Review: Vectors

A scalar, for the purpose of this course, is simply a real number

e.g. 1, 3, 8, $\frac{2}{3}$, $-\frac{1}{7}$, $\sqrt{2}$ are all scalars.

A vector, $\underline{r}$, is a collection of scalars, x, y, z:

e.g. $\underline{r} \equiv \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ is an example of a 3D vector.

A vector is a shorthand notation for a point in a vector space, e.g. $\underline{r}$ above encodes the 3D Cartesian coordinate (x, y, z). It has magnitude and direction in contrast to a scalar which has only magnitude, e.g. velocity is a vector quantity but speed is a scalar.

The components of a vector (x, y, z) are often indicated by an index notation. An index, i, is added to the vector notation, where i runs from one to the dimension of the vector space. In the 3D example above we may write the components of $\underline{r}$ as r^i where $i = 1, 2$ or 3.

e.g. $r^1 = x$, $r^2 = y$, $r^3 = z$.

N.B. compare this notation with that of matrices: M^{ab} which have two indices and scalars which have no indices.

Vectors "live" in a vector space. Vectors may be:

(i) multiplied by a scalar, a:

$$a\underline{r} \equiv a \begin{pmatrix} x \\ y \\ z \end{pmatrix} \equiv \begin{pmatrix} ax \\ ay \\ az \end{pmatrix}.$$

(ii) added to other vectors in the vector space:

$$\underline{r}_1 + \underline{r}_2 \equiv \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} \equiv \begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \\ z_1 + z_2 \end{pmatrix}$$

These properties may be combined to give linear addition, e.g.

$$a\underline{r}_1 + b\underline{r}_2 = \begin{pmatrix} ax_1 + bx_2 \\ ay_1 + by_2 \\ az_1 + bz_2 \end{pmatrix}.$$

The result of all these operations gives another vector in the vector space.

Example.

Three-dimensional Cartesian space can be defined as the vector space resulting from all linear combinations of:

$$\underline{e}_x \equiv \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \underline{e}_y \equiv \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \underline{e}_z \equiv \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

$\{\underline{e}_x, \underline{e}_y, \underline{e}_z\}$ are an example of a basis. In a D-dimensional vector space a basis consists of D linearly independent vectors. This means that if $\{\underline{e}_1, \underline{e}_2, \underline{e}_3 \ldots \underline{e}_D\}$ are a basis then one cannot find D scalars, $\{a_1, \ldots a_D\}$, such that:

$$a_1 \underline{e}_1 + a_2 \underline{e}_2 + \ldots a_D \underline{e}_D = 0.$$

If this were possible then one could express one of the basis vectors as a linear sum of the others

$$\text{e.g.} \quad \underline{e}_D = -1/a_D \left(a_1 \underline{e}_1 + a_2 \underline{e}_2 + \ldots a_{D-1} \underline{e}_{D-1} \right).$$

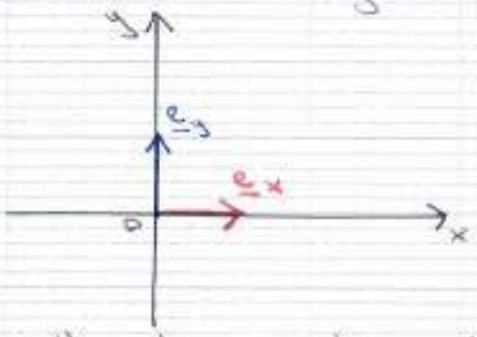
A degree of freedom would be lost and it would no longer possible to "span" a D-dimensional vector space. For example if this were true for 3D Cartesian space and one could write:

$$\underline{e}_z = 1/a_z \left(a_x \underline{e}_x + a_y \underline{e}_y \right)$$

Then by moving in the x and y directions one could achieve a z-direction motion! This is clearly impossible so if $\{\underline{e}_x, \underline{e}_y, \underline{e}_z\}$ were not linearly independent they would not be a basis of the 3D space.

A Simpler Example.

The Cartesian plane in two dimensions has a basis consisting of two vectors, which may be chosen as $\underline{e}_x \equiv \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\underline{e}_y \equiv \begin{pmatrix} 0 \\ 1 \end{pmatrix}$:



Every point in the plane can be reached from the origin, O, by a

linear combination of $\underline{e}_x$ and $\underline{e}_y$.

E.g. the point (x,y) may be reached from O by adding:

$$x\underline{e}_x + y\underline{e}_y = x\begin{pmatrix}1\\0\end{pmatrix} + y\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}x\\0\end{pmatrix} + \begin{pmatrix}0\\y\end{pmatrix} = \begin{pmatrix}x\\y\end{pmatrix} \equiv \underline{r}$$

In this simple example both $\underline{e}_x$ and $\underline{e}_y$ have unit length and so are called "normal" vectors.

They are also orthogonal to each other (at right-angles) and so are "orthogonal" vectors.

$\Rightarrow$ They form an "orthonormal" basis.

The Scalar Product (also called the dot product).

The scalar product of two vectors, $\underline{r}_1$ & $\underline{r}_2$, gives a scalar, $\underline{r}_1 \cdot \underline{r}_2$.

It satisfies two properties:

(i) $\underline{r}_1 \cdot \underline{r}_2 = \underline{r}_2 \cdot \underline{r}_1$ SYMMETRIC/COMMUTATIVE.

(ii) $\underline{r}_1 \cdot (\underline{r}_2 + \underline{r}_3) = \underline{r}_1 \cdot \underline{r}_2 + \underline{r}_1 \cdot \underline{r}_3$. DISTRIBUTIVE.

These properties imply that the scalar product is bilinear.

E.g. $(a\underline{r}_1) \cdot (b\underline{r}_2) = (ab)(\underline{r}_1 \cdot \underline{r}_2)$.

If, in addition, the scalar product is positive definite (which will be the case for this course) then:

$$\underline{r} \cdot \underline{r} \geq 0 \quad \left(\text{and } \underline{r} \cdot \underline{r} = 0 \text{ only when } \underline{r} = \underline{0} \equiv \begin{pmatrix}0\\0\\0\end{pmatrix}\right).$$

Example:

A positive definite scalar product in three dimensions is defined by:

$$\underline{r}_1 \cdot \underline{r}_2 = x_1x_2 + y_1y_2 + z_1z_2$$

Where $\underline{r}_1 \equiv \begin{pmatrix}x_1\\y_1\\z_1\end{pmatrix}$, $\underline{r}_2 \equiv \begin{pmatrix}x_2\\y_2\\z_2\end{pmatrix}$

Check that this is a scalar product (e.g. check (i.) & (ii.)).

E.g. if $\underline{r}_1 \equiv \begin{pmatrix}1\\1\\1\end{pmatrix}$ and $\underline{r}_2 = \begin{pmatrix}-1\\2\\3\end{pmatrix}$ then $\underline{r}_1 \cdot \underline{r}_2 = (1)(-1) + (1)(2) + (1)(3)$

$= -1 + 2 + 3$

$= 4.$ //

The length of a vector is defined in terms of the scalar product:

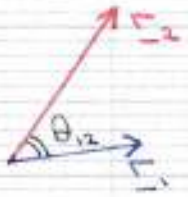
$$|\underline{r}| \equiv \sqrt{\underline{r} \cdot \underline{r}}$$

↖ Indicates length, or modulus, of $\underline{r}$.

There is a geometric interpretation of the dot product:

$$\underline{r}_1 \cdot \underline{r}_2 \equiv |\underline{r}_1||\underline{r}_2| \cos\theta_{12}.$$

Where θ_{12} is the angle between $\underline{r}_1$ and $\underline{r}_2$:



N.B. given two vectors we can find the angle between them using the scalar product.

Since $\cos(\pi/2) = 0$, orthogonal vectors, $\underline{r}_1$ and $\underline{r}_2$, have $\underline{r}_1 \cdot \underline{r}_2 = |\underline{r}_1||\underline{r}_2| \cos(\pi/2) = 0$. //

Recall the 3D orthonormal basis $\underline{e}_x = \begin{pmatrix}1\\0\\0\end{pmatrix}$, $\underline{e}_y = \begin{pmatrix}0\\1\\0\end{pmatrix}$, $\underline{e}_z = \begin{pmatrix}0\\0\\1\end{pmatrix}$.

These satisfy $\underline{e}_x \cdot \underline{e}_y = \underline{e}_x \cdot \underline{e}_z = \underline{e}_y \cdot \underline{e}_z = 0$. (CHECK.)

The Vector Product (also called the cross product).

The vector product of two vectors, $\underline{r}_1$ & $\underline{r}_2$, gives a vector, $\underline{r}_1 \wedge \underline{r}_2$, which is orthogonal to both $\underline{r}_1$ and $\underline{r}_2$.

i.e. $\underline{r}_1 \cdot (\underline{r}_1 \wedge \underline{r}_2) = 0$

and $\underline{r}_2 \cdot (\underline{r}_1 \wedge \underline{r}_2) = 0$.

The vector product satisfies:

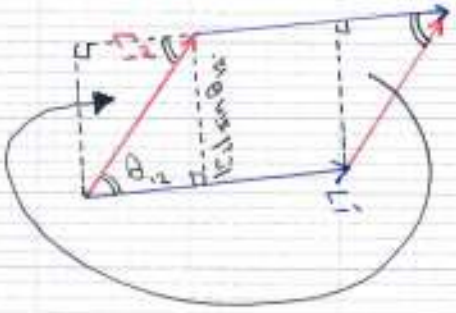
(i) $\underline{r}_1 \wedge \underline{r}_2 = -\underline{r}_2 \wedge \underline{r}_1$. ANTISYMMETRIC / ANTICOMMUTATIVE.

(ii) $\underline{r}_1 \wedge (\underline{r}_2 + \underline{r}_3) = \underline{r}_1 \wedge \underline{r}_2 + \underline{r}_1 \wedge \underline{r}_3$. DISTRIBUTIVE OVER ADDITION.

Geometrically,

$$|\underline{r}_1 \wedge \underline{r}_2| \equiv |\underline{r}_1||\underline{r}_2| \sin\theta_{12}.$$

In two dimensions this is the area of a parallelogram, two of whose edges are $\underline{r}_1$ and $\underline{r}_2$:

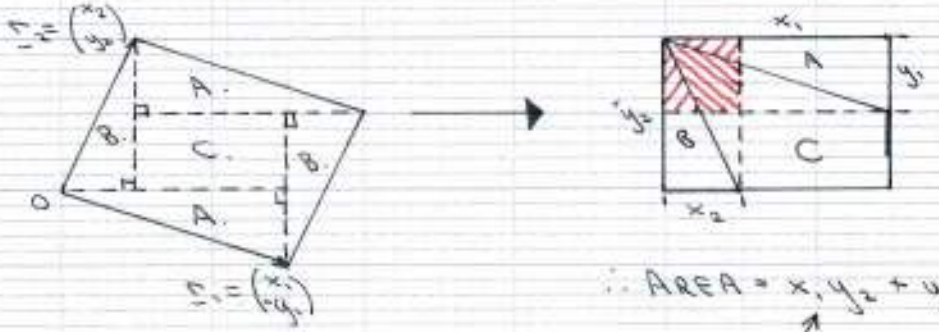


Area $= |\underline{r}_1||\underline{r}_2| \sin\theta_{12}$.

Recall that in two dimensions the area of a parallelogram is also given by the determinants of a 2-by-2 matrix whose columns define the edges. In our case the matrix columns would be $\underline{r}_1$ and $\underline{r}_2$:

$$\det \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix} = x_1 y_2 - y_1 x_2 .$$

This is because one can reorganise the parts of a parallelogram into a rectangle with an internal rectangle undercounted as so:



$$\therefore \text{AREA} = x_1 y_2 + y_1 x_2 .$$

Large rectangle. Overlapping part.

However in general $\underline{r}_1 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$ having positive components this changes our area formula as $-y_1 \to y_1$ $\therefore$ AREA $= x_1 y_2 - y_1 x_2$. This is really a matter of sign conventions. We can write the length of the crossproduct in terms of the vector components:

$$|\underline{r}_1 \wedge \underline{r}_2| = \left| \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \wedge \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \right| = \left(x_1 y_2 - y_1 x_2 \right).$$

Using this we can find an expression for the 3D cross product in terms of components.

We know that $\underline{r}_1 \wedge \underline{r}_2$ is a vector orthogonal to both $\underline{r}_1$ and $\underline{r}_2$ with length $x_1 y_2 - y_1 x_2$. If $\underline{r}_1$ and $\underline{r}_2$ are not parallel we must move to working in three dimensions - this means adding another row to our vectors:

$$\underline{r}_1 = \begin{pmatrix} x_1 \\ y_1 \\ 0 \end{pmatrix}, \quad \underline{r}_2 = \begin{pmatrix} x_2 \\ y_2 \\ 0 \end{pmatrix}.$$

The third row corresponds to the direction orthogonal to both $\underline{r}_1$ and $\underline{r}_2$. Now a vector in this direction with length $x_1 y_2 - y_1 x_2$ is:

$$\begin{pmatrix} x_1 \\ y_1 \\ 0 \end{pmatrix} \wedge \begin{pmatrix} x_2 \\ y_2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \pm (x_1 y_2 - y_1 x_2) \end{pmatrix}.$$

We have yet to determine the plus or minus sign.

Insert: Alternative area of a parallelogram derivation.

Consider a parallelogram defined by the two vectors $\underline{r}_1 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$ and $\underline{r}_2 = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$:



The area is: $|\underline{r}_1 \wedge \underline{r}_2| = |\underline{r}_1||\underline{r}_2| \sin\theta$.

= Base × Height.

To find an expression for the area in terms of components we can use the scalar product:

$$\underline{r}_1 \cdot \underline{r}_2 = |\underline{r}_1||\underline{r}_2| \cos\theta.$$

$$= x_1 x_2 + y_1 y_2.$$

We can use the identity:

$$\sin^2\theta + \cos^2\theta = 1$$

To find $|\underline{r}_1||\underline{r}_2| \cos\theta$:

$$|\underline{r}_1|^2 |\underline{r}_2|^2 \cos^2\theta = |\underline{r}_1|^2 |\underline{r}_2|^2 - |\underline{r}_1|^2 |\underline{r}_2|^2 \sin^2\theta.$$

$$\| \quad (x_1 x_2 + y_1 y_2)^2$$

N.B. $|\underline{r}_1|^2 = x_1^2 + y_1^2$, $|\underline{r}_2|^2 = x_2^2 + y_2^2$.

Giving,

$$|\underline{r}_1|^2 |\underline{r}_2|^2 \sin^2\theta = (x_1^2 + y_1^2)(x_2^2 + y_2^2) - (x_1 x_2 + y_1 y_2)^2$$

$$= x_1^2 x_2^2 + x_1^2 y_2^2 + y_1^2 x_2^2 + y_1^2 y_2^2$$
$$- x_1^2 x_2^2 - 2 x_1 x_2 y_1 y_2 - y_1^2 y_2^2.$$

$$= (x_1 y_2 - y_1 x_2)^2.$$

$$\therefore \quad |\underline{r}_1 \wedge \underline{r}_2| = x_1 y_2 - y_1 x_2. \quad //$$

However there are two other ways we could have added zeroes to $\underline{r}_1$ and $\underline{r}_2$ to extend into the third direction. We would have written:

$$\begin{pmatrix} x_1 \\ 0 \\ z_1 \end{pmatrix} \wedge \begin{pmatrix} x_2 \\ 0 \\ z_2 \end{pmatrix} = \pm \begin{pmatrix} 0 \\ x_1 z_2 - x_2 z_1 \\ 0 \end{pmatrix}$$

and $$\begin{pmatrix} 0 \\ y_1 \\ z_1 \end{pmatrix} \wedge \begin{pmatrix} 0 \\ y_2 \\ z_2 \end{pmatrix} = \pm \begin{pmatrix} y_1 z_2 - y_2 z_1 \\ 0 \\ 0 \end{pmatrix}.$$

Putting these three together, we have:

$$\underline{r}_1 \wedge \underline{r}_2 = \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} \wedge \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} \equiv \begin{pmatrix} \pm(y_1 z_2 - y_2 z_1) \\ \pm(x_1 z_2 - x_2 z_1) \\ \pm(x_1 y_2 - x_2 y_1) \end{pmatrix} \equiv \underline{r}_3.$$

What about the +/− signs? Well we know the result must be orthogonal to both $\underline{r}_1$ and $\underline{r}_2$ - we can check this with the dot product i.e. we need that $\underline{r}_1 \cdot \underline{r}_3 = \underline{r}_2 \cdot \underline{r}_3 = 0$.

