambda x\dot{x}\\ m\dot{y}\ddot{y} &= -\lambda y\dot{y} \end{aligned}$$

(where λ may **not** be assumed constant.)

$$\frac{d}{dt}\frac{1}{2}m\left(\dot{x}^2+\dot{y}^2\right) = -\frac{\lambda}{2}\frac{d}{dt}\left(x^2+y^2\right) = 0 \text{ by constraint}$$

$$\therefore \dot{x}^2+\dot{y}^2 = \text{const.} = v^2, \text{ say}$$

Add $x(24)+y(25)$:

$$m\left(x\ddot{x}+y\ddot{y}\right) = -\lambda\left(x^2+y^2\right) = -\lambda\ell^2$$

But

$$x\ddot{x}+y\ddot{y} = \frac{d}{dt}\left(x\dot{x}+y\dot{y}\right) - \left(\dot{x}^2+\dot{y}^2\right) = \frac{d^2}{dt^2}\left(x^2+y^2\right) - v^2 = -v^2$$

But by (24), (25),

$$\text{LHS} = -\frac{\lambda}{m}\left(x^2+y^2\right) = -\frac{\lambda}{m}\ell\%^2$$

$$\therefore \lambda = \frac{mv^2}{\ell^2}$$

so $\lambda$is constant in this case. Solution of (24), (25) is just the usual motion in a circle with constant speed, $-\lambda\,(x,\,y)$ = tension force of magnitude $\frac{mv^2}{\ell}$ in radial inward direction. It is nice to see the method working, but there is not much point int he above other than as a pedagogic exercise.

Lecture 9

10 Hamiltonian Mechanics

This is an alternative formulation of mechanics which is not necessarily more useful for solving specific simple problems, but:

1. It is more general, and so useful for understanding the mathematic structure ("symplectic geometry") and attempting to answer "in principle" questions such as:

- Is the solar system stable?
- When does a mechanical system yield chaotic motion? When is it ordered (the opposite)?

2. It gives further understanding of constants of motion going beyond symmetries of the Lagrangian

3. It is particularly close to the Hamiltonian formulation of quantum mechanics, and its equations can be reinterpreted as quantum equations.

The Lagrangian approach is based on configuration space described by n generalised co-ordinates $\mathbf{q}_i = (q_1 \dots q_n)$ and second-order equations in the time. The choice of generalised co-ordinates may be made for convenience, with the quantity L having in general a different functional form depending on the choice of generalised co-ordinates. Alternatively, the Lagrange equation may trivially be rewritten as $2n$ first-order equations b treating the generalised velocities $(\dot{q}_i \dots \dot{q}_n)$ as an additional set of co-ordinates $= (v_1 \dots v_n)$.

The Hamiltonian formulation is based on taking the n further "co-ordinates" as the n generalised momenta conjugate to the q_i, i.e. $p_i = \frac{\partial L}{\partial \dot{q}_i}$. This choice turns out to expose the p_i as very similar in their properties to the q_i (just some sign changes). Because of this similarity between q's and p's, Hamilton's equations allow a very much larger choice of "**co-ordinates in phase space**", the space in which the Hamiltonian formalism replaces configuration space of the Lagrangian formalism. In the Lagrangian formalism we can change generalised co-ordinates arbitrarily $Q_i = f_i(q_1, q_2, \dots q_n, t)$ (point transformation) with the P_i following from $P_i = \frac{\partial L}{\partial \dot{Q}_i}$. In the Hamiltonian formalism the now $(Q_1 \dots Q_n, P_1 \dots P_n)$ can be chosen in more general ways with Qs functions of q's and p's (i.e. not just q's) and Ps similarly. Allowed transformations are called **canonical transformations**, and we shall have more to say about that later. First we show how Hamilton's formalism may be deduced from Lagrange's.