So, $$\underline{r}_1 \cdot (\underline{r}_1 \wedge \underline{r}_2) = x_1\left(\pm\underline{(y_1 z_2 - y_2 z_1)}\right) + y_2\left(\pm\underline{(x_1 z_2 - x_2 z_1)}\right) + z_3\left(\pm\underline{(x_1 y_2 - x_2 y_1)}\right).$$

$$= 0.$$

For this to be zero the terms underlined in red must have the opposite sign to the one underlined in blue.

We would find the same condition from checking $\underline{r}_2 \cdot (\underline{r}_1 \wedge \underline{r}_2)$.

Finally we pick a sign convention: the red terms are +ve and the blue terms −ve. This corresponds to defining positive and negative areas. Finally we have the usual definition:

$$\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} \wedge \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} \equiv \begin{pmatrix} y_1 z_2 - y_2 z_1 \\ -(x_1 z_2 - x_2 z_1) \\ x_1 y_2 - x_2 y_1 \end{pmatrix}.$$

Note that since $|\underline{r}_1 \wedge \underline{r}_2| = |\underline{r}_1||\underline{r}_2| \sin\theta_{12}$, if $\underline{r}_1$ and $\underline{r}_2$ are parallel ($\theta = 0$) then $|\underline{r}_1 \wedge \underline{r}_2| = 0 \Rightarrow \underline{r}_1 \wedge \underline{r}_2 = \underline{0}$ if $\underline{r}_1$ and $\underline{r}_2$ are parallel.

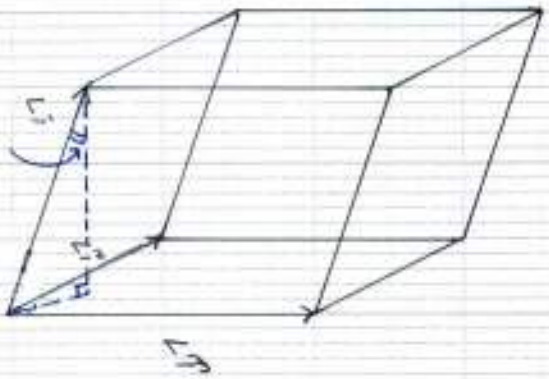
Triple Products.

A. The Scalar Triple Product.

The scalar triple product takes three vectors and gives a scalar: $\underline{r}_1 \cdot (\underline{r}_2 \wedge \underline{r}_3)$

It is the volume of a parallelapiped whose defining edges are $\underline{r}_1$, $\underline{r}_2$ and $\underline{r}_3$.

Parallelapiped:



VOLUME = AREA OF BASE × HEIGHT.

AREA OF BASE = Area of the parallelogram defined by $\underline{r}_2$ and $\underline{r}_3$.

$$= |\underline{r}_2 \wedge \underline{r}_3|$$

$$= |\underline{r}_2||\underline{r}_3| \sin\theta_{23}.$$

HEIGHT $= |\underline{r}_1| \cos\theta_{1,213}$.

(N.B. $\theta_{1,213}$ is the angle between $\underline{r}_1$ and the perpendicular to $\underline{r}_2$ and $\underline{r}_3$, i.e. $\underline{r}_2 \wedge \underline{r}_3$).

$\therefore$ Volume $= |\underline{r}_1||\underline{r}_2||\underline{r}_3| \sin\theta_{23} \cos\theta_{1,213}$.

Now consider,

$$\underline{r}_1 \cdot (\underline{r}_2 \wedge \underline{r}_3) = |\underline{r}_1||\underline{r}_2 \wedge \underline{r}_3| \cos\theta_{1,213}$$

$$= |\underline{r}_1||\underline{r}_2||\underline{r}_3| \sin\theta_{23} \cos\theta_{1,213}. \;//$$

$$= \text{Volume of a parallelapiped.}$$

Since the scalar triple product gives a volume then we can cyclically change the order of the vectors to get the same result:

i.e. $\underline{r}_1 \cdot (\underline{r}_2 \wedge \underline{r}_3) = \underline{r}_2 \cdot (\underline{r}_3 \wedge \underline{r}_1) = \underline{r}_3 \cdot (\underline{r}_1 \wedge \underline{r}_2)$.

N.B. If we changed the order of the vectors we would find minus the volume. The scalar triple product gives the determinant

of the 3×3 matrix whose column vectors are $\underline{r}_1, \underline{r}_2$ and $\underline{r}_3$.

B. The Vector Triple Product

The vector triple product is defined as the cross product of one vector with the cross product of another two vectors:

$$\underline{r}_1 \wedge (\underline{r}_2 \wedge \underline{r}_3)$$

It satisfies the identity:

$$\underline{r}_1 \wedge (\underline{r}_2 \wedge \underline{r}_3) = (\underline{r}_1 \cdot \underline{r}_3)\,\underline{r}_2 - (\underline{r}_1 \cdot \underline{r}_2)\,\underline{r}_3.$$

This is called the triple product expansion.

Elementary Mechanics.

Newton's three laws of motion are:

NI./ A particle will remain at rest or move at constant velocity in a straight line unless it is acted on by an external force.

NII./ In an inertial reference frame, the rate of change of momentum of a particle is equal to, and in the direction of, the net force acting on it.

NIII./ Every action has an equal and opposite reaction.

Comments.

NI gives the definition of an 'inertial reference frame': an inertial reference frame is one in which the first Law holds true.*

As an example of a reference frame that is not inertial consider a passenger in a stationary train carriage which accelerates when the train departs the station. The passenger may observe the coffee trolley to begin to roll backwards along the train without observing an external force acting, in violation of NI. The carriage is not an inertial reference frame because it accelerates.

As an example of an inertial frame consider a lecture hall in King's College or, more excitingly, a free-falling elevator full of second year undergraduates falling at constant velocity due to air resistance. If you roll a marble in idealised versions of these environments it will move in a straight line at constant velocity (friction neglected).

Given an inertial frame there are an infinite number of other inertial frames related to it by 10 independent transformations:

* In special relativity an inertial reference frame is one in which all the physical laws are valid not just NI.

- 3 space translations and 1 time.
- 3 boosts: $x \to x + u_x t$, $y \to y + u_y t$, $z \to z + u_z t$.
- 3 rotations.

The validity of NI does not change if we make any of these transformations, which together make up the Galilean group. Indeed since NI, NII and NIII are valid in inertial frames we may say that Newton's laws of motion are invariant under the Galilean group.

- NIII is the statement that if a body A exerts a force on body B then body B exerts an equal force in the opposite direction on A.
- NII is more familiarly expressed as $\underline{F} = m\underline{a}$:

  NII states that $\underline{F} = \frac{d}{dt}(\underline{p})$ where $\underline{p} \equiv m\underline{v} \equiv m\dot{\underline{r}}$ is momentum.

$$\therefore \underline{F} = \frac{d}{dt}(m\dot{\underline{r}})$$
$$= \dot{m}\dot{\underline{r}} + m\ddot{\underline{r}}$$

In most cases the mass, m, is constant over the motion of a body but in some important cases it is not (e.g. rocket science or the motion of a snail) so we must generally use $\underline{F} = \dot{\underline{p}}$. But when m is constant then $\dot{m} = 0$ and we have:

$$\underline{F} = m\ddot{\underline{r}} \equiv m\underline{a}.$$

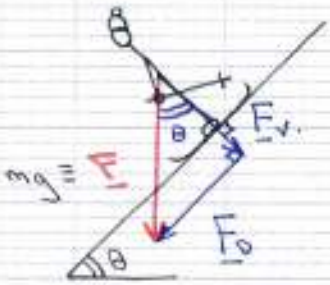
The force is a function of position, $\underline{r}$, and velocity, $\dot{\underline{r}}$. If we are given initial value data at t_0 for $\underline{r}$ and $\dot{\underline{r}}$ we can solve this differential equation for $\underline{r}$ at all times! This is the goal of classical dynamics.

Familiar Examples: Solving $\underline{F} = m\ddot{\underline{r}}$ (A reminder.)

The Skier with no friction.

Consider a skier descending a slope inclined at an angle of θ to the horizontal who is subject to no forces other than the force of gravity. Let's use Newton's laws to discover the skier's equation of motion.

We commence with a diagram indicating the forces on the skier:



First we observe that the skier remains on the surface of the mountain, that is despite the force of gravity, $\underline{F}$, acting in a direction that pulls the skier vertically downwards and through the mountain the skier does not move in that direction. We are able to make this observation due to our innate knowledge that skiers and mountains are held together by forces stronger than the gravitational pull of the Earth (the electromagnetic force is far stronger than gravity). The component of the gravitational force acting perpendicular to the mountain on the skier is cancelled out by NIII: there is an equal but opposite force $(-\underline{F}_v)$ acting at the point of contact between the skier and the mountain. There is only one unbalanced force acting on the skier, the component of $\underline{F}$ down the mountain, $\underline{F}_0$. The system has become one-dimensional - we let r denote the distance down the mountain. By NII we write the equation of motion in the r-direction:

$$|\underline{F}_0| = mg\sin\theta = m\ddot{r}$$

$$\Rightarrow \quad \ddot{r} = g\sin\theta$$

This is an inhomogeneous, second-order differential equation. It is inhomogeneous because the differential terms in $r(\ddot{r}, \dot{r}, r)$ when collected together are not equal to zero. The term on the right is a constant, $g\sin\theta$. In this example we simply integrate with respect to time twice to find $r(t)$:

$$\dot{r} = tg\sin\theta + A.$$

$$\Rightarrow \quad r = \tfrac{1}{2}t^2 g\sin\theta + At + B.$$

A, B are the two constants of integration which must be fixed by the initial conditions.

Suppose the skier has speed $2\,ms^{-1}$ at $100\,m$ down the slope what are A and B?

We may use t to parameterise $r(t)$ as we prefer: let us choose $t=0$ to correspond to $r = 100\,m$. Then:

$$r(0) = 100 = B.$$

By substituting $\dot{r}(0) = 2$ we can find A:

$$\dot{r}(0) = 2 = A.$$

We have,

$$r(t) = \tfrac{1}{2}t^2 g\sin\theta + 2t + 100. \;//$$

We have found the position of the skier for all time t, this is the aim of classical mechanics. However in other situations the differential equations may not be so easy to solve and we may prefer to use our understanding of constants of the motion to find the dynamics of a system.

For linear dynamics when there are no forces acting the linear momentum is a conserved quantity:

$$\underline{F} = \frac{d}{dt}(\underline{p})$$

If $\underline{F} = 0$ then the linear momentum is conserved over the motion.

Angular Momentum.

The angular momentum of a particle about O is defined as:

$$\underline{L} \equiv \underline{r} \wedge \underline{p}.$$

i.e. $\underline{L}$ is a vector orthogonal to both $\underline{r}$ and $\underline{p}$ and with length $|\underline{L}| = |\underline{r}||\underline{p}| \sin\theta_{\underline{r},\underline{p}}$.

e.g. Consider a planet orbiting a star; which we put at the origin, O:



If the planet orbits the star in a plane, its angular momentum vector lies perpendicular to that plane and passes through the origin: it appears to be the axis of rotation.

The moment of force, or torque, is defined to be

$$\underline{N} \equiv \underline{r} \wedge \underline{F}.$$

Notice that both angular momentum and torque are defined in terms of the cross product and will have a different sign if the order of the vectors is reversed: so be careful!

How does torque relate to angular momentum?

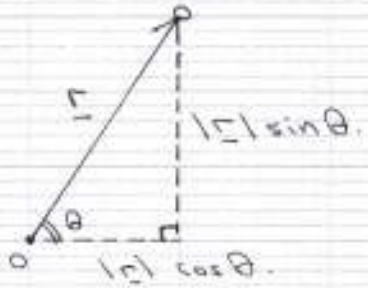
Recall that by NII, $\underline{F} = \dot{\underline{p}}$ and that $\underline{p} = m\dot{\underline{r}}$

$$\therefore \underline{N} = \underline{r} \wedge \underline{F}$$
$$= \underline{r} \wedge \dot{\underline{p}}$$
$$= \frac{d}{dt}(\underline{r} \wedge \underline{p}) \quad \left[\text{as } \frac{d}{dt}(\underline{r} \wedge \underline{p}) = \dot{\underline{r}} \wedge \underline{p} + \underline{r} \wedge \dot{\underline{p}}\right.$$
$$= \dot{\underline{L}}. \qquad = \dot{\underline{r}} \wedge m\dot{\underline{r}} + \underline{r} \wedge \dot{\underline{p}}$$
$$= \underline{0} + \underline{r} \wedge \dot{\underline{p}}. \left.\right]$$

This is NII written in terms of the angular quantities $\underline{N}$ and $\underline{L}$.

Let us now consider the example of uniform circular motion in order to see the role played by these quantities and compare them with their linear counterparts. Our aim is to see angular momentum, torque and the moment of inertia ($I = mr^2$) as angular analogues of momentum, force and mass.

Example: Constant Circular Motion.



Let's rewrite $|\underline{r}|$ as r, $|\underline{r}| \equiv r$.

Then in the Cartesian plane,

$$\underline{r} = \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix}$$

For circular motion r is a constant so $\frac{d}{dt}|\underline{r}| = 0$

N.B. $\frac{d}{dt}|\underline{r}| = 0 \not\Rightarrow |\dot{\underline{r}}| = 0$.

The linear momentum is $\underline{p} = m\dot{\underline{r}}$.

What is $\dot{\underline{r}}$?

$$\dot{\underline{r}} = \begin{pmatrix} -r\sin\theta\,\dot{\theta} \\ r\cos\theta\,\dot{\theta} \end{pmatrix}$$

Aside: $\underline{p}$ is perpendicular to $\underline{r}$;

as $\underline{p}\cdot\underline{r} = m\dot{\theta}(r^2\cos\theta\sin\theta - r^2\sin\theta\cos\theta)$

$= 0$. //

N.B. when $\theta = 0$, $\underline{r} \equiv \begin{pmatrix} r \\ 0 \end{pmatrix}$ and $\dot{\underline{r}} \equiv \begin{pmatrix} 0 \\ r\dot{\theta} \end{pmatrix}$

notice that $|\dot{\underline{r}}|_{\theta=0} = r\dot{\theta}$.

Aside II: $\dot{\theta} \equiv 2\pi f$ where f is the frequency; or cycles per second, of the motion.

For constant circular motion $\dot{\theta}$ is a constant and so is $|\underline{r}| \equiv r$ - does this mean $|\dot{\underline{r}}| = r\dot{\theta}$ over the whole motion i.e. all values of θ?

In fact $|\dot{\underline{r}}|^2 = r^2\sin^2\theta\,\dot{\theta}^2 + r^2\cos^2\theta\,\dot{\theta}^2$

$= r^2\dot{\theta}^2(\sin^2\theta + \cos^2\theta)$

$= r^2\dot{\theta}^2$

$\Rightarrow |\dot{\underline{r}}| = r\dot{\theta}$.

Now notice that even for a system moving with constant angular speed there is still an acceleration:

$$|\ddot{\underline{r}}|^2 = \left|\begin{pmatrix} -r\cos\theta\,\dot{\theta}^2 \\ -r\sin\theta\,\dot{\theta}^2 \end{pmatrix}\right|^2$$

$= r^2\dot{\theta}^4$

$\Rightarrow |\ddot{\underline{r}}| = r\dot{\theta}^2$.

Note that $\underline{r} = -\ddot{\underline{r}}$ the acceleration points towards the origin.

There is an associated force inherent in a system moving with constant angular speed, namely:

$$\underline{F} = m\ddot{\underline{r}} = -m\underline{r}.$$

Its magnitude is:

$$|\underline{F}| = mr\dot{\theta}^2 = m\dot{r}^2/r.$$

This is a centripetal force it is not associated to an angular acceleration (i.e. $\ddot{\theta} \neq 0$).

Angular momentum defines the torque which is associated to an angular acceleration.

Recall, angular momentum, $\underline{L} = \underline{r} \wedge \underline{p}$

Here, $\underline{r} = \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix}$, $\underline{p} = m\dot{\underline{r}} = m\begin{pmatrix} -r\sin\theta\,\dot{\theta} \\ r\cos\theta\,\dot{\theta} \end{pmatrix}$

(Adding a third dimension in the z-direction)

$$\underline{L} = \underline{r} \wedge \underline{p} = m\begin{pmatrix} r\cos\theta \\ r\sin\theta \\ 0 \end{pmatrix} \wedge \begin{pmatrix} -r\sin\theta\,\dot{\theta} \\ r\cos\theta\,\dot{\theta} \\ 0 \end{pmatrix}$$

$$= m\begin{pmatrix} 0 \\ 0 \\ r^2\cos^2\theta\,\dot{\theta} + r^2\sin^2\theta\,\dot{\theta} \end{pmatrix}$$

$$= m\begin{pmatrix} 0 \\ 0 \\ r^2\dot{\theta} \end{pmatrix}$$

Now $|\underline{L}| = mr^2\dot{\theta} = (mr)r\dot{\theta} = (mr)|\dot{\underline{r}}| = \underbrace{(mr^2)}\,\dot{\theta} = I\dot{\theta}$

MOMENT OF INERTIA (The equivalent of mass in a rotating system, denoted by I).