10.1 Hamilton's equations from Lagrange's

Let a system be defined b the Lagrangian $L(\{q_i\}, \{\dot{q}_i\}, t)$. Then δL is given by the chain rule

$$\delta L = \sum_i \left(\frac{\partial L}{\partial q_i} \delta q_i + \frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i \right) + \frac{\partial L}{\partial t} \delta t$$

Now define

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

so

$$\delta H = \sum_i \left(\underline{p_i \delta \dot{q}_i} + \delta p_i \dot{q}_i \right) - \sum_i \left(\underline{\frac{\partial L}{\partial q_i} \delta q_i} + \frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i \right) - \frac{d}{dt} \delta t$$

Now $p_i = \frac{\partial L}{\partial \dot{q}_i}$ by definition, so the underlined terms cancel. Further, we must consider H defined above as functions of p's, q's and t, so

$$\delta H = \sum_i \left(\frac{\partial H}{\partial p_i} \delta p_i + \frac{\partial H}{\partial q_i} \delta q_i \right) + \frac{\partial H}{\partial t} \delta t$$

Comparing the two expressions for δH, we conclude

$$\frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}$$

$$\boxed{\begin{aligned} \dot{q}_i &= \frac{\partial H}{\partial p_i} \\ \dot{p}_i &= -\frac{\partial H}{\partial q_i} \end{aligned}}$$

where the last equation follows from the Lagrange equation $\dot{p}_i = \frac{\partial L}{\partial \dot{q}_i}$.

The boxed equations, exhibiting the pleasing "symmetry" between p's and q's (but note the important $-$ sign!) are Hamilton's equations of motion, $2n$ first-order equations entirely equivalent to the n second-order equations of Lagrange.

Example: To show using the Hamiltonian equations that if H has no explicit dependence on t, then it is a constant of the motion.

$$\frac{dH}{dt} = \sum_i \left(\frac{\partial H}{\partial p_i} \dot{p}_i + \frac{\partial H}{\partial q_i} \dot{q}_i \right)$$

But $\dot{p}_i = -\frac{\partial H}{\partial q_i}$, $\dot{q}_i = \frac{\partial H}{\partial p_i}$, so $\frac{dH}{dt} = 0$ giving H to be a constant of the motion.

10.2 Hamiltonian for a charged particle in an electromagnetic field

We already obtained the Lagrangian

$$L = \frac{1}{2} m \dot{\mathbf{r}}^2 + Q \left[\dot{\mathbf{r}}.\mathbf{A} \left(\mathbf{r}, t \right) - \phi \left(\mathbf{r}, t \right) \right]$$

so the momentum conjugate to $\dot{\mathbf{r}}$ is not the mechanical momentum $m\dot{\mathbf{r}}$, but

$$\mathbf{p} = m\dot{\mathbf{r}} + Q\mathbf{A}$$

[Binney's problem set 2, question 7 (difficult!) shows that (for static fields) $Q\mathbf{A}$ is the moment stored in the electromagnetic field surrounding the charge Q. Therefore $\mathbf{p}$ is the sum of mechanical and electromagnetic field momentum.]

We find

$$\begin{aligned} H &= \dot{\mathbf{r}}.\mathbf{p} - L, \quad \text{expressed in terms of } \mathbf{p} \text{ and } \mathbf{r}. \\ &= \frac{1}{2m}\left(\mathbf{p} - Q\mathbf{A}\right)^2 + Q\phi \end{aligned}$$

This very important formula is used in quantum mechanics with $\mathbf{p} \to -i\hbar\nabla$ to obtain the Schrödinger equation for atomic electrons in interaction with the electromagnetic field (static fields and/or radiation). The only missing ingredients are electron spin and relativity.

10.3 Hamiltonian for a particle in a rotating co-ordinate system.

We had for the Lagrangian of a particle in terms of rotating Cartesian co-ordinates $\mathbf{r}' = (x',\, y',\, z')$

$$L = \frac{1}{2}m\left(\dot{\mathbf{r}}' + \boldsymbol{\omega} \wedge \mathbf{r}'\right)^2 - V\left(\mathbf{r}',\, t\right)$$

where the time dependence of V may be either intrinsic or induced by the rotating frame or both. We get, therefore, $H = \mathbf{p}'.\dot{\mathbf{r}}' - L$ expressed in terms of $\mathbf{p}'$, $\mathbf{r}'$ where $\mathbf{p}' = m\left(\dot{\mathbf{r}}' + \boldsymbol{\omega} \wedge \mathbf{r}'\right)$. This gives

$$H = \frac{\mathbf{p}'^2}{2m} - \boldsymbol{\omega}.\left(\mathbf{r}' \wedge \mathbf{p}'\right) + V\left(\mathbf{r}',\, t\right)$$

Now $\mathbf{p}' = m\left(\dot{\mathbf{r}}' + \boldsymbol{\omega} \wedge \mathbf{r}'\right) = m($**velocity** in **non**-rotating frame coincident at the given time with the rotating frame), so$\mathbf{r}' \wedge \mathbf{p}' = \mathbf{L}$, the angular momentum (**not** the Lagrangian!!). So,

$$H = \frac{\mathbf{p}'^2}{2m} - \boldsymbol{\omega} \wedge \mathbf{L} + V\left(\mathbf{r}',\, t\right)$$

where $\mathbf{L}$ is the orbital angular momentum.