And torque, $\underline{N} = \underline{r} \wedge \underline{F}$.

For circular motion we have $\underline{F} = m\ddot{\underline{r}} = -m\underline{r}$

$$\therefore \underline{N} = -\underline{r} \wedge m\underline{r}$$

$$= 0. \;//$$

This is what we wanted we now have an angular analogue of Newton's second law (NII):

Linear motion.	Constant $\underline{p}$	$\Leftrightarrow$ no Force, $\underline{F}$.	$\underline{F} = m\underline{a}$.
Angular motion.	constant $\underline{L}$	$\Leftrightarrow$ no torque, $\underline{N}$.	$\lvert\underline{N}\rvert = I\ddot{\theta}$.

Recall in general $\underline{N} = \dot{\underline{L}}$ and here $|\dot{\underline{L}}| = mr^2\ddot{\theta}\,(=0) = I\ddot{\theta}$

Conserved Quantities.

We have already seen that linear momentum is conserved when there are no forces acting on a particle (by NII). We may wonder what the equivalent conserved quantity is for a rotating particle. It certainly isn't linear momentum which varies cyclically about a circular orbit. Angular momentum is conserved when there is no torque. Recall the angular version of NII:

$$\underline{N} = \dot{\underline{L}}$$

If the torque, $\underline{N}$, is zero then the angular momentum, $\underline{L}$, is a conserved quantity.

Let us introduce the concept of work*. When a force is applied to a body the work, W, is equal to the product of the force and the distance moved in the direction of the force. If the position of a particle is given by $\underline{r}(t)$ and a force is applied then the work done while the particle moves $\Delta\underline{r}$ is:

$$\Delta W = |\underline{F}||\Delta\underline{r}|\cos\theta \equiv \underline{F}\cdot\Delta\underline{r}$$

Where θ is the angle between $\Delta\underline{r}$ and $\underline{F}$: $|\Delta\underline{r}|\cos\theta$ is then the distance moved in the direction of the force. If the particle moves from $\underline{r}(t_1)$ to $\underline{r}(t_2)$ then the total work over the motion is the line integral:

$$W_{12} = \int_{t_1}^{t_2} \underline{F}\cdot d\underline{r} \qquad [1.]$$

In terms of components, writing: $\underline{F} = \begin{pmatrix} F_x \\ F_y \end{pmatrix}$, $\underline{r} = \begin{pmatrix} x \\ y \end{pmatrix}$ for a 2D system we have

$$W_{12} = \int_{t_1}^{t_2} (dx\,F_x + dy\,F_y)$$

The line integral is a compact notation for a sum of "ordinary", scalar integrals.

Let us use NII to rewrite the integral for work ([1.]) by substituting $\underline{F} = \dot{\underline{p}}$ into equation [1.]. Let us assume that the

* As a second year undergraduate you will already be familiar with work!

mass, m, is constant (so that $\underline{F} = m\ddot{\underline{r}}$). We have:

$$W = \int \underline{F} \cdot d\underline{r}$$

$$= m \int \ddot{\underline{r}} \cdot \dot{\underline{r}}\, dt. \qquad \left(\text{N.B. } d\underline{r} = \left(\frac{d\underline{r}}{dt}\right) dt = \dot{\underline{r}}\, dt\right).$$

$$= m \int \frac{d}{dt}\left(\frac{1}{2}\dot{\underline{r}}^2\right) dt. \qquad \left(\text{N.B. } \frac{d}{dt}\left(\tfrac{1}{2}\dot{\underline{r}}^2\right) = \tfrac{1}{2}\frac{d}{dt}\left(\dot{\underline{r}} \cdot \dot{\underline{r}}\right)\right.$$

$$= \tfrac{1}{2}\left(\ddot{\underline{r}} \cdot \dot{\underline{r}} + \dot{\underline{r}} \cdot \ddot{\underline{r}}\right)$$

$$= \ddot{\underline{r}} \cdot \dot{\underline{r}}\Big).$$

$$= \frac{1}{2} m \int \frac{d}{dt}\left(\dot{\underline{r}}^2\right) dt. \quad //$$

Substituting the limits t_1 and t_2 to find the work done W_{12} as the particle moves from $\underline{r}(t_1)$ to $\underline{r}(t_2)$ we have:

$$W_{12} = \frac{1}{2} m \left[\dot{\underline{r}}^2\right]_{t_1}^{t_2}$$

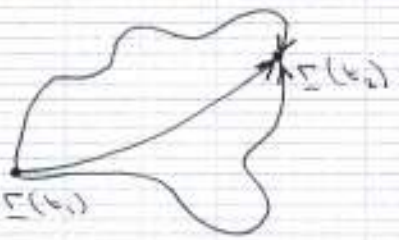
$$= \frac{1}{2} m \left(\dot{\underline{r}}^2(t_2) - \dot{\underline{r}}^2(t_1)\right) \quad //$$

Where $\dot{\underline{r}}^2 = |\dot{\underline{r}}|^2$ is a scalar. The quantity $\frac{1}{2} m \dot{\underline{r}}^2$ is called the kinetic energy of the particle, and is frequently denoted by a "T".

The work done as the system moves from time 1, t_1, to time 2, t_2, is equal to the change in kinetic energy:

$$\boxed{W_{12} = T_2 - T_1 \equiv \Delta T.}$$

The work depends on the path taken between $\underline{r}(t_1)$ and $\underline{r}(t_2)$:



However whenever the work done is independent of the path taken, that is it depends only on the start $\underline{r}(t_1)$ and end $\underline{r}(t_2)$ points, then the force causing the work is conservative.

If the work done moving from $\underline{r}(t_1)$ to $\underline{r}(t_2)$ is conservative i.e. the work depends only on the endpoints then the reverse path from $\underline{r}(t_2)$ to $\underline{r}(t_1)$ also depends only on the endpoints and is the negative of W_{12}.

i.e. $W_{12} = T_2 - T_1$

$\Rightarrow W_{21} = T_1 - T_2 = -W_{12}$.

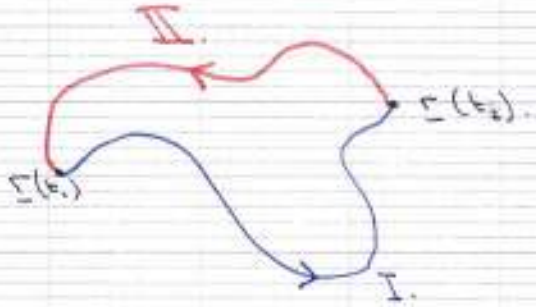
Consequently the work done moving around a closed circuit under the influence of a conservative force is zero, independent of the path taken, that is:

$$\oint \underline{F} \cdot d\underline{r} = 0.$$

$\oint$ Indicates a closed path.

e.g. $\oint \underline{F} \cdot d\underline{r} = \int_{t_1}^{t_2} \underline{F} \cdot d\underline{r} + \int_{t_2}^{t_1} \underline{F} \cdot d\underline{r}$.

Path I. Path II.



Since $\oint \underline{F} \cdot d\underline{r}$ is a sum of work integrals along a closed circuit then if $\oint \underline{F} \cdot d\underline{r} = 0$ (and the force is conservative) then it must be composed of both positive and negative terms. Consequently any force that gives only positive integrals $\int \underline{F} \cdot d\underline{r}$ around a closed path cannot give a vanishing $\oint \underline{F} \cdot d\underline{r}$ and so is not a conservative force. For example a force due to friction always gives a positive contribution to $\oint \underline{F} \cdot d\underline{r}$ so is not a conservative force. Consider pushing a filing cabinet around an office at every point in the path one has to do work against the friction: the friction never pulls the filing cabinet and certainly not because cabinet is on the return leg of its journey. Some forces do have this property, gravity is a conservative force: if one cycles down one side of a hill before returning to the top, neglecting all forces other than gravity, the work around the closed path is zero. Gravity is a central force and, as we shall see, all central forces are also conservative.

A necessary and sufficient condition for a force to be conservative is that it can be expressed as the gradient of some scalar function which depends only on position, $V(\underline{r})$. It is convenient to normalise the force to be in the opposite direction to the gradient:

$$\underline{F} = -\underline{\nabla} V(\underline{r}). \qquad [2]$$

$V(\underline{r})$ is the potential, or potential energy. Let us prove that our notions of path-independent work and that the force may be expressed as the gradient of the potential are equivalent.

Suppose $\underline{F}$ may be written as in equation [2] then,

$$\int_{\underline{r}_1}^{\underline{r}_2} \underline{F} \cdot d\underline{r} = -\int_{\underline{r}_1}^{\underline{r}_2} \underline{\nabla} V \cdot d\underline{r}$$

In three dimensions,

$$\underline{\nabla} V = \begin{pmatrix} \partial V/\partial x \\ \partial V/\partial y \\ \partial V/\partial z \end{pmatrix}$$

$$d\underline{r} = \begin{pmatrix} \partial x/\partial t \\ \partial y/\partial t \\ \partial z/\partial t \end{pmatrix} dt.$$

$$\therefore \int_{t_1}^{t_2} \underline{F} \cdot d\underline{r} = -\int_{t_1}^{t_2} \left(\frac{\partial V}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial V}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial V}{\partial z}\frac{\partial z}{\partial t} \right) dt.$$

$$= -\int_{t_1}^{t_2} dV.$$

$$= -\left[V(\underline{r}) \right]_{t_1}^{t_2}$$

$$= -V(\underline{r}(t_2)) + V(\underline{r}(t_1))$$

$$W_{12} = V_1 - V_2.$$

That is, assuming [2] the work becomes a function of the beginning and endpoints of the path.

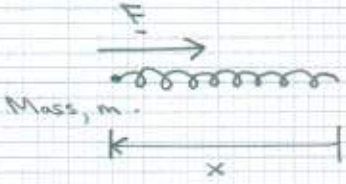
For a conservative force we now have two expressions for W_{12} namely: $W_{12} = T_2 - T_1 = V_1 - V_2$.

Rearranging we have:

$$\boxed{T_1 + V_1 = T_2 + V_2.}$$

Examples: Conservation of Energy.

Example 1: The Simple Harmonic Oscillator.



A mass m is subject to a force $F = -kx$. This is a one-dimensional problem, the problem of whether $\underline{F} = -\underline{\nabla} V$ for some $V(x,y,z)$ becomes whether $F = -\frac{\partial}{\partial x} V(x)$ for some $V(x)$. Clearly $V(x) = \frac{1}{2}kx^2$ will work.

(A constant.)

Check: $-\frac{\partial}{\partial x}\left(\frac{1}{2}kx^2\right) = -kx \equiv F$. //

As $F \neq 0$ linear momentum is not conserved. However since $F = -\frac{\partial}{\partial x}(V)$ then F is conservative and energy, $E\ (=T+V)$, is conserved for this motion:

$$E = T + V$$
$$= \frac{1}{2}m\dot{x}^2 + \frac{1}{2}kx^2 \quad //$$

Example 2: The Simple Harmonic Oscillator with Friction.

Let us amend the F of example 1 to include a force due to friction: $F = -kx - A\dot{x}$ k, A are constants.

Recall that friction terms give rise to a path-dependent work function: the longer the path, the more work done against friction. The force is not conservative $\Rightarrow$ Energy is not a constant during the motion. Of course with friction the harmonic oscillator lose energy and slows down to an eventual standstill.

Example 3: A Central, Conservative Force – Gravity.

Gravitational force, $\underline{F} = \left(-G.M_1M_2/r^2\right)\underline{r}/|\underline{r}| = -GM_1M_2/r^3\ \underline{r}$.

Is derived from a potential, $V = -GM_1M_2/r$.

As $-\underline{\nabla} V = \underline{F}$ the force is conservative. $\underline{F}$ is also central $(\underline{F} \propto -\underline{r})$

$\therefore$ Torque $\underline{N} = \underline{r} \wedge \underline{F} = 0 \Rightarrow$ (by $\underline{N} = \dot{\underline{L}}$) that angular momentum, $\underline{L}$, is conserved.

This is the energy conservation theorem:

> If the forces acting on a particle are conservative then $T+V$, the total energy, is conserved.

Hence we see why conservative forces are called conservative. The total energy is an example of a conserved quantity in a dynamical system. Often knowledge of the conserved quantities in a system offer a simple way of evaluating properties of the system which is much easier than solving the 2nd order differential equations contained in $\underline{F} = m\ddot{\underline{r}}$.

Many-Particle Systems

So far we have considered only the dynamics of single particle systems. In this section we will generalise the concepts we have defined to multiparticle systems. One important difference is that because the system will be composed of many parts it will be able to exert forces and torques on itself. Such forces arising from interactions of the system of particles with itself will be called internal forces. The forces we have been considering so far will be called external forces, since they originate from a source external to the system. Consider two oppositely charged particles moving in a magnetic field. The force of attraction between the particles is an internal force, while the force on the particles due to their motion in a magnetic field is an external force.

Consider a system of n particles. We will denote the external force on particle i $(i=1 \cdots n)$ by $\underline{F}_i^{ext}$ and the internal force on particle i due to an interaction with particle j by $\underline{F}_{ij}^{int}$. By NII the total force on particle i will be equal to the rate of change of its momentum, $\dot{\underline{p}}_i$.

$$\therefore \underline{F}_i^{ext} + \sum_{j=1}^{n} \underline{F}_{ij}^{int} = \dot{\underline{p}}_i$$

[3.]

Where the self-interaction of i with itself is zero ($\underline{F}^{int}_{ii} = 0$).
It will be assumed that the internal forces obey NIII,

i.e. $\underline{F}^{int}_{ij} = -\underline{F}^{int}_{ji}$.

We are interested in the dynamics of the full system of particles not just one particle within a background of other particles. Consequently we must sum the internal and external forces over all the particles. We sum [3.] over i:

$$\sum_{ij} \underline{F}^{int}_{ij} + \sum_{i} \underline{F}^{ext}_{i} = \sum_{i} \dot{\underline{p}}_i$$

$$= \sum_{i} m_i \ddot{\underline{r}}_i$$

Where we have assumed that while the masses of the particles may not be identical they are all constant. We have:

$$\sum_{ij} \underline{F}^{int}_{ij} + \sum_{i} \underline{F}^{ext}_{i} = \sum_{i} m_i \ddot{\underline{r}}_i$$

Notice that $\sum_{ij} \underline{F}^{int}_{ij} = 0$ as for every $\underline{F}^{int}_{ij}$ there is an $\underline{F}^{int}_{ji}$ in the sum which cancels it out. Also $\underline{F}^{int}_{ii} = 0$.

The equations simplify if instead of $\underline{r}_i$ we use a vector $\underline{R}$ called the centre of mass:

$$\underline{R} \equiv \sum_i m_i \underline{r}_i \Big/ \sum_i m_i = \sum_i m_i \underline{r}_i \Big/ M.$$

Where $M = \sum m_i$ is the total mass of the system. $\underline{R}$ is a weighted average of the $\underline{r}_i$ and the weighting is in proportion to the particle's mass.

Now, $\dot{\underline{R}} = \sum m_i \dot{\underline{r}}_i / M$

and $\ddot{\underline{R}} = \sum m_i \ddot{\underline{r}}_i / M.$

$$\Rightarrow M\ddot{\underline{R}} = \sum_i m_i \ddot{\underline{r}}_i$$

So that we have,

$$\sum_i \underline{F}^{ext}_i = M\ddot{\underline{R}}$$

This tells us that the many-particle system behaves as if all the mass, m_i, were concentrated at the centre of mass, $\underline{R}$. Under the same external forces the centre of mass of a particle, a

Football or a strawberry trifle all follow the same shaped path.

A nice example is that of a space rocket: in order for the motion of the centre of mass to be unaltered when the rocket jettisons exhaust gases the body of the rocket is propelled "ahead" of the centre of mass while the gases "fall behind" the centre of mass.

We can define a linear momentum for a system of particles using the centre of mass:

$$\underline{P} = \sum_i m_i \frac{d\underline{r}_i}{dt}$$

$$\equiv M\dot{\underline{R}}.$$

And so $\sum_i \underline{F}_i^{ext} = \dot{\underline{P}}$ and if the external forces are zero then the total linear momentum is conserved.

Angular momentum for a multiparticle system is found from $\underline{L}_i = \underline{r}_i \wedge \underline{p}_i$ summed over all the particles.

$$\sum_i (\underline{r}_i \wedge \underline{p}_i) \equiv \underline{L}$$

But what is the total torque? Can we find $\underline{N} = \dot{\underline{L}}$?

Total torque, $\underline{N} = \sum_i \left(\underline{r}_i \wedge \left(\underline{F}_i^{ext} + \sum_j \underline{F}_{ij}^{int} \right) \right)$

$$= \sum_i (\underline{r}_i \wedge \underline{F}_i^{ext}) + \sum_{ij} (\underline{r}_i \wedge \underline{F}_{ij}^{int}).$$

N.B. $\dot{\underline{L}} = \sum_i (\dot{\underline{r}}_i \wedge \underline{p}_i + \underline{r}_i \wedge \dot{\underline{p}}_i) = \sum_i \left(\underline{r}_i \wedge \left(\underline{F}_i^{ext} + \sum_j \underline{F}_{ij}^{int} \right) \right) = \underline{N}.$

($\dot{\underline{r}}_i \wedge \underline{p}_i \to 0$ as $\underline{p}_i = m\dot{\underline{r}}_i$).

We are concerned that the total torque seems to depend on the internal forces while the other properties we have considered are functions of the external forces only. Under a minor assumption we can remove $\underline{F}_{ij}^{int}$ from the total torque.

The term we would like to remove can be written as a sum of pairs of torques:

$$\underline{r}_i \wedge \underline{F}_{ij}^{(int)} + \underline{r}_j \wedge \underline{F}_{ji}^{(int)}$$

$$= (\underline{r}_i - \underline{r}_j) \wedge \underline{F}_{ij}^{(int)}.$$

If we assume that the internal force between particle i and particle j

is in the direction of the straight line between $\underline{r}_i$ and $\underline{r}_j$ then:

$$\underline{F}_{ij}^{(int)} \propto (\underline{r}_i - \underline{r}_j).$$

In which case $(\underline{r}_i - \underline{r}_j) \wedge \underline{F}_{ij}^{(int)} = \underline{0}$, and the internal torque does not contribute to the total torque:

$$\underline{N} = \sum_i (\underline{r}_i \wedge \underline{F}_i^{(ext)}).$$

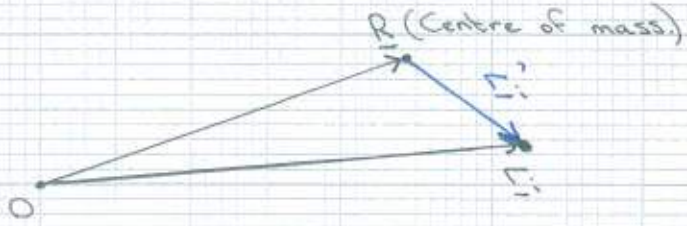
Under this assumption the rate of change of angular velocity of a system of particles is dependent only on the external forces applied.

We saw that the total linear momentum was very simple for a multiparticle system, it is as if all the mass were at the centre of mass (and moving with it). There is not as simple a statement for the angular momentum

Let's write the position of a single particle in a system, $\underline{r}_i$, as

$$\underline{r}_i = \underline{r}_i' + \underline{R}$$

Where $\underline{r}_i'$ is the vector displacement of $\underline{r}_i$ from the centre of mass:



Now,

$$\begin{aligned}
\underline{L} &= \sum_i (\underline{r}_i \wedge m_i \dot{\underline{r}}_i) \\
&= \sum_i (\underline{r}_i \wedge m_i \dot{\underline{r}}_i' + \underline{r}_i \wedge m_i \dot{\underline{R}}) \\
&= \sum_i ((\underline{r}_i' + \underline{R}) \wedge m_i \dot{\underline{r}}_i' + (\underline{r}_i' + \underline{R}) \wedge m_i \dot{\underline{R}}) \\
&= \sum_i (\underline{R} \wedge m_i \dot{\underline{R}}) + \sum_i (\underline{r}_i' \wedge m_i \dot{\underline{r}}_i') + \sum_i (m_i \underline{r}_i' \wedge \dot{\underline{R}}) + \sum_i (\underline{R} \wedge m_i \dot{\underline{r}}_i')
\end{aligned}$$

In fact this is more simple than it presently appears. Using the definition of the centre of mass we have:

$$M\underline{R} = \sum_i m_i \underline{r}_i = \sum_i m_i (\underline{r}_i' + \underline{R}) = \sum_i m_i \underline{r}_i' + M\underline{R}$$

$$\therefore \sum_i m_i \underline{r}_i' = 0 \; //$$

And the last two terms vanish leaving:

$$\underline{L} = \sum_i (\underline{R} \wedge m_i \dot{\underline{R}}) + \sum_i (\underline{r}_i' \wedge m_i \dot{\underline{r}}_i').$$

The total angular momentum is the sum of two terms. The first is the angular momentum of the centre of mass, the second term is the angular momentum of the individual particles about the centre of mass, all summed.

Let's finish our consideration of multiparticle systems by considering the work, energy and potentials in this setting.

If we move a system of particles at positions $\underline{r}_i(t_1)$ to $\underline{r}_i(t_2)$ the total work done is:

$$W_{12} = \sum_i \int_{t_1}^{t_2} \underline{F}_i \cdot d\underline{r}_i$$

$$\left[= \sum_i \int_{t_1}^{t_2} \left(\underline{F}_i^{(ext)} \cdot d\underline{r}_i + \sum_j \underline{F}_{ij}^{(int)} \cdot d\underline{r}_i \right) \right] \quad -(*)$$

As for the single particle case we substitute $\underline{F}_i = m_i \ddot{\underline{r}}_i$,

$$\therefore W_{12} = \sum_i \int_{t_1}^{t_2} m_i \ddot{\underline{r}}_i \cdot \dot{\underline{r}}_i \, dt.$$

$$= \sum_i \int_{t_1}^{t_2} m_i \frac{d}{dt}\left(\tfrac{1}{2} \dot{\underline{r}}_i^2\right) dt.$$

$$= \sum_i \left(\tfrac{1}{2} m_i \dot{\underline{r}}_i^2(t_2) - \tfrac{1}{2} m_i \dot{\underline{r}}_i^2(t_1) \right)$$

$$= T_2 - T_1.$$

Where $T_1 = \sum_i \tfrac{1}{2} m_i \dot{\underline{r}}_i^2(t_1)$ is the total kinetic energy at time t_1.

If we carried on the analysis we would find that if both internal and external forces are conservative then there would be a conservation of total energy $E = T + V$ where V would be the potential associated to both internal and external forces. Note from $(*)$ above that the internal forces do work as $\underline{r}_i(t)$ move. However for rigid bodies the internal part of the work function vanishes and V has only an external component in the conserved quantity E. In fact for rigid bodies $V^{(int)}$ will be a constant and so may be neglected when considering the dynamics.

2. LAGRANGIAN MECHANICS.

Towards a Lagrangian: Constraints.

Thus far we have considered dynamical systems whose motion can be described by $\underline{F} = \dot{\underline{p}}$ or, where appropriate, $\underline{N} = \dot{\underline{L}}$. The motion of the dynamical systems we have seen can be solved completely from these equations and two states of the system (usually $\underline{r}(t_0)$ and $\dot{\underline{r}}(t_0)$). In the single particle examples we have dealt mainly with free particles, we have only placed constraints on particles in a few cases, e.g. circular motion (the particle was constrained to move with fixed $|\underline{r}|$) or the harmonic oscillator. In general the constraints on the motion of a particle can be difficult to express, at least in the usual Cartesian coordinates. Consider how you would go about finding the motion of a bead on an arbitrarily curved string under gravity. Clearly the curve of the string plays an important role in the dynamics of the system but it is difficult to express in the coordinates used so far. In fact the problem is only a little more complicated than the example of the skier who was constrained to ski on the surface of the mountain. For multiparticle systems a very common constraint occurs for rigid bodies where the distance between particles is held fixed throughout the motion. Another example of a constraint is gas molecules constrained to move within a container. In short there are many dynamical systems whose motion is complicated by constraints.

Of course there are constraints placed on the motion of particles in all our examples due to forces. In the case of the harmonic oscillator a spring constrained the motion of the particle - in this example the origin of the constraining force was made known to us,

Compare this to the constrained radial length in circular motion where no physical explanation was offered - we were told at the outset that the motion was constrained such that $|\underline{r}|$ was constant. There could be a number of explanations for the constraining forces but note that we were able to work with the system without an understanding of the microscopic forces acting. In general when we impose constraints we will cover up some microscopic physics - it is not that we will neglect any physics but rather a constraint will essentially place some of the microscopic rules behind a black box that replicates the physics exactly. With a constraint we will effectively say: "I want to understand the system this far but no further." In practise constraints will allow us to simplify very complicated systems.

Constraints fit into a classification scheme. If the constraint can be expressed in terms of the positions of the particles and time as:

$$f(\underline{r}_1, \underline{r}_2, \ldots, t) = 0$$

Then the constraint is called "holonomic". The constraint in our circular motion example was holonomic as it could be written as:

$$|\underline{r}| = r$$

$$\Rightarrow \quad \underline{r}^2 - r^2 = 0. \quad //$$

The constraints for a rigid body (that the separation between particles) is a constant may be written:

$$|\underline{r}_i - \underline{r}_j| = c_{ij}.$$

$$\Rightarrow (\underline{r}_i - \underline{r}_j)^2 - c_{ij}^2 = 0.$$

This is another holonomic constraint; there is one for every pair of particles in a rigid body.

For a bead on a curved wire the constraint describes the curve of the wire in terms of the components of $\underline{r}$. It is also holonomic.

Constraints of a different type are called "nonholonomic". The constraints of a gas moving in a box are nonholonomic.

When a holonomic constraint exists the coordinates of the particles are no longer independent. Consider a rigid body having constituent particles we can single out one particle and then write down $(n-1)$ constraints specifying the fixed distances of all the other particles from the first one. By using the constraints we no longer need to solve n equations of motion, instead we solve the equation of motion of the singled out particle and use the $(n-1)$ constraints to rewrite the solution for the $(n-1)$ other particles.

Let us now make the leap from Cartesian coordinates to "generalised coordinates".

Our aims so far have been to solve Newton's equation $\underline{F} = \dot{\underline{p}}$ to find $\underline{r}(t)$, where $\underline{r}(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix}$ has been a three-dimensional vector of Cartesian coordinates. Or, for multiparticle systems, we have aimed to find $\underline{r}_i(t)$ where $i = \{1, 2, \ldots n\}$ and n is the number of particles in the system. We have used Cartesian coordinates because they are well-known to us but in most of our examples we have complicated the situation by doing so. In fact due to the presence of implicit constraints (For "constraints" think: "sets of relations between the coordinates") not all the Cartesian coordinates are independent and the solution $\underline{r}_i(t)$ depends on only a subset of the coordinates. Let us quantify these comments. For a system of n particles there are $3n$ Cartesian coordinates to specify to identify a configuration. If there are k constraints then

the configuration of the system is given by only $3n-k$ of the coordinates.

Let us reflect on a familiar example: the skier going down the mountain in a straight line. We count the coordinates and the implicit constraints:

Coordinates: We must find $\underline{r}(t)$ $\Rightarrow$ 3 coordinates.

Constraints: (1.) The skier skis down the mountain's surface (the equation describing the surface of the mountain is a constraint on x, y, z).

(2.) The skier moves in a straight line.

$\therefore$ 3 coordinates minus 2 constraints gives a one-dimensional solution. There is a canonical single coordinate that best describes the solution, it is the distance moved down the slope. Notice that in our example we automatically focussed on this coordinate, it is not x, y or z - it is an example of a generalised coordinate:

The configuration of a system of n particles subject to k constraints is given by $3n-k$ generalised coordinates.

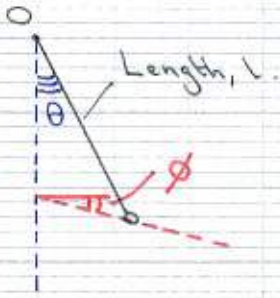
We are freed from the constraints of Cartesian coordinates and may now pick whichever $3n-k$ parameters that most easily describe the system. We are really parameterising the $\underline{r}_i$ by the generalised coordinates (labelled $q_1, q_2 \dots q_{3n-k}$), k constraints and, of course, time, t:

$$\underline{r}_i = \underline{r}_i(q_1, q_2, \dots q_{3n-k}, t)$$

Together with the k constraints we have specified the system.

Example: The pendulum attached by a rigid rod.

We consider a 3D system of 1 particle, representing the pendulum bob, acting under gravity $\Rightarrow$ $\underline{r}(t)$ has 3 components.



There is only 1 constraint: that the pendulum is constructed from a rigid rod. If the rod has length l then the motion of the bob is always a distance l from O indicated on the diagram above.

i.e. there are 3 Cartesian coordinates and 1 constraint

$\Rightarrow$ 2 generalised coordinates are needed for the motion.

In this example a sensible choice of coordinates would be spherical coordinates with r fixed equal to l leaving θ, ϕ as coordinates.

N.B. by giving θ and ϕ the position of the bob is specified.

Generalised coordinates need not be chosen from any formal sets of coordinates that you have seen before. But they should be chosen to make calculations easier and they must specify the position of the system at time t completely. A bead constrained to move on a wire can be described by an unconventional coordinate: distance moved along the wire. Of course it could also be described by x, y, or z the other two coordinates being deduced from the equation describing the shape of the wire.

Configuration Space.

Having said that a system of n particles with k constraints may be described by $3n-k$ generalised coordinates it is appropriate to introduce the concept of a configuration space, although we will not use the idea until later in the course. Given $3n-k$ generalised coordinates (which are scalars) we can plot a point in a $(3n-k)$-dimensional vector space given by those coordinates. The $3n-k$ numbers specify completely the positions of the n particles at a time t, also called the configuration of the

system. In fact each point in configuration space corresponds to some state of the multiparticle system. The dynamical motion of the system in 3D space is given by a line in configuration space which is paramaterised by t. At any particular t one could read off the generalised coordinates of all the particles in the system from the curve in configuration space. Later in the course we will make use of configuration space in a derivation of LaGrange's equations using variational calculus.

Virtual Work, Forces of Constraint & LaGrange's Equations.

A dynamical system of particles may have forces acting on its particles which are changing with time. In this case the definition of work begins to seem complicated. Instead we begin to work with the concept of virtual work. When calculating the virtual work done one considers the system frozen at an instant of time, t. Then the work done against the constant forces at time t to move the system in space a distance $\delta \underline{r}_i$, subject to the constraints, is called the virtual work. It is a form of work calculated from the forces acting at time t.

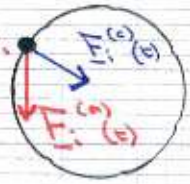
As we have stated a constraint encodes the action of a force. In this section we will make the force of the constraint explicit. Let us denote the force on a particle i as:

$$\underline{F}_i = \underline{F}_i^{(a)} + \underline{F}_i^{(c)}$$

$\underline{F}_i^{(a)}$: Applied force.

$\underline{F}_i^{(c)}$: Force of constraint.

As an example consider the force acting on a particle constrained to move on a vertical circular frame under gravity.



The force of constraint $\underline{F}_i^{(c)}(t)$ imposes the circular motion and points towards the centre of the circle. The applied force $\underline{F}_i^{(a)}(t)$ is gravity and is constant.

The total force on the particle is:

$$\underline{F}_i(t) = \underline{F}_i^{(a)}(t) + \underline{F}_i^{(c)}(t).$$

(The index i is superfluous here.)

Notice that $\delta\underline{r}_i(t)$ which is tangent to the circle at time, t, is orthogonal to the force of constraint, so that the force of constraint does not contribute to the virtual work over $\delta\underline{r}_i(t)$:

$$\underline{F}_i(t) \cdot \delta\underline{r}_i(t) = \underline{F}_i^{(a)}(t) \cdot \delta\underline{r}_i(t) + \underline{F}_i^{(c)}(t) \cdot \delta\underline{r}_i(t)$$

$$\underline{F}_i^{(c)}(t) \cdot \delta\underline{r}_i(t) \to 0.$$

In fact whenever a particle is constrained to move along a surface the force of the constraint (which cause the particle to be "stuck" to the surface) will act perpendicularly to the translation along the surface $\delta\underline{r}_i$. So that its contribution to the virtual work will vanish.