It is amusing to note that if $\boldsymbol{\omega} =$ const. and V is independent of the angle about $\boldsymbol{\omega}$ and of t then H is conserved but not the energy (Energy $= \frac{\mathbf{p}'^2}{2m} + V$), but in this case $\boldsymbol{\omega}.\mathbf{L}$ is conserved, so the "new" conservation law is **not** new. On the other, if V is independent of t in the inertial frame, then energy is conserved, but in general H is not conserved if V has a time-dependence in the non-inertial frame.

Lecture 10

10.4 H for the harmonic oscillator

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

$$p = \frac{\partial L}{\partial q} = m\dot{q}$$

$$\begin{aligned} H &= p\dot{q} - L \quad \text{expressed in terms of } p,\, q \\ &= \frac{p^2}{2m} + \frac{1}{2}m\omega^2 q^2 \end{aligned}$$

H has no explicit dependence on t, so H is a constant of the motion, $H = E$.

The **orbit in phase space** is therefore an ellipse, and Hamilton's equations are

$$\begin{aligned} \dot{p} &= -m\omega^2 q \\ \dot{q} &= \frac{p}{m} \end{aligned}$$

Differentiating the second equation and substituting in the first, $\ddot{q} = -\omega^2 q$. This has general solution

$$\begin{aligned} q &= a\cos\omega t + b\sin\omega t \\ p &= m\dot{q} = ma\left(-a\sin\omega t + b\cos\omega t\right) \end{aligned}$$

Orbit traced clockwise around the ellipse.

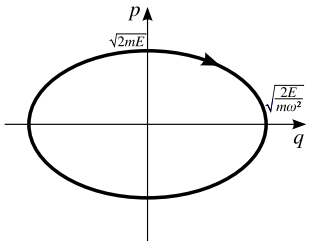


Figure 7: The orbit in phase space of the harmonic oscillator

This all goes to show that the harmonic oscillator is simple to solve however you do it! Now I wish to use the harmonic oscillator as a peg on which to illustrate some general properties of Hamiltonian theory.

10.4.1 "Volume" in phase space independent of time

Imagine many different oscillators, all governed by the same Hamiltonian, with different initial conditions $(q(0), p(0))$ so that they form a "dense gas" of points in phase space, occupying at time $t = 0$ a "volume" $V = \iint dp(0)\, dq(0)$. At time t they will occupy the volume $V' = \iint dp(t)\, dq(t)$. We show explicitly that $V' = V$. (The shape of the region occupied in phase space may alter, but its "volume" stays fixed. It moves in phase space like an incompressible fluid!)

$$\begin{aligned} p(t) &= m\omega\left(-a\sin\omega t + b\cos\omega t\right) \\ q(t) &= a\cos\omega t + b\sin\omega t \end{aligned}$$

We need to show that the Jacobian determinant

$$J = \left|\frac{\partial\left(q(t),\, p(t)\right)}{\partial\left(q(0),\, p(0)\right)}\right|$$

is 1. To do this, use

$$J = \left|\tfrac{\partial(q(t),\, p(t))}{\partial(a,\, b)}\right| \Big/ \left|\tfrac{\partial(q(0),\, p(0))}{\partial(a,\, b)}\right| \tag{26}$$

$$\frac{\partial\left(q\left(t\right),\,p\left(t\right)\right)}{\partial\left(a,\,b\right)}=\left|\begin{array}{cc}\cos\omega t & -m\omega\sin\omega t\\ \sin\omega t & m\omega\cos\omega t\end{array}\right|=m\omega,\text{ independent of }t$$

So the ratio (26) = 1.