We are now ready to show that LaGrange's equations describe the dynamics of a system.

From (NII) we have, $\underline{F}_i = \dot{\underline{p}}_i$

$$\therefore \underline{F}_i - \dot{\underline{p}}_i = \underline{0}$$

And so, $$\sum_i (\underline{F}_i - \dot{\underline{p}}_i) \cdot \delta\underline{r}_i = 0.$$

If we resolve $\underline{F}_i$ into applied forces and forces of constraint we have:

$$\sum_i (\underline{F}_i^{(a)} - \dot{\underline{p}}_i) \cdot \delta\underline{r}_i + \sum_i \underline{F}_i^{(c)} \cdot \delta\underline{r}_i = 0.$$

Now we make an assumption that the virtual work due to the forces of constraint is zero:

$$\sum_i (\underline{F}_i^{(a)} - \dot{\underline{p}}_i) \cdot \delta\underline{r}_i = 0. \qquad -(*)$$

By our assumption we have removed the forces of constraint from our equation. We will now develop this expression to find the equations of motion of the system.

In a system without constraints all the $\delta \underline{r}_i$ would be independent. The only way for the expression above to vanish (if $\delta \underline{r}_i \neq 0$) is if

$$\underline{F}_i^{(a)} - \dot{\underline{p}}_i = \underline{0}$$

However we are interested in systems of particles with constraints. If we have holonomic constraints then not all the $\delta \underline{r}_i$ are independent. Alternatively, by definition, if we work with generalised coordinates q_i then there are $3n-k$ of them and they are linearly dependent. If, somehow, we were able to rewrite the equation denoted (*) on the previous page as a sum of virtual work functions done with respect to generalised coordinates it would have the form:

$$\sum_{i=1}^{3n-k} (\#_i)\, \delta q_i = 0 \qquad \text{---(**)}$$

And since the δq_i are independent then (**) could only be true if all the functions $\#_i = 0$. These would be the equations of motion of the system. These $\#_i$ are what we shall now find.

Recall that $\underline{r}_i \equiv \underline{r}_i(q_1, q_2, \cdots q_{3n-k}, t)$.

So that a variation in $\underline{r}_i$ is:

$$\delta \underline{r}_i = \sum_{j=1}^{3n-k} \frac{\partial \underline{r}_i}{\partial q_j} \delta q_j + \frac{\partial \underline{r}_i}{\partial t} \cdot \delta t.$$

Since virtual work is measured at an exact instant of time, $\delta t = 0$.
Substituting into (*) gives:

$$\sum_{i=1}^{n} \left(\underline{F}_i^{(a)} - \dot{\underline{p}}_i \right) \cdot \sum_{j=1}^{3n-k} \frac{\partial \underline{r}_i}{\partial q_j} \delta q_j = 0.$$

We have introduced δq_i as in (**) now let us rewrite the rest of the equation in terms of generalised coordinates.
First we will simplify our notation, by introducing the "generalised force" Q_j:

$$Q_j \equiv \sum_{i=1}^{n} \underline{F}_i^{(a)} \cdot \frac{\partial \underline{r}_i}{\partial q_j}$$

Then our equation becomes:

$$\sum_{j=1}^{3n-k} Q_j \cdot \delta q_j - \sum_{i=1}^{n} \dot{\underline{p}}_i \cdot \sum_{j=1}^{3n-k} \left(\frac{\partial \underline{r}_i}{\partial q_j}\right) \delta q_j = 0. \qquad (***)$$

Let us rewrite the second term using $\underline{p}_i = m_i \underline{\dot{r}}_i$:

$$\sum_{i=1}^{n} \dot{\underline{p}}_i \cdot \sum_{j=1}^{3n-k} \frac{\partial \underline{r}_i}{\partial q_j} \delta q_j = \sum_{i=1}^{n} m_i \ddot{\underline{r}}_i \cdot \sum_{j=1}^{3n-k} \frac{\partial \underline{r}_i}{\partial q_j} \delta q_j.$$

(We have assumed $\dot{m}_i = 0$.)

Momentarily ignoring the sum over j we can rewrite this term using:

$$\sum_{i=1}^{n} m_i \ddot{\underline{r}}_i \cdot \frac{\partial \underline{r}_i}{\partial q_j} \delta q_j = \sum_{i=1}^{n} \left\{ \frac{d}{dt}\left(m_i \dot{\underline{r}}_i \cdot \frac{\partial \underline{r}_i}{\partial q_j} \right) - m_i \dot{\underline{r}}_i \cdot \frac{d}{dt}\left(\frac{\partial \underline{r}_i}{\partial q_j} \right) \right\} \delta q_j$$

Since $\underline{r}_i = \underline{r}_i(q_1, \ldots q_{3n-k}, t)$ then, by the chain rule,

$$\dot{\underline{r}}_i = \sum_{j=1}^{3n-k} \frac{\partial \underline{r}_i}{\partial q_j} \dot{q}_j + \frac{\partial \underline{r}_i}{\partial t}.$$

And so $$\frac{\partial}{\partial q_j}\left(\dot{\underline{r}}_i\right) = \frac{\partial \underline{r}_i}{\partial q_j}.$$

We use this and the observation that $\frac{d}{dt}\left(\frac{\partial \underline{r}_i}{\partial q_j}\right) = \frac{\partial}{\partial q_j}\left(\frac{d \underline{r}_i}{dt}\right) = \frac{\partial}{\partial q_j}\left(\dot{\underline{r}}_i\right)$

to find:

$$\sum_{i=1}^{n} m_i \ddot{\underline{r}}_i \cdot \frac{\partial \underline{r}_i}{\partial q_j} \delta q_j = \sum_{i=1}^{n} \left\{ \frac{d}{dt}\left(m_i \dot{\underline{r}}_i \cdot \frac{\partial \dot{\underline{r}}_i}{\partial \dot{q}_j} \right) - m_i \dot{\underline{r}}_i \cdot \frac{\partial \dot{\underline{r}}_i}{\partial q_j} \right\} \delta q_j.$$

Now we can rewrite $(***)$ as:

$$\sum_{j=1}^{3n-k} \left[Q_j - \frac{d}{dt}\left\{ \frac{\partial}{\partial \dot{q}_j}\left(\sum_{i=1}^{n} \frac{1}{2} m_i \dot{\underline{r}}_i^2 \right) \right\} + \frac{\partial}{\partial q_j}\left\{ \sum_{i=1}^{n} \frac{1}{2} m_i \dot{\underline{r}}_i^2 \right\} \right] \cdot \delta q_j = 0.$$

We recognise $\sum_i \frac{1}{2} m_i \dot{\underline{r}}_i^2$ as the kinetic energy of the system, denoted T.

Then we have:

$$\sum_{j=1}^{3n-k} \left[Q_j - \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j} \right) + \frac{\partial T}{\partial q_j} \right] \delta q_j = 0.$$

We have finally found our equations of motion (denoted #, earlier):

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j} \right) - \frac{\partial T}{\partial q_j} - Q_j = 0.$$

If the applied forces, $\underline{F}_i^{(a)}$, are conservative then they are derived from a scalar potential, V:

$$\underline{F}_i^{(a)} = -\underline{\nabla}_i V.$$

Where $\underline{\nabla}_i = \begin{pmatrix} \partial/\partial x_i \\ \partial/\partial y_i \\ \partial/\partial z_i \end{pmatrix}$.

Consequently the generalised forces Q_j are:

$$Q_j = \sum_i \underline{F}_i \cdot \frac{\partial \underline{r}_i}{\partial q_j} = -\sum_i \underline{\nabla}_i V \cdot \frac{\partial \underline{r}_i}{\partial q_j}$$

Since V is a function of all the positions of the particles $\underline{r}_i(t)$

i.e. $V \equiv V(\underline{r}_1, \underline{r}_2, \ldots \underline{r}_n, t) \equiv V((x_1,y_1,z_1),(x_2,y_2,z_2),\ldots(x_n,y_n,z_n),t)$.

We have:

$$Q_j = -\sum_i \begin{pmatrix} \partial V/\partial x_i \\ \partial V/\partial y_i \\ \partial V/\partial z_i \end{pmatrix} \cdot \begin{pmatrix} \partial x_i/\partial q_j \\ \partial y_i/\partial q_j \\ \partial z_i/\partial q_j \end{pmatrix}$$

$$= -\sum_i \left(\frac{\partial V}{\partial x_i}\frac{\partial x_i}{\partial q_j} + \frac{\partial V}{\partial y_i}\frac{\partial y_i}{\partial q_j} + \frac{\partial V}{\partial z_i}\frac{\partial z_i}{\partial q_j} \right)$$

$$= -\frac{\partial V}{\partial q_j}$$

∴ the equation of motion becomes:

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

Since V is defined as being independent of the velocities or generalised velocities then $\partial V/\partial \dot{q}_j = 0$ and we can write:

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

Defining $T-V \equiv L$ as a new function the Lagrangian we arrive at Lagrange's equations (also known as the Euler-Lagrange equations):

$$\boxed{\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} = 0.}$$

Comment: The Lagrangian, L, is not unique for a system.

If L satisfies Lagrange's equations then so does

$$L' = L + \frac{dF}{dt}(q_1 \cdots q_{3n-k}, t).$$

N.B. $\dfrac{dF}{dt} = \displaystyle\sum_{i=1}^{3n-k} \frac{\partial F}{\partial q_i}\cdot\dot{q}_i + \frac{\partial F}{\partial t} \Rightarrow \frac{\partial}{\partial \dot{q}_j}\left(\frac{dF}{dt}\right) = \frac{\partial F}{\partial q_j}$.

Now substitute L' into Lagrange's equations: (instead of L)

$$\frac{d}{dt}\left(\frac{\partial L'}{\partial \dot{q}_j}\right) - \frac{\partial L'}{\partial q_j} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j} + \frac{\partial}{\partial \dot{q}_j}\left(\frac{dF}{dt}\right)\right) - \frac{\partial}{\partial q_j}\left(L + \frac{dF}{dt}\right)$$

$$= \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} + \frac{d}{dt}\frac{\partial F}{\partial q_j} - \frac{\partial}{\partial q_j}\frac{dF}{dt}$$

The first two terms vanish by Lagrange's equations, and the second two cancel each other out (the derivatives commute: $\frac{\partial^2 F}{\partial t \partial q} = \frac{\partial^2 F}{\partial q \partial t}$.).

So for any Lagrangian, L, there are always others, L', related as shown and also satisfying Lagrange's equations.

Our aim now is to write down the Lagrangian and use Lagrange's equations to find the equations of motion for a number of examples.

Example 1A: The Single Particle in Cartesian Coordinates.

We will identify the coordinates and from these write down the kinetic energy. The particle is free => no potential.

$$\therefore L = T - V = T.$$

We work with Cartesian coordinates: x, y and z. (N.B. $q_1 \equiv x, q_2 \equiv y, q_3 \equiv z$)

$$\therefore T = \tfrac{1}{2} m (\dot{x}^2 + \dot{y}^2 + \dot{z}^2) \equiv \tfrac{1}{2} \sum_{i=1}^{3} m_i \dot{q}_i^2$$

$$= L.$$

Notice that with no constraints we expect to find Newton's equation.

We use Lagrange's equations to find the equations of motion:

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

Let us write the 3 equations out for our generalised coordinates:

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

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

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{z}}\right) - \frac{\partial L}{\partial z} = 0.$$

Using $L = \tfrac{1}{2} m (\dot{x}^2 + \dot{y}^2 + \dot{z}^2)$ we have:

$$m\ddot{x} = 0.$$

$$m\ddot{y} = 0.$$

$$m\ddot{z} = 0.$$

Had we included an unspecified conservative force, $F = -\nabla V$ then

we would have found for the x-coordinate:
(using $L = T - V$)

$$\frac{d}{dt}\left(\frac{\partial}{\partial \dot{x}}(T-V)\right) - \frac{\partial (T-V)}{\partial x} = 0.$$

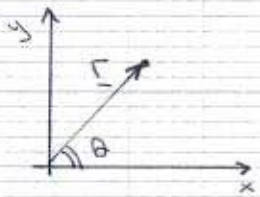
$$\frac{d}{dt}\left(m\dot{x}\right) + \frac{\partial V}{\partial x} = 0.$$

$$\therefore \quad m\ddot{x} = -\frac{\partial V}{\partial x} \equiv F_x \;//$$

Where $\underline{F} = -\nabla V = -\begin{pmatrix} \partial V/\partial x \\ \partial V/\partial y \\ \partial V/\partial z \end{pmatrix} = \begin{pmatrix} F_x \\ F_y \\ F_z \end{pmatrix}$.

So we are able to reproduce Newton's equations. //

Example 18: The Single Particle in Plane Polar coordinates.



In plane polar coordinates our generalised coordinates are r and θ.
We must still find the kinetic energy but now in terms of $\dot{r}$ and $\dot{\theta}$.
In Cartesian coordinates:

$$T = \tfrac{1}{2} m \dot{x}^2 + \tfrac{1}{2} m \dot{y}^2.$$

In terms of polar coordinates:

$$x = r\cos\theta.$$
$$y = r\sin\theta.$$
$$\therefore \quad \dot{x} = \dot{r}\cos\theta - r\sin\theta\,\dot{\theta}.$$
$$\dot{y} = \dot{r}\sin\theta + r\cos\theta\,\dot{\theta}.$$
$$\therefore T = \tfrac{1}{2} m\left(\dot{r}^2\cos^2\theta - 2r\dot{r}\cos\theta\sin\theta\,\dot{\theta} + r^2\sin^2\theta\,\dot{\theta}^2\right)$$
$$+ \tfrac{1}{2} m\left(\dot{r}^2\sin^2\theta + 2r\dot{r}\cos\theta\sin\theta\,\dot{\theta} + r^2\cos^2\theta\,\dot{\theta}^2\right)$$
$$= \tfrac{1}{2} m\dot{r}^2 + \tfrac{1}{2} m r^2\dot{\theta}^2.$$

One could have arrived at this by considering the displacements found by varying the generalised coordinates:

$$d\underline{r} = \hat{\underline{r}}\,dr + r\hat{\underline{\theta}}\,d\theta$$

i.e. in the $\hat{\underline{r}}$ direction a change in r is dr, while in the $\hat{\underline{\theta}}$ direction a change in θ, $d\theta$ gives a change of $r\,d\theta$ to $d\underline{r}$.

$$\therefore \quad \dot{\underline{r}} = \dot{r}\hat{\underline{r}} + r\dot{\theta}\hat{\underline{\theta}}$$

and $(\dot{\underline{r}})^2 = \dot{r}^2 + r^2\dot{\theta}^2$. //

As before let's find the equations of motion in the absence of a force.

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

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

$\therefore \quad m\ddot{r} - mr\dot{\theta}^2 = 0.$

The second term is the centripetal acceleration term.

For the θ equation: $\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right) - \frac{\partial L}{\partial \theta} = \frac{d}{dt}(mr^2\dot{\theta}) = 0.$

$\therefore \quad 2mr\dot{r}\dot{\theta} + mr^2\ddot{\theta} = 0.$

In fact if we had included arbitrary forces we could have used the equation: $\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_i}\right) - \frac{\partial T}{\partial q_i} = Q_i$

The equations above would have been the same for r and θ but we would have Q_r and Q_θ on the right-hand-side instead of zero.

Recall that $Q_i = \sum_j \underline{F}_j \cdot \frac{\partial \underline{r}_j}{\partial q_i}$ are the generalised forces. In this example:

$$Q_r = \underline{F} \cdot \frac{\partial \underline{r}}{\partial r} = \underline{F} \cdot \hat{\underline{r}} = F_r.$$

$$Q_\theta = \underline{F} \cdot \frac{\partial \underline{r}}{\partial \theta} = \underline{F} \cdot r\hat{\underline{\theta}} = rF_\theta$$

$\therefore$ the equations of motion, in the presence of a force, are:

$$m\ddot{r} - mr\dot{\theta}^2 = F_r.$$

$$2mr\dot{r}\dot{\theta} + mr^2\ddot{\theta} = rF_\theta.$$

N.B. recall that $|\underline{L}| = mr^2\dot{\theta}$ $\therefore$ $|\dot{\underline{L}}| = 2mr\dot{r}\dot{\theta} + mr^2\ddot{\theta}$.