10.4.2 Canonical transformations

If $P\left(p,\,q\right),\,Q\left(p,\,q\right)$ are such that, defining $G\left(P,\,Q\right)=H\left(p,\,q\right)$ (i.e. the same physical quantity but in general a different function of the two variables - c.f. discussion of $L'\left(q',\,\dot{q}'\right)=L\left(q,\,\dot{q}\right)$), we have

$$\left.\begin{array}{l}\dfrac{\partial H}{\partial p}=\dot{q}\\ -\dfrac{\partial H}{\partial q}=\dot{p}\end{array}\right\}\Leftrightarrow\left\{\begin{array}{l}\dfrac{\partial G}{\partial P}=\dot{Q}\\ -\dfrac{\partial G}{\partial Q}=\dot{P}\end{array}\right.\tag{27}$$

i.e. "Gamilton" equations for $(Q,\,P)$ are equivalent to Hamilton equations for $(q,\,p)$, we say that $(p,\,q)\rightarrow(P,\,Q)$ is a (time-independent) **canonical transformation**. (27) is equivalent to saying that Hamilton's equations are **covariant** with respect to canonical transformations.

Examples:

1. $P=\frac{p}{\sqrt{m\omega}},\,Q=\sqrt{m\omega}q$

$$H\left(p,\,q\right)=\frac{p^2}{2m}+\frac{1}{2}m\omega^2q^2=\frac{1}{2}\omega\left(P^2+Q^2\right)=G\left(P,\,Q\right)$$

$$\dot{q}=\frac{1}{\sqrt{m\omega}}\dot{Q},\,\dot{p}=\sqrt{m\omega}\dot{P},\,\text{so}\,\dot{q}=\frac{p}{m}\quad\Rightarrow\quad\dot{Q}=\sqrt{\frac{\omega}{m}}\dot{p}=\omega P=\frac{\partial G}{\partial P}$$

$$\text{and similarly}\qquad\dot{P}=-\omega Q=-\frac{\partial G}{\partial Q}$$

In fact, we always have a canonical transformation if $q,\,p$ are scaled in opposite ways $Q=kq,\,P=\frac{p}{k}$. This particular transformation corresponds to a covariance of Lagrange's equation, since $Q=kq\Rightarrow\frac{\partial L'}{\partial Q}=\frac{1}{k}\frac{\partial L}{\partial Q}$.
Note that $dPdQ=dpdq$, so the volume element in phase space is unaltered by canonical transformation.

2. I now show that "**rotation in phase space**" is a canonical transformation. I start from $(Q,\,P)$ and $G\left(Q,\,P\right)=\frac{1}{\omega(P^2+Q^2)}$ just because the algebra is a little easier, but in fact as we shall see later, whether or not a transformation is canonical does **not** depend on the form of the Hamiltonian! (This is not obvious in what I have said so far.)

Let

$$\begin{aligned} Q' &= Q\cos\delta + P\sin\delta \quad (28)\\ P' &= -Q\sin\delta + P\cos\delta \quad (29) \end{aligned}$$

(rotation of Q, P in phase space; note again that the "volume" element in phase space is unaltered). From (28), (29),

$$\begin{aligned} \frac{\partial}{\partial Q'} &= \cos\delta\frac{\partial}{\partial Q} - \sin\delta\frac{\partial}{\partial P} \quad (30)\\ \frac{\partial}{\partial P'} &= \sin\delta\frac{\partial}{\partial Q} + \cos\delta\frac{\partial}{\partial P} \end{aligned}$$

Now

$$\begin{aligned} \dot{Q}' &= \cos\delta\dot{Q} + \sin\delta\dot{P}\\ &= \cos\delta\frac{\partial G}{\partial P} - \sin\delta\frac{\partial G}{\partial Q} \text{ by Gamilton equation}\\ &= \frac{\partial G'}{\partial P'} \end{aligned}$$

by (30), where $G'(Q', P') = G(Q, P)$; and similarly

$$\dot{P}' = -\frac{\partial G'}{\partial Q'}$$

Note that the covariance of Hamilton's equations under rotation in phase space is **not** connected to a covariance of Lagrange's equations, because it is mixing up Q and P.