$\therefore$ The second equation is really

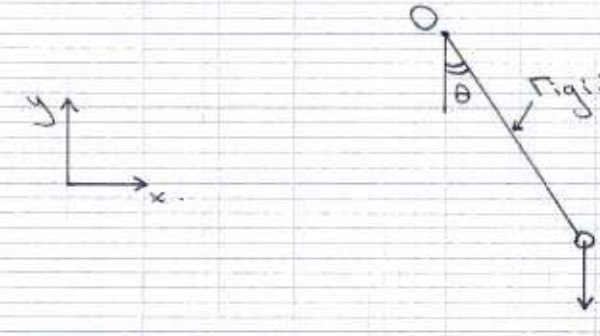
$$|\dot{\underline{L}}| = |\underline{N}|$$

Where $|\underline{N}| = rF_\theta$. //

Example 2: The Pendulum with a rigid rod, in the vertical plane.

The motion is constrained to the plane and by the rigidity of the rod.

We need to pick $(3n-k) = 3-2 = 1$ generalised coordinate:



θ is a convenient choice. Now we must write down the Lagrangian, $L = T - V$.

$$T = \tfrac{1}{2} m (\dot{x}^2 + \dot{y}^2).$$

In terms of our generalised coordinate, θ:

$$x = l \sin\theta.$$
$$y = -l \cos\theta.$$
$$\therefore \quad \dot{x} = l \cos\theta . \dot{\theta}$$
$$\dot{y} = l \sin\theta . \dot{\theta}$$
$$\therefore T = \tfrac{1}{2} m \left(l^2 \cos^2\theta . \dot{\theta}^2 + l^2 \sin^2\theta . \dot{\theta}^2 \right)$$
$$= \tfrac{1}{2} m l^2 \dot{\theta}^2.$$

There is a gravitational force (a conservative force) acting and so there is an associated potential, V.

What is V?

$$\underline{F} = -\nabla V = \begin{pmatrix} 0 \\ -mg \end{pmatrix}.$$

$$\therefore \quad V = +mgy. = -mgl\cos\theta.$$

Hence the Lagrangian for the pendulum is:

$$L = T - V$$
$$= \tfrac{1}{2} m l^2 \dot{\theta}^2 + mgl\cos\theta.$$

We find the equation of motion from Lagrange's equation $\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0$. Where $q_i = \theta$:

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

$$\therefore \quad \frac{d}{dt}\left(ml^2\dot{\theta}\right) + mgl\sin\theta = 0.$$

$$\therefore \quad ml^2\ddot{\theta} + mgl\sin\theta = 0.$$

$$\Rightarrow \quad \ddot{\theta} = -\left(\frac{g}{l}\right)\sin\theta.$$

This is the equation of motion for the pendulum. It is a complicated equation and it is often simplified to the approximation at small angles, θ.

N.B. $\sin(\theta) = \theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \ldots$

If $\theta \ll 1$ then $\sin(\theta) \simeq \theta$.

In this case, $\ddot{\theta} = -\left(\frac{g}{l}\right)\theta$.

And we can solve for θ to find:

$$\theta = A\cos\left(\sqrt{\tfrac{g}{l}}\,t\right) + B\sin\left(\sqrt{\tfrac{g}{l}}\,t\right).$$

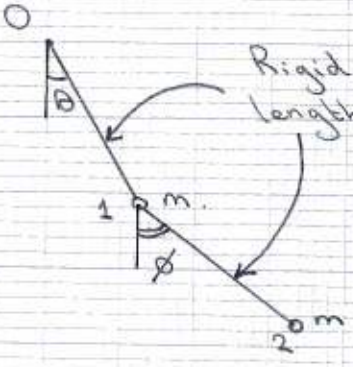
$\Rightarrow$ For small angles the frequency of θ's motion is $\sqrt{\frac{g}{l}}$.

One can use this to construct a clock using a pendulum or perhaps estimate the gravitational acceleration, g.

Example 3: The Compound Pendulum.

Let us consider a more complicated system but we will be able to follow the same systematic Lagrangian analysis.

Consider two pendula, one attached to the other like so:



The generalised coordinates are θ and ϕ.

If the position of bob 1 is $\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$ and that of bob 2 is $\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$ then the total kinetic energy of the system is:

$$T = \tfrac{1}{2}m\left(\dot{x}_1^2 + \dot{y}_1^2\right) + \tfrac{1}{2}m\left(\dot{x}_2^2 + \dot{y}_2^2\right).$$

We must write this in terms of θ and ϕ.

So, $x_1 = l\sin\theta$.

$y_1 = -l\cos\theta$.

$x_2 = l\sin\theta + l\sin\phi$.

$y_2 = -l\cos\theta - l\cos\phi$.

Taking the time derivatives:

$\dot{x}_1 = +l\cos\theta.\dot{\theta}$.

$\dot{y}_1 = l\sin\theta.\dot{\theta}$.

$\dot{x}_2 = l\cos\theta.\dot{\theta} + l\cos\phi.\dot{\phi}$.

$\dot{y}_2 = l\sin\theta.\dot{\theta} + l\sin\phi.\dot{\phi}$.

$$\therefore T = \tfrac{1}{2}m\left(l^2\cos^2\theta.\dot{\theta}^2 + l^2\sin^2\theta.\dot{\theta}^2 + l^2\cos^2\theta.\dot{\theta}^2 + l^2\cos^2\phi.\dot{\phi}^2 + 2l^2\cos\theta.\cos\phi.\dot{\theta}\dot{\phi} + l^2\sin^2\theta.\dot{\theta}^2 + l^2\sin^2\phi.\dot{\phi}^2 + 2l^2\sin\theta.\sin\phi.\dot{\theta}\dot{\phi}\right)$$

$$= \tfrac{1}{2}m\left(2l^2\dot{\theta}^2 + l^2\dot{\phi}^2 + 2l^2\dot{\theta}\dot{\phi}\cos(\theta-\phi)\right).$$

Where we have used: $\sin^2\theta + \cos^2\theta = 1$.

and: $\cos(\theta-\phi) = \cos\theta\cos\phi + \sin\theta\sin\phi$.

$$\therefore \underline{T = ml^2\dot{\theta}^2 + \tfrac{1}{2}ml^2\dot{\phi}^2 + ml^2\dot{\theta}\dot{\phi}\cos(\theta-\phi).}$$

Let us find V. We use $V = mgy$ for each pendulum bob:

$V = V_1 + V_2$

$= mgy_1 + mgy_2$

$= -mgl\cos\theta - mgl\cos\theta - mgl\cos\phi$.

$$\underline{= -2mgl\cos\theta - mgl\cos\phi.}$$

In all we have that the Lagrangian of the compound pendulum is:

$$L = T - V$$

$$\underline{= ml^2\dot{\theta}^2 + \tfrac{1}{2}ml^2\dot{\phi}^2 + ml^2\dot{\theta}\dot{\phi}\cos(\theta-\phi) + 2mgl\cos\theta + mgl\cos\phi.}$$

The Lagrange equations are:

(For θ) $\frac{d}{dt}\left(\frac{\partial L}{\partial\dot{\theta}}\right) - \frac{\partial L}{\partial\theta} = 0.$

$$\therefore \frac{d}{dt}\left(2ml^2\dot{\theta} + ml^2\dot{\phi}\cos(\theta-\phi)\right) + ml^2\dot{\theta}\dot{\phi}\sin(\theta-\phi) + 2mgl\sin\theta = 0.$$

$$\therefore 2ml^2\ddot{\theta} + ml^2\ddot{\phi}\cos(\theta-\phi) - ml^2\dot{\phi}\sin(\theta-\phi).(\dot{\theta}-\dot{\phi}) + ml^2\dot{\theta}\dot{\phi}\sin(\theta-\phi) + 2mgl\sin\theta = 0.$$

$\therefore 2ml^2\ddot{\theta} + ml^2\ddot{\phi}\cos(\theta-\phi) + ml^2\dot{\phi}^2\sin(\theta-\phi) + 2mgl\sin\theta = 0.$

$\therefore \quad \underline{\ddot{\theta} + \tfrac{1}{2}\ddot{\phi}\cos(\theta-\phi) + \tfrac{1}{2}\dot{\phi}^2\sin(\theta-\phi) + \left(\tfrac{g}{l}\right)\sin\theta = 0.}$

(For ϕ) $\qquad \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\phi}}\right) - \frac{\partial L}{\partial \phi} = 0.$

$\Rightarrow \frac{d}{dt}\left(ml^2\dot{\phi} + ml^2\dot{\theta}\cos(\theta-\phi)\right) - ml^2\dot{\theta}\dot{\phi}\sin(\theta-\phi) + mgl\sin\phi = 0.$

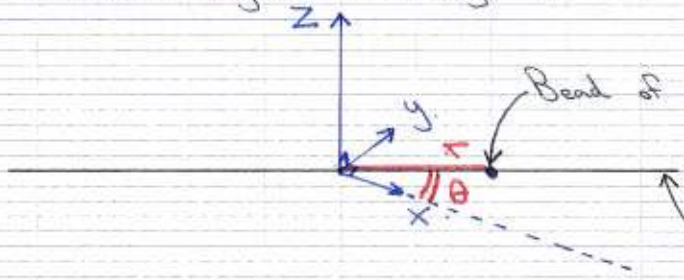
$\therefore ml^2\ddot{\phi} + ml^2\ddot{\theta}\cos(\theta-\phi) - ml^2\dot{\theta}\sin(\theta-\phi)(\dot{\theta}-\dot{\phi}) - ml^2\dot{\theta}\dot{\phi}\sin(\theta-\phi) + mgl\sin\phi = 0.$

$\therefore ml^2\ddot{\phi} + ml^2\ddot{\theta}\cos(\theta-\phi) - ml^2\dot{\theta}^2\sin(\theta-\phi) + mgl\sin\phi = 0.$

$\therefore \quad \underline{\ddot{\phi} + \ddot{\theta}\cos(\theta-\phi) - \dot{\theta}^2\sin(\theta-\phi) + \left(\tfrac{g}{l}\right)\sin\phi = 0.}$

Example 4: A Bead on a Rotating Wire.

In this example we will introduce a time-dependent constraint. Consider a straight wire rotating in the horizontal xy-plane about the z-axis. Let us find the equations of motion of a bead constrained to move, without friction, along the wire; the wire rotates at constant angular velocity, $\dot{\theta}$.



(Let us consider the wire rotating in an anticlockwise ↺ direction in the xy-plane.)

There is no motion in the z-direction so we will discard this coordinate from our analysis.

There are two ordinary coordinates which we pick to be r and θ, the polar coordinates in the xy-plane. In terms of the Cartesian coordinates we have:

$$x = r\cos\theta.$$

$$y = r\sin\theta.$$

The constraint that the bead lies on the rotating wire places a time-dependent constraint on θ, namely $\theta = \dot{\theta}t$. Where we have chosen θ such that $\theta = 0$ when $t = 0$.

We have:

$$x = r\cos(\dot{\theta}t).$$

$$y = r\sin(\dot{\theta}t).$$

We take the time derivatives:

$$\dot{x} = \dot{r}\cos(\dot{\theta}t) - r\dot{\theta}\sin(\dot{\theta}t) \qquad \text{Recall } \ddot{\theta} = 0.$$

$$\dot{y} = \dot{r}\sin(\dot{\theta}t) + r\dot{\theta}\cos(\dot{\theta}t)$$

The kinetic energy T is:

$$T = \tfrac{1}{2} m \left(\dot{x}^2 + \dot{y}^2 \right)$$

$$= \tfrac{1}{2} m \left((\dot{r}^2 \cos^2(\dot{\theta}t) - 2r\dot{r}\dot{\theta}\sin(\dot{\theta}t)\cos(\dot{\theta}t) + r^2\dot{\theta}^2\sin^2(\dot{\theta}t) \right.$$
$$\left. + (\dot{r}^2\sin^2(\dot{\theta}t) + 2r\dot{r}\dot{\theta}\sin(\dot{\theta}t)\cos(\dot{\theta}t) + r^2\dot{\theta}^2\cos^2(\dot{\theta}t) \right)$$

$$= \tfrac{1}{2} m \left(\dot{r}^2 + r^2\dot{\theta}^2 \right).$$

There is no potential specified in the problem and so $L = T$.

The one generalised coordinate is r, so the equation of motion is:

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

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

$$\equiv m\ddot{r} - mr\dot{\theta}^2.$$

Rearranging,

$$\ddot{r} = r\dot{\theta}^2.$$

Which is satisfied by $r = A e^{\dot{\theta}t} + B e^{-\dot{\theta}t}$.

If at $t = 0$ we have $r = r_0$ and $\dot{r} = 0$ then,

$$A + B = r_0.$$

And,

$$\dot{r} = A\dot{\theta}e^{\dot{\theta}t} - B\dot{\theta}e^{-\dot{\theta}t}.$$

$$\therefore \quad \dot{r}(t=0) = (A - B)\dot{\theta} = 0 \quad \Rightarrow A = B.$$

$$\therefore \quad r = \tfrac{1}{2} r_0 \left(e^{\dot{\theta}t} + e^{-\dot{\theta}t} \right)$$

$$\equiv r_0 \cosh(\dot{\theta}t).$$

If $r_0 \neq 0$ then the bead moves towards radial infinity as $t \to \infty$.

If $r_0 = 0$ the bead does not move from $r = 0$.

N.B. this is similar to circular motion without a force of constraint (the centripetal force) to cancel out the radial force, so the bead accelerates towards infinity.

Example 5: The Marble in a Bowl.

What are the equations of motion of a marble in a hemisphere, or bowl, under gravity?

The holonomic constraint is the surface of the bowl which we choose to be a hemisphere:

$$z^2 = x^2 + y^2. \qquad (\text{where } z \leq R, \text{ the radius of the bowl}).$$

There are 2 generalised coordinates, which due to the rotational symmetry of the problem, may be sensibly chosen as r, θ the polar coordinates in the xy-plane.

$$\therefore \quad x = r\cos\theta.$$

$$y = r\sin\theta.$$

$$z = (y^2 + x^2)^{1/2} = r.$$

The kinetic energy is:

$$T = \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2).$$

$$= \tfrac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2 + \dot{r}^2)$$

$$= m\dot{r}^2 + \tfrac{m}{2}r^2\dot{\theta}^2.$$

The gravitational potential energy is:

$$V = mgz$$

$$= mgr.$$

$\therefore$ the Lagrangian, $L \equiv T - V$, is:

$$L = m\dot{r}^2 + \tfrac{m}{2}r^2\dot{\theta}^2 - mgr.$$

The Lagrange equations are:

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

$$= 2m\ddot{r} - mr\dot{\theta}^2 - mg = 0.$$

(for θ) $$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right) - \frac{\partial L}{\partial \theta} = \frac{d}{dt}(mr^2\dot{\theta}) = 0 \qquad \text{(Conservation of angular momentum.)}$$

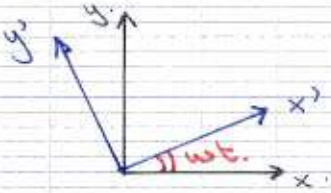
$$= 2mr\dot{r}\dot{\theta} + mr^2\ddot{\theta}$$

$$= 0.$$

Note that if we had specified that the marble had no angular velocity ($\dot{\theta} = 0$) then the equations are those of the simple pendulum, whose length is the radius of the bowl, R. We have solved the problem using polar coordinates instead of spherical coordinates (or a subset thereof) which we used when we considered the simple pendulum. If we allow $\dot{\theta} \neq 0$ then we have analysed a system equivalent to the spherical pendulum. We will consider the spherical pendulum later but first we will have an interlude on coordinate frames.

Example 6: The Coriolis Effect.

As we did in the previous example we will consider a transformation to cylindrical polar coordinates, but this time we will consider the Lagrangian of a free particle. The significant difference in this example is that we will consider the dynamics with respect to a rotating coordinate frame, rotating at constant angular velocity, $\dot{\theta} = \omega$.



$$\therefore \quad x = x'\cos(\omega t) - y'\sin(\omega t).$$
$$y = x'\sin(\omega t) + y'\cos(\omega t).$$
$$z = z'. \qquad \text{(Rotation about the z-axis.)}$$

So that,

$$\dot{x} = \dot{x}'\cos(\omega t) - x'\omega\sin(\omega t) - \dot{y}'\sin(\omega t) - y'\omega\cos(\omega t).$$
$$\dot{y} = \dot{x}'\sin(\omega t) + x'\omega\cos(\omega t) + \dot{y}'\cos(\omega t) - y'\omega\sin(\omega t).$$
$$\dot{z} = \dot{z}'$$

As $L = T - V$ we have,

$$T = \tfrac{1}{2}m\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right)$$
$$= \tfrac{1}{2}m\Big(\dot{x}'^2\cos^2(\omega t) - 2\dot{x}'x'\omega\sin(\omega t)\cos(\omega t) - 2\dot{x}'\cos(\omega t)\dot{y}'\sin(\omega t) - 2\dot{x}'\cos^2(\omega t)y'\omega$$
$$+ x'^2\omega^2\sin^2(\omega t) + 2x'\dot{y}'\omega\sin^2(\omega t) + 2x'y'\omega^2\sin(\omega t)\cos(\omega t)$$
$$+ \dot{y}'^2\sin^2(\omega t) + 2\dot{y}'y'\omega\cos(\omega t)\sin(\omega t) + y'^2\omega^2\cos^2(\omega t)$$
$$+ \dot{x}'^2\sin^2(\omega t) + 2\dot{x}'x'\omega\cos(\omega t)\sin(\omega t) + 2\dot{x}'\dot{y}'\cos(\omega t)\sin(\omega t)$$
$$- 2\dot{x}'\dot{y}'\omega\sin^2(\omega t) + x'^2\omega^2\cos^2(\omega t) + 2x'\dot{y}'\omega\cos^2(\omega t)$$
$$- 2x'\omega^2y'\cos(\omega t)\sin(\omega t) + \dot{y}'^2\cos^2(\omega t) - 2\dot{y}'y'\omega\cos(\omega t)\sin(\omega t)$$
$$+ y'^2\omega^2\sin^2(\omega t) + \dot{z}'^2\Big).$$
$$= \tfrac{1}{2}m\left(\dot{x}'^2 - 2\dot{x}'y'\omega + x'^2\omega^2 + 2x'\dot{y}'\omega + \dot{y}'^2 + y'^2\omega^2 + \dot{z}'^2\right).$$
$$L = \tfrac{1}{2}m\left((\dot{x}' - \omega y')^2 + (\dot{y}' + \omega x')^2 + \dot{z}'^2\right).$$

We can write this notation in vector notation:

Let, $\underline{\omega} \equiv \begin{pmatrix} 0 \\ 0 \\ \omega \end{pmatrix}$, $\underline{r}' \equiv \begin{pmatrix} \dot{x}' \\ \dot{y}' \\ \dot{z}' \end{pmatrix}$ & $\underline{r}' \equiv \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}$

Then, $\underline{\omega} \wedge \underline{r}' = \begin{pmatrix} 0 \\ 0 \\ \omega \end{pmatrix} \wedge \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} -\omega y' \\ \omega x' \\ 0 \end{pmatrix}$.

$$\therefore \quad L = \tfrac{1}{2} m \left(\underline{\dot{r}}' + (\underline{\omega} \wedge \underline{r}') \right)^2 .$$

We can now write the equation of motion in vector notation:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \underline{\dot{r}}'}\right) - \frac{\partial L}{\partial \underline{r}'} = \underline{0}.$$

N.B. $$\frac{\partial L}{\partial \underline{r}'} = m\left(\underline{\dot{r}}' + (\underline{\omega} \wedge \underline{r}')\right) \wedge \underline{\omega}$$

$$= m\left(\underline{\dot{r}}' \wedge \underline{\omega} - \underline{\omega} \wedge (\underline{\omega} \wedge \underline{r}')\right).$$

and $$\frac{d}{dt}\left(\frac{\partial L}{\partial \underline{\dot{r}}'}\right) = m\frac{d}{dt}\left(\underline{\dot{r}}' + (\underline{\omega} \wedge \underline{r}')\right) = m\left(\underline{\ddot{r}}' + (\underline{\omega} \wedge \underline{\dot{r}}')\right).$$

Giving: $$m\underline{\ddot{r}}' + m(\underline{\omega} \wedge \underline{\dot{r}}') - m(\underline{\dot{r}}' \wedge \underline{\omega}) + m(\underline{\omega} \wedge (\underline{\omega} \wedge \underline{r}')) = \underline{0}.$$

$$\therefore \quad \underline{m\ddot{r}}' + 2m(\underline{\omega} \wedge \underline{\dot{r}}') + m(\underline{\omega} \wedge (\underline{\omega} \wedge \underline{r}')) = \underline{0}.$$

Of course this kind of vector differentiation can be checked by working with the component equations. For reference we find:

(x'-equation) $$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}'}\right) - \frac{\partial L}{\partial x'} = \frac{d}{dt}(m\dot{x}' - m y'\omega) - m x'\omega^2 - m y'\omega = 0.$$

$$\therefore \quad m\ddot{x}' - m\dot{y}'\omega - m x'\omega^2 - m\dot{y}'\omega = 0.$$

$$\therefore \quad \underline{m\ddot{x}' - 2m\dot{y}'\omega - m x'\omega^2 = 0.}$$

(y'-equation) $$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{y}'}\right) - \frac{\partial L}{\partial y'} = \frac{d}{dt}\left(m x'\omega + m y'\right) + m\dot{x}'\omega - m y'\omega^2 = 0.$$

$$\therefore \quad m\omega\dot{x}' + m\ddot{y}' + m\dot{x}'\omega - m y'\omega^2 = 0.$$

$$\underline{m\ddot{y}' + 2m\omega\dot{x}' - m y'\omega^2 = 0.}$$

(z'-equation) $$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{z}'}\right) - \frac{\partial L}{\partial z'} = \frac{d}{dt}(m\dot{z}') = 0.$$

$$\therefore \underline{m\ddot{z}' = 0.}$$

These three equations agree with the vector equation of motion above, but took a little longer to derive. However if in doubt it is better to take a little more time and use components. The vector differential notation is also a little confusing. For example the variation of L above with respect to x' in the vector notation should be derived slowly:

$$L = \tfrac{1}{2} m\left(\underline{\dot{r}}' + (\underline{\omega} \wedge \underline{r}')\right) \cdot \left(\underline{\dot{r}}' + (\underline{\omega} \wedge \underline{r}')\right)$$

Then, varying x' only:

$$\begin{aligned}
\delta L &= \tfrac{1}{2} m \delta_{x'}\left(\dot{\underline{r}}' + (\underline{\omega} \wedge \underline{r}')\right) \cdot \left(\dot{\underline{r}}' + (\underline{\omega} \wedge \underline{r}')\right) \\
&\quad + \tfrac{1}{2} m \left(\dot{\underline{r}}' + (\underline{\omega} \wedge \underline{r}')\right) \cdot \delta_{x'}\left(\dot{\underline{r}}' + (\underline{\omega} \wedge \underline{r}')\right) \\
&= \tfrac{1}{2} m \left(\delta_{x'}(\underline{\omega} \wedge \underline{r}')\right) \cdot \left(\dot{\underline{r}}' + (\underline{\omega} \wedge \underline{r}')\right) \\
&\quad + \tfrac{1}{2} m \left(\dot{\underline{r}}' + (\underline{\omega} \wedge \underline{r}')\right) \cdot \delta_{x'}(\underline{\omega} \wedge \underline{r}') \\
&= \tfrac{1}{2} m \left(\underline{\omega} \wedge \begin{pmatrix} \delta x' \\ 0 \\ 0 \end{pmatrix}\right) \cdot \left(\dot{\underline{r}}' + (\underline{\omega} \wedge \underline{r}')\right) \\
&\quad + \tfrac{1}{2} m \left(\dot{\underline{r}}' + (\underline{\omega} \wedge \underline{r}')\right) \cdot \left(\underline{\omega} \wedge \begin{pmatrix} \delta x' \\ 0 \\ 0 \end{pmatrix}\right) \\
&= m \left(\underline{\omega} \wedge \begin{pmatrix} \delta x' \\ 0 \\ 0 \end{pmatrix}\right) \cdot \dot{\underline{r}}' + m \left(\underline{\omega} \wedge \begin{pmatrix} \delta x' \\ 0 \\ 0 \end{pmatrix}\right) \cdot (\underline{\omega} \wedge \underline{r}') \\
&= m \left(\dot{\underline{r}}' \wedge \underline{\omega}\right) \cdot \begin{pmatrix} \delta x' \\ 0 \\ 0 \end{pmatrix} + m \left((\underline{\omega} \wedge \underline{r}') \wedge \underline{\omega}\right) \cdot \begin{pmatrix} \delta x' \\ 0 \\ 0 \end{pmatrix} \\
&= m \left[(\dot{\underline{r}}' \wedge \underline{\omega}) - \underline{\omega} \wedge (\underline{\omega} \wedge \underline{r}')\right] \cdot \begin{pmatrix} \delta x' \\ 0 \\ 0 \end{pmatrix}.
\end{aligned}$$

We were ultra-careful in this calculation in order to demonstrate the meaning of the vector equation earlier. The manipulation need not be this lengthy but one should take care with variations of vectors.

Returning to the equation of motion for the free particle:

$$m\ddot{\underline{r}}' + 2m(\underline{\omega} \wedge \dot{\underline{r}}') + m(\underline{\omega} \wedge (\underline{\omega} \wedge \underline{r}')) = \underline{0}$$

First note that if there were no relative angular velocity between the two coordinate frames, $\underline{\omega} = \underline{0}$ and the last two terms vanish giving the usual free particle equation of motion.

The second two terms are due to the coordinate frame and are sometimes called "fictitious forces."

If there were no radial velocity ($\dot{\underline{r}}' = 0$) the second term would vanish and we would be considering a circular motion about $\underline{\omega}$. Therefore we recognise the centrifugal force due to a circular motion: $-m\underline{\omega} \wedge (\underline{\omega} \wedge \underline{r}') = \underline{F}_{\text{centrifugal}}$.

When the particle is moving the second term appears and acts as a force which is perpendicular to $\underline{\omega}$ and to $\dot{\underline{r}}'$. The action of this term to rotate the particle from its path in the direction of $\dot{\underline{r}}'$ is called the Coriolis effect.

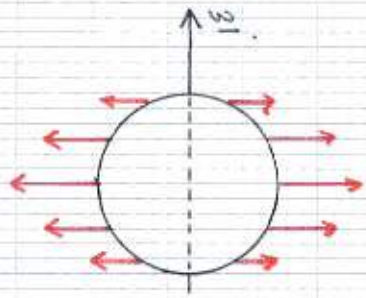
The Coriolis effect is responsible for the deflection of particles as judged from the rotating frame of the Earth.



Anticlockwise rotation of the Earth.

There are 2 forces acting on the free particle due to the Earth's rotation: the centrifugal force and the Coriolis force.

The centrifugal force always acts outward (radially) from the axis of rotation.

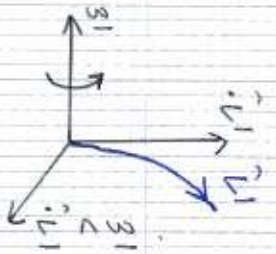


Cross-section of the rotating Earth indicating the centrifugal force acting at a number of points on the surface (N.B. the force is proportional to r the distance from the axis of rotation.)

The centrifugal force means you feel lighter at the equator than at the poles, about 0.3% lighter.

The Coriolis force, $\underline{F}_{coriolis} = -2m(\underline{\omega} \wedge \underline{\dot{r}}')$, causes a clockwise deflection at the north pole and an anticlockwise deflection at the south pole:

NORTH POLE:



(horizontal trajectory).

One simply imagines the Earth rotating under the projectile at the north pole. At the south pole the direction of $\underline{\omega}$ is reversed and the force causes an anticlockwise deviation.

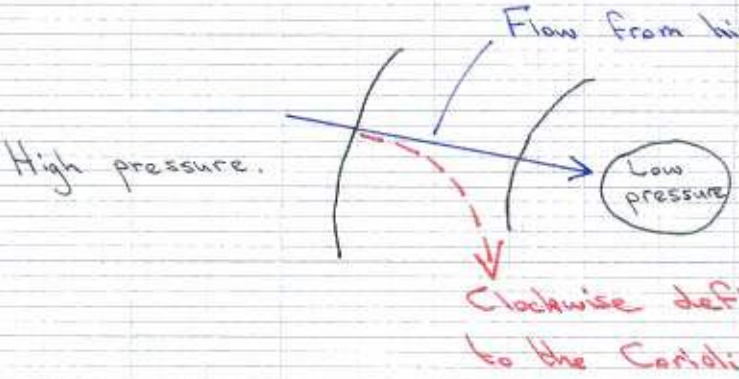
If one travels East the Coriolis force counteracts gravity a small amount and makes you feel lighter, if you go West you feel heavier. This is the Eötvös effect

The Coriolis effect causes the circulation of waters and air on the Earth. Despite the effect causing particles to be deflected in a clockwise direction in the northern hemisphere, the Coriolis effect is responsible for cyclones which rotate in an anticlockwise direction in the northern hemisphere. In the southern hemisphere the anticlockwise deflection causes clockwise cyclones to form.

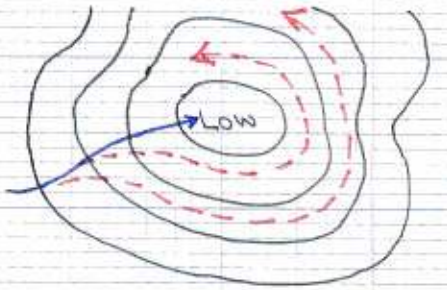
Why is this?

Consider a northern hemisphere example:

Wind is the flow of air from regions of high pressure to low pressure, indicated by isobars (lines of constant, equal pressure):



As the air molecules move from high to low pressure regions they experience a force rotating them in a clockwise direction in the northern hemisphere. As the molecules' paths turn the Coriolis force begins to be balanced by the force causing the flow from high to low pressure regions, until the force is exactly balanced and the air flows in a region between the isobars (or parallel to the isobars in otherwords). If you turn to the right at a roundabout you go around in an anticlockwise direction! Similarly the Coriolis force causes the wind to circulate in an anticlockwise direction in the northern hemisphere.



The Conjugate Momentum.

We are now in a position to define the conjugate momentum, p_i for each generalised coordinate.

Consider the Lagrangian for a free particle in Cartesian coordinates:

$$L = \tfrac{1}{2} m (\dot{x}^2 + \dot{y}^2 + \dot{z}^2).$$

If we calculate the Lagrange equation for x:

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

We begin by finding $\frac{\partial L}{\partial \dot{x}} = m\dot{x} \equiv p_x$ which is just the momentum in the x-direction. This is where the following definition of generalised or conjugate* momenta originates. We define the conjugate momentum p_i associated to a coordinate q_i as:

$$p_i \equiv \frac{\partial L}{\partial \dot{q}_i}$$

The Lagrange equations can now be written as:

$$\frac{d}{dt}(p_i) - \frac{\partial L}{\partial q_i} = 0$$

$$\text{Or,} \qquad \dot{p}_i = \frac{\partial L}{\partial q_i}$$

We will make use of this notation when we formulate the Hamiltonian approach to dynamics.

The Energy Function and the Conservation of Energy.

The Lagrangian is a function of q_i, $\dot{q}_i$ and perhaps t. What is the total time derivative of L?

Recall that,

$$\frac{dL}{dt} = \sum_j \frac{\partial L}{\partial q_j}\frac{\partial q_j}{\partial t} + \sum_j \frac{\partial L}{\partial \dot{q}_j}\frac{\partial \dot{q}_j}{\partial t} + \frac{\partial L}{\partial t}.$$

From Lagrange's equations,

$$\frac{\partial L}{\partial q_j} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right)$$

So we can write,

$$\frac{dL}{dt} = \sum_j \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right)\frac{\partial q_j}{\partial t} + \sum_j \frac{\partial L}{\partial \dot{q}_j}\frac{\partial \dot{q}_j}{\partial t} + \frac{\partial L}{\partial t}.$$

* Sometimes even canonical momenta.

$$\therefore \quad \frac{dL}{dt} = \sum_j \frac{d}{dt}\left(\dot{q}_j \frac{\partial L}{\partial \dot{q}_j}\right) + \frac{\partial L}{\partial t}.$$

So that, $$\frac{d}{dt}\left(\sum_j \dot{q}_j \frac{\partial L}{\partial \dot{q}_j} - L\right) + \frac{\partial L}{\partial t} = 0.$$

The last term is zero if L is not an explicit function of t. When this is the case the equation above is a statement saying that energy is conserved. The object in brackets is called the energy function, and denoted with an uncapitalised h:

$$h(q_i, \dot{q}_i; t) \equiv \sum_j \dot{q}_j \frac{\partial L}{\partial \dot{q}_j} - L \qquad \text{The energy function.}$$

N.B. L is explicitly independent of t if it is written in terms of q_i, $\dot{q}_i$ only and not t. Even though q_i, $\dot{q}_i$ are time dependent $\partial L/\partial t = 0$ in this case; remember that all other variables are treated as constants in a partial derivative of a given variable.