10.5 Infinitesimal generators of canonical transformations

Let

$$\begin{aligned} q \to q' &= q + \delta q\\ p \to p' &= p + \delta p \end{aligned}$$

or, equivalently,

$$\begin{aligned} q &= q' - \delta q\\ p &= p' - \delta p \end{aligned}$$

where

$$\begin{aligned} \delta q &= \delta\lambda\frac{\partial K}{\partial p}\\ \delta p &= -\delta\lambda\frac{\partial K}{\partial q} \end{aligned}$$

with $K(p, q)$ **any** given function of p, q and $\delta\lambda$ infinitesimal. I show later, using Poisson brackets, that $(q, p) \to (q', p')$ is (to first order in $\delta\lambda$) a canonical transformation. $K(p, q)$ is called an **infinitesimal generator of canonical transformation**. The general solution $(q(\lambda), p(\lambda))$ to the simultaneous differential equations

$$\frac{dq}{d\lambda} = \frac{\partial K}{\partial p} \tag{31}$$

$$\frac{dp}{d\lambda} = -\frac{\partial K}{\partial q} \tag{32}$$

for arbitrary $(q(0), p(0))$ gives, for each λ, a (finite) canonical transformation. Note the similarity of these equations to Hamilton's equations ($\lambda \to t$, $K \to H$). So the Hamiltonian may be regarded as the infinitesimal generator of canonical transformations giving the path of the system in phase space. Solutions $(q(\lambda), p(\lambda))$ to the simultaneous equations (31) and (32) are called "**flows**".

11 Poisson Brackets, Constants of Motion, and Symmetries (Invariance)

How does $M(p, q)$ change under the infinitesimal canonical transformation generated by $K(p, q)$?

$$\begin{aligned} \delta M &= \frac{\partial M}{\partial p}\delta p + \frac{\partial M}{\partial q}\delta q \\ &= \left(-\frac{\partial M}{\partial p}\frac{\partial k}{\partial q} + \frac{\partial M}{\partial q}\frac{\partial K}{\partial p}\right)\delta\lambda \end{aligned}$$

The quantity $\left(-\frac{\partial M}{\partial p}\frac{\partial k}{\partial q} + \frac{\partial M}{\partial q}\frac{\partial K}{\partial p}\right)$ is denoted $[M, K]$, the **Poisson bracket** of M with K.

Suppose now we choose as our two functions:

1. $K(p, q)$ (as above), the infinitesimal generator of canonical transformations
2. $H(p, q)$ the Hamiltonian.

Then,

1. $\delta K = \delta\lambda [K, H]$ under infinitesimal canonical transformations generated by H
2. $\delta H = -\delta\lambda [K, H]$ under infinitesimal canonical transformations generated by K

So

1. implies that if $[K,\, H] = 0$, K does not change under the transformation generated by H. But this is just the path of the system in phase space as governed by the Hamiltonian H, i.e. K is conserved.

2. implies that if $[K\, H] = 0$ the Hamiltonian **does not change in form** $(G = H)$ under the canonical transformation generated by K (i.e. symmetry or invariance). Thus the vanishing of the Poisson bracket $[K,\, H]$ implies bot K is a constant of the motion for Hamiltonian H and H is symmetric (invariant) under the infinitesimal canonical transformation generated by K.

   This is the relationship between symmetries and constants of the motion in the Hamiltonian formalism.

Lecture 11

11.1 Invariance of the Poisson Bracket under Canonical Transformation

The definition

$$[K,\, L] = \frac{\partial K}{\partial q}\frac{\partial L}{\partial p} - \frac{\partial K}{\partial p}\frac{\partial L}{\partial q} \tag{33}$$

seems to depend on which set of co-ordinates in phase space $(q,\, p)$ are chosen. With a different choice $(q',\, p')$ and

$$\begin{aligned} K'\,(q',\, p') &= K\,(q,\, p) \\ L'\,(q',\, p') &= L\,(q,\, p) \end{aligned}$$

(i.e. same physical quantities but different functions)

$$[K',\, L'] = \frac{\partial K'}{\partial q'}\frac{\partial L'}{\partial p'} - \frac{\partial K'}{\partial p'}\frac{\partial L'}{\partial q'} \tag{34}$$

If (33) and (34) were different, the Poisson bracket would not be a property of the physical quantities but would depend on the co-ordinates in phase space chosen. In fact (33)=(34), so the Poisson bracket has an invariant meaning in the same way as in 3-dimensional vector analysis,

$$\begin{aligned} \nabla u.\nabla v+ &= \frac{\partial u}{\partial x}\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\frac{\partial v}{\partial y} + \frac{\partial u}{\partial z}\frac{\partial v}{\partial z} \\ &= \begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \end{pmatrix} I \begin{pmatrix} \frac{\partial v}{\partial x} \\ \frac{\partial v}{\partial y} \\ \frac{\partial v}{\partial z} \end{pmatrix} \end{aligned} \tag{35}$$

(where I is the unit matrix). $[K,\, L]$ may be written, analogously to (35)

$$[K,\,L] = \begin{pmatrix} \frac{\partial K}{\partial q} & \frac{\partial K}{\partial p} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} \frac{\partial L}{\partial q} \\ \frac{\partial L}{\partial p} \end{pmatrix} \tag{36}$$

Now the form (35) is invariant under orthogonal transformations $O^T O = 1$. The analogous transformations in 2-dimensional phase space are via 2×2 matrices $\underline{\underline{S}}$ with $S^T \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} S = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Such matrices $\underline{\underline{S}}$ are called **symplectic matrices**. In general, the entries in $\underline{\underline{S}}$ are functions of q and p (or equivalently of q' and p'). It is easy to check, however, that the following **constant** matrices are symplectic (and in fact, any 2×2 matrix of determinant 1 is symplectic):

1. $\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$ rotation in phase space
2. $\begin{pmatrix} e^{\lambda} & 0 \\ 0 & e^{-\lambda} \end{pmatrix}$ scale transformation in opposite senses of q, p
3. $\begin{pmatrix} \cosh\theta' & \sinh\theta' \\ \sinh\theta' & \cosh\theta' \end{pmatrix}$ "hyperbolic rotation"

So (a), (b), (c) all correspond to canonical transformations. (a) and (b) we have met before; these are the transformations generated by $K = q^2 + p^2$ and $K = qp$ respectively. (c) is new and is generated by $K = q^2 - p^2$.

Any canonical transformations corresponds to a symplectic matrix, and vice-versa. I show part of this by demonstrating that, if $(p,\,q) \to (P,\,Q)$ preserves the Poisson bracket, then Hamilton's equations in $(p,\,q)$ imply Hamilton's equations in $(P,\,Q)$. Recall that this was our original definition of a canonical transformation, one for which Hamilton's equations were covariant. To show this, note that if Hamilton's equations in p, q are satisfied, then

$$\begin{aligned} \dot{P} = \frac{\partial P}{\partial p}\dot{p} + \frac{\partial P}{\partial q}\dot{q} &= \frac{\partial P}{\partial p}\frac{\partial H}{\partial q} + \frac{\partial P}{\partial q}\frac{\partial H}{\partial p} \\ &= [P,\,H]_{p,\,q} \end{aligned}$$

(where the subscript means the Poisson bracket defined in terms of derivatives with respect to p and q).

But we are to assume $(p,\,q) \to (P,\,Q)$ preserves the Poisson bracket (i.e. this is a symplectic transformation)

$$\begin{aligned} \therefore \dot{P} = [P,\,H]_{P,\,Q} &= \frac{\partial P}{\partial P}\frac{\partial H}{\partial Q} - \frac{\partial P}{\partial Q}\frac{\partial H}{\partial P} \\ &= \frac{\partial H}{\partial Q} \end{aligned}$$

and similarly

$$\dot{Q} = -\frac{\partial H}{\partial P}$$

(Note, here I have used H for the same quantity, whether or not it is treated as a function of p, q or P, Q. In our previous treatment, I would have called $\frac{\partial H}{\partial Q}$ "$\frac{\partial G}{\partial Q}$" where the function $G(P, Q)$ was the "Gamiltonian".)

We have one further loose end to tie up. We asserted that infinitesimal generators do indeed generate canonical transformations, but did not prove this at the time. We do this now by showing that the infinitesimal transformation generated by $K(p, q)$ viz

$$\begin{aligned} q' &= q + \delta\lambda \frac{\partial K}{\partial p} \\ p' &= p - \delta\lambda \frac{\partial K}{\partial q} \end{aligned}$$

is symplectic, i.e. if $M = \begin{pmatrix} \frac{\partial p'}{\partial p} & \frac{\partial p'}{\partial q} \\ \frac{\partial q'}{\partial p} & \frac{\partial q'}{\partial q} \end{pmatrix}$ then $M^T J M = J$ where $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

We work out $M^T J M$ using $M = \begin{pmatrix} 1 - \delta\lambda \frac{\partial^2 K}{\partial p \partial q} & -\delta\lambda \frac{\partial^2 K}{\partial q^2} \\ \delta\lambda \frac{\partial^2 K}{\partial p^2} & 1 + \delta\lambda \frac{\partial^2 K}{\partial p \partial q} \end{pmatrix}$, giving

$$\begin{aligned} M^T J M &= \begin{pmatrix} 1 - \delta\lambda \frac{\partial^2 K}{\partial p \partial q} & -\delta\lambda \frac{\partial^2 K}{\partial p^2} \\ \delta\lambda \frac{\partial^2 K}{\partial q^2} & 1 + \delta\lambda \frac{\partial^2 K}{\partial p \partial q} \end{pmatrix} \begin{pmatrix} \delta\lambda \frac{\partial^2 K}{\partial p^2} & 1 + \delta\lambda \frac{\partial^2 K}{\partial p \partial q} \\ -1 + \delta\lambda \frac{\partial^2 K}{\partial p \partial q} & \delta\lambda \frac{\partial^2 K}{\partial q^2} \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \end{aligned}$$

to first order, as anticipated.

11.2 Variational principle for motion in phase space

Recall that Lagrange's equations follow from the action principle: $\int L\, dt$ is stationary for the actual motion (keeping initial and final positions in configuration space fixed). Now,

$$H = \sum_i \dot{q}_i p_i - L \Rightarrow L = \sum_i \dot{q}_i p_i - H(q, p)$$

so let us consider the consequences of a **path in phase space** making $\int L\, dt$ stationary (the "action principle" in phase space)

$$\delta \int_{t_1}^{t_2} L\, dt = \int_{t_1}^{t_2} \sum_i \left(\dot{q}_i \delta p_i + p_i \delta \dot{q}_i - \frac{\partial H}{\partial q_i} \delta q_i - \frac{\partial H}{\partial p_i} \delta p_i \right)$$

Rewrite $p_i \delta \dot{q}_i = \frac{d}{dt}(p_i \delta q_i) - \dot{p}_i \delta q_i$, and note that

$$\int_{t_1}^{t_2} \frac{d}{dt}(p_i \delta q_i) = p_i(t_2)\, \delta q_i(t_2) - p_i(t_1)\, \delta q_i(t_1)$$

which is zero if

$$\delta q_i(t_1) = 0 = \delta q_i(t_2) \tag{37}$$

(no restriction on δp_i!). The variations $(\delta q_i, \delta p_i)$ in paths are arbitrary, restricted only by (37). So we can conclude that the coefficients of δq_i and δp_i are zero. This gives

$$\dot{p}_i = -\frac{\partial H}{\partial q}, \quad \dot{q}_i = \frac{\partial H}{\partial p}$$

i.e. Hamilton's equations; i.e. Hamilton's equations are the "Euler-Lagrange" equations for this new variation principle.

This variational principle is the starting point for the treatment of what are called the generators of canonical transformations, this time finite not infinitesimal. We have no time to discuss this; suffice it to say that these generators are to this variational principle the analogies of the $F(t)$ of problem set 2, question 1.

12 The Hamilton-Jacobi Equation

I am afraid I have not had time to discuss this. Suffice it to say that it is yet another formulation of classical mechanics:

Lagrangian method:	n second-order ordinary differential equations
Hamiltonian method:	$2n$ first-order ordinary differential equations
Hamilton-Jacobi method:	1 partial differential equation in n variables (and possibly time)

The independent variable which is the subject of the Hamilton-Jacobi equation is a generator S of a particular canonical transformation (finite rather than infinitesimal, for which the original momenta p_i satisfy $p_i = \frac{\partial S}{\partial q_i}$. The (time-independent) equation is

$$H\left(\left\{\frac{\partial S}{\partial q_i}\right\}, \{q_i\}\right) = E$$

This is very close to the Schrödinger equation with $\psi \sim e^{iS/\hbar}$. In fact, Schrödinger's equation is exactly equivalent to the Hamilton-Jacobi equation with an extra term in it which vanishes in the limit $\hbar \to 0$, called the "quantum potential".

13 Envoy

A course of four weeks necessarily has limitations. Whilst exposing you to the different ideas, I am conscious that you have not had much chance to develop problem-solving techniques. You may remedy this deficiency for yourselves by trying your hand at more problems. Kibble's book, for example, has an interesting selection of problems, many of which have interesting physics connected with them.