Note that in the absence of a parameter (t) a partial derivative of L ($\partial L/\partial t$) vanishes and we find a conserved quantity. This is true more generally and is a part of Noether's theorem.

Noether's Theorem.

We have commented that the conservation of energy is a consequence of a Lagrangian manifestly independent of t (i.e. such that $\frac{\partial L}{\partial t} = 0$). From Lagrange's equations we have already seen a number of conserved quantities when L is independent of a generalised coordinate, q_i. Such a coordinate is called ignorable.

E.g. Lagrange's equation for q_i: $\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0.$

If q_i is an ignorable quantity then $\partial L/\partial q_i = 0$ and:

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

When a coordinate is ignorable its conjugate momentum is a conserved quantity of the motion.

The absence of a coordinate gives rise to a symmetry of the Lagrangian, namely the Lagrangian is unchanged by a variation of this ignorable coordinate. Noether's theorem relates a conserved quantity to every continuous symmetry of a Lagrangian. To understand the theorem we must define what we mean by a continuous symmetry of a Lagrangian.

A transformation of the coordinates:

$$q_i(t) \longrightarrow \hat{q}_i(s,t) \qquad s \in \mathbb{R}$$

such that $\hat{q}_i(0,t) = q_i(t)$, is a continuous symmetry of the Lagrangian if:

$$\frac{\partial}{\partial s}\left(L\left(\hat{q}_i(s,t), \dot{\hat{q}}_i(s,t), t \right)\right) = 0.$$

In other words if we deform the coordinates and velocities away from their originals (and we parameterise this deformation by the continuous variable s) then if the deformed Lagrangian is independent of s the coordinate transformation is a continuous symmetry of L.

Noether's theorem asserts that for each continuous symmetry of L there is a conserved quantity.

Proof.

$$\frac{\partial L}{\partial s} = \sum_i \frac{\partial L}{\partial \hat{q}_i}\frac{\partial \hat{q}_i}{\partial s} + \sum_i \frac{\partial L}{\partial \dot{\hat{q}}_i}\frac{\partial \dot{\hat{q}}_i}{\partial s}$$

Suppose the continuous symmetry exists and $\frac{\partial L}{\partial s} = 0$, then it is zero also when it is evaluated at $s = 0$:

$$\begin{aligned}
0 &= \left.\frac{\partial L}{\partial s}\right|_{s=0} \\
&= \sum_i \left.\frac{\partial L}{\partial \hat{q}_i}\frac{\partial \hat{q}_i}{\partial s}\right|_{s=0} + \sum_i \left.\frac{\partial L}{\partial \dot{q}_i}\frac{\partial \dot{\hat{q}}_i}{\partial s}\right|_{s=0} \\
&= \sum_i \left(\frac{\partial L}{\partial q_i}\right)\left.\left(\frac{\partial \hat{q}_i}{\partial s}\right)\right|_{s=0} + \sum_i \frac{\partial L}{\partial \dot{q}_i}\left.\left(\frac{\partial \dot{\hat{q}}_i}{\partial s}\right)\right|_{s=0} \\
&= \sum_i \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right)\left.\left(\frac{\partial \hat{q}_i}{\partial s}\right)\right|_{s=0} + \sum_i \frac{\partial L}{\partial \dot{q}_i}\left.\left(\frac{\partial \dot{\hat{q}}_i}{\partial s}\right)\right|_{s=0} \\
&= \frac{d}{dt}\left(\sum_i \frac{\partial L}{\partial \dot{q}_i}\left.\left(\frac{\partial \hat{q}_i}{\partial s}\right)\right|_{s=0}\right)
\end{aligned}$$

This is the conserved quantity associated to the continuous symmetry.

Let us now look at how some common symmetries lead to familiar conserved quantities.

The homogeneity of space.

Let us suppose that the laws of dynamics do not vary from point to point in space, or, in other words, if we translated our coordinate system we would expect Lagrange's equations to be unaltered. This gives rise to the translation:

$$\underline{r}_i \to \underline{r}_i + s\underline{a} \equiv \hat{\underline{r}}_i$$

Where $\underline{a}$ is a constant vector.

Noether's theorem states that the associated conserved quantity is:

$$\sum_i \frac{\partial L}{\partial \dot{q}_i}\left(\frac{\partial \hat{q}_i}{\partial s}\right)\Bigg|_{s=0}$$

As $\frac{\partial L}{\partial \dot{q}_i} \equiv p_i$, the conjugate momentum, and in this example

$$\left(\frac{\partial \hat{\underline{r}}_i}{\partial s}\right)\Bigg|_{s=0} = \underline{a}.$$

Then the conserved quantity is $\sum_i \underline{p}_i \cdot \underline{a}$. This is the total linear momentum in the direction $\underline{a}$. Since it is true for all $\underline{a}$ we conclude that total momentum is conserved.

> Homogeneity of space (translation invariance of L) $\Rightarrow$ Conservation of total linear momentum.

The isotropy of space.

Suppose that the Lagrangian is invariant under a rotation. The coordinate transformation is infinitesimally:

$$\underline{r}_i \to \underline{r}_i + s(\underline{a} \wedge \underline{r}_i). \equiv \hat{\underline{r}}_i$$

Where $\underline{a}$ is constant. We saw in the example of the Coriolis effect that this corresponded to a rotation about $\underline{a}$. Let's remind ourselves:

$$\underline{a} \wedge \underline{r}_i = \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix} \wedge \begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix} = \begin{pmatrix} a_y z_i - y_i a_z \\ a_z x_i - z_i a_x \\ a_x y_i - x_i a_y \end{pmatrix}$$

Restrict to a simple example a rotation of θ about the z-axis.

In this case $\underline{a} = \begin{pmatrix} 0. \\ 0. \\ \theta. \end{pmatrix}$ and $\underline{a} \wedge \underline{r}_i = \begin{pmatrix} -y_i \theta. \\ x_i \theta. \\ 0. \end{pmatrix}$

Then $\hat{\underline{r}}_i = \underline{r}_i + s(\underline{a} \wedge \underline{r}_i) = \begin{pmatrix} x_i - s y_i \theta. \\ y_i + s x_i \theta. \\ z_i. \end{pmatrix}$

If we had rotated the xy-plane an angle $s\theta$ about the z-axis we would find:

$$\hat{x}_i = x_i \cos(s\theta) - y_i \sin(s\theta).$$

$$\hat{y}_i = x_i \sin(s\theta) + y_i \cos(s\theta).$$

When $s\theta$ is small, $\cos(s\theta) \simeq 1$, $\sin(s\theta) \simeq s\theta$, and we have:

$$\hat{x}_i \simeq x_i - y_i s\theta$$

$$\hat{y}_i \simeq y_i + s x_i \theta.$$

Hence the map $\underline{r}_i \to \underline{r}_i + s(\underline{a} \wedge \underline{r}_i)$ corresponds to an infinitesimal rotation about the z-axis. For general $\underline{a}$ it is an infinitesimal rotation about $\underline{a}$.

But what is the associated conserved quantity?

We have to calculate:

$$\sum_i \frac{\partial L}{\partial \dot{q}_i} \cdot \left(\frac{\partial \hat{q}_i}{\partial s} \right)\Big|_{s=0} = \sum_i \underline{p}_i \cdot (\underline{a} \wedge \underline{r}_i)$$

$$= \sum_i \underline{a} \cdot (\underline{r}_i \wedge \underline{p}_i)$$

This is the component of the total angular momentum in the direction $\underline{a}$. As $\underline{a}$ can be any direction (if space is isotropic) then we find:

Isotropy of space. (Rotational invariance of L)	$\Rightarrow$ Conservation of total angular momentum.

Homogeneity of time.

In fact we have seen this one already. The transformation is:

$$t \to t + s a.$$

But we only considered continuous transformations of q_i and $\dot{q}_i$

When we proved Noether's theorem earlier. If the time translation is a continuous symmetry of L then:

$$\frac{\partial}{\partial s}\left(L(q_i, \dot{q}_i; t+sa)\right) = 0.$$

In this case,

$$\frac{\partial L}{\partial s} = \sum_i \frac{\partial L}{\partial q_i}\frac{\partial q_i}{\partial s} + \sum_i \frac{\partial L}{\partial \dot{q}_i}\frac{\partial \dot{q}_i}{\partial s} + \frac{\partial L}{\partial t}\frac{\partial t}{\partial s}.$$

($\frac{\partial q_i}{\partial s} \to 0$, $\frac{\partial \dot{q}_i}{\partial s} \to 0$)

$$= \left(\frac{\partial L}{\partial t}\right) a.$$

$$= 0.$$

If time translation is a symmetry then $\frac{\partial L}{\partial t} = 0$ and the Lagrangian is not an explicit function of time. We saw earlier that when this is true the energy function, h, is conserved:

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

You can convince yourself that this is the energy (in brackets above) by substituting $p_i \equiv m_i \dot{q}_i$ and $L = T - V$ then,

$$h \equiv \sum \dot{q}_i p_i - L$$

$$= \sum_i m \dot{q}_i^2 - T + V$$

$$= 2T - T + V$$

$$= T + V.$$

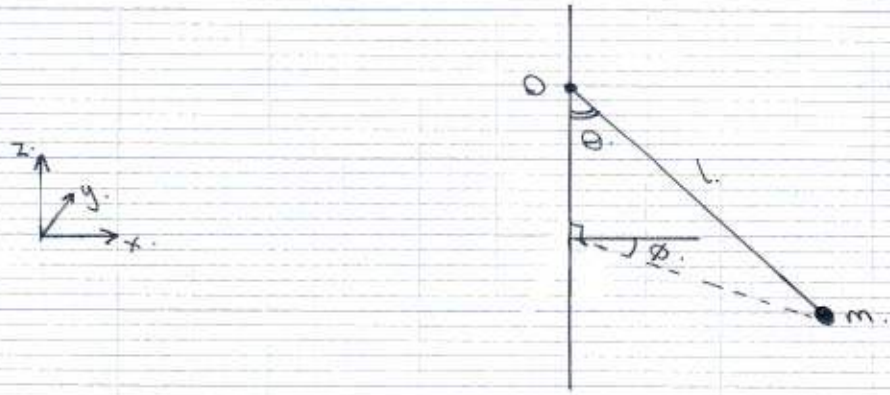
Therefore,

> Homogeneity of time $\Rightarrow$ Conservation of energy.
> $(\partial L / \partial t = 0)$

Let us now look at the motion of a spherical pendulum and use Noether's theorem to help analyse the dynamics. Note that the motion of a spherical pendulum is identical to that of the marble rolling in a bowl/hemisphere that we considered in example 5. However we will use a different set of coordinates to describe the pendulum so that the equations of motion will look a little different.

Example 7: The Spherical Pendulum.

The spherical pendulum is not constrained to move in the plane but may move in 3D space:



$0 \leq \theta < \pi$.

$0 \leq \phi < 2\pi$.

There is one holonomic constraint: the length of the rigid rod is l. Hence there are 2 generalised coordinates, which we choose to be θ and ϕ as indicated in the diagram. The Cartesian coordinates of the bob are:

$$x = l\sin\theta\cos\phi.$$

$$y = l\sin\theta\sin\phi.$$

$$z = -l\cos\theta.$$

$$\therefore \quad \dot{x} = l\cos\theta\cos\phi\,\dot{\theta} - l\sin\theta\sin\phi\,\dot{\phi}$$

$$\dot{y} = l\cos\theta\sin\phi\,\dot{\theta} + l\sin\theta\cos\phi\,\dot{\phi}.$$

$$\dot{z} = l\sin\theta\,\dot{\theta}.$$

Using $V = +mgz$ we have:

$$L = \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - mgz.$$

$$= \tfrac{1}{2}m\left(l^2\cos^2\theta\cos^2\phi\,\dot{\theta}^2 - 2l^2\cos\theta\sin\theta\cos\phi\sin\phi\,\dot{\theta}\dot{\phi} + l^2\sin^2\theta\sin^2\phi\,\dot{\phi}^2\right.$$
$$+ l^2\cos^2\theta\sin^2\phi\,\dot{\theta}^2 + 2l^2\cos\theta\sin\theta\sin\phi\cos\phi\,\dot{\theta}\dot{\phi} + l^2\sin^2\theta\cos^2\phi\,\dot{\phi}^2$$
$$\left. + l^2\sin^2\theta\,\dot{\theta}^2\right) + mgl\cos\theta.$$

$$= \tfrac{1}{2}m\left(l^2\cos^2\theta\,\dot{\theta}^2 + l^2\sin^2\theta\,\dot{\theta}^2 + l^2\sin^2\theta\,\dot{\phi}^2\right) + mgl\cos\theta.$$

$$= \tfrac{1}{2}ml^2\left(\dot{\theta}^2 + \sin^2\theta\,\dot{\phi}^2\right) + mgl\cos\theta.$$

Notice that ϕ does not appear explicitly in the Lagrangian, and neither does $\dot{\phi}$. These coordinates are "ignorable". By Noether's theorem we know that (1.) $\partial L/\partial\dot{\phi} \equiv$ (The component of angular momentum ($\underline{L}$) in the $\hat{\underline{\phi}}$ direction) is conserved, and (2.) As $\partial L/\partial t = 0$, energy is conserved.

From the Lagrange equation for ϕ:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\phi}}\right) = 0.$$

And, $\frac{\partial L}{\partial \dot{\phi}} = m l^2 \sin^2\theta . \dot{\phi} \equiv J$ is a constant.

The equation of motion for θ gives:

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

$$\therefore \frac{d}{dt}\left(m l^2 \dot{\theta}\right) - m l^2 \sin\theta . \cos\theta . \dot{\phi}^2 + mgl\sin\theta = 0.$$

$$\therefore m l^2 \ddot{\theta} - m l^2 \sin\theta . \cos\theta . \dot{\phi}^2 + mgl\sin\theta = 0.$$

We can use our constant J to find an equation in terms of only θ and its derivatives:

$$\dot{\phi}^2 = J^2 \Big/ m^2 l^4 \sin^4\theta .$$

$$\therefore m l^2 \ddot{\theta} - \left(\cos\theta . J^2\right) \Big/ \left(m l^2 \sin^3\theta .\right) + mgl\sin\theta = 0.$$

$$\therefore \ddot{\theta} = J^2 . \cos\theta \Big/ m^2 l^4 \sin^3\theta . - \left(g/l\right)\sin\theta .$$

This can be integrated with respect to theta, θ, and if we define $\ddot{\theta} \equiv -\partial V_{eff}/\partial\theta$, where V_{eff} is the effective potential we find:

$$V_{eff} = J^2 \Big/ 2 m^2 l^4 \sin^2\theta . - \left(g/l\right)\cos\theta .$$

On top of J being conserved we also have that energy is conserved. We can use the effective potential to write out the energy function.

$$h \equiv \sum \dot{q}_i p_i - L$$

$$= \dot{\theta}\left(\frac{\partial L}{\partial \dot{\theta}}\right) + \dot{\phi}\left(\frac{\partial L}{\partial \dot{\phi}}\right) - L.$$

$$= \tfrac{1}{2} m l^2 \left(\dot{\theta}^2 + \sin^2\theta . \dot{\phi}^2\right) - mgl\cos\theta .$$

$$= \tfrac{1}{2} m l^2 \dot{\theta}^2 + \tfrac{1}{2} m l^2 \sin^2\theta \left(J^2 \Big/ m^2 l^4 \sin^4\theta\right) - mgl\cos\theta .$$

$$= \tfrac{1}{2} m l^2 \dot{\theta}^2 + m l^2 \left(V_{eff}\right).$$

N.B. We defined V_{eff} in terms of $\dot{\theta}$ without the constants m and l hence we have h written as above instead of $h = T + V_{eff}$., as m and l^2 are constants that also premultiply $\tfrac{1}{2}\dot{\theta}^2$ it makes no difference to have defined V_{eff} in this way. It may even simplify our expressions.

We can rewrite this equation as:

$$\frac{d\theta}{dt} \equiv \dot{\theta} = \left(\frac{2h}{ml^2} - 2V_{eff}(\theta) \right)^{1/2}.$$

or (collecting the terms dependent on θ) and integrating:

$$\int_{\theta_1}^{\theta_2} d\theta \Big/ \left(\frac{2h}{ml^2} - 2V_{eff}(\theta) \right)^{1/2} = \int_{t_1}^{t_2} dt. \quad = t_2 - t_1.$$

When one can write the solution in this form it is called "reducing the problem to quadrature". It is clear that it is still not easy to solve the left-hand-side of the equation above. If we could find a solution for $\theta(t)$ we would also have $\phi(t)$, using $\dot{\phi}^2 = J^2/m^2l^4 \sin^4\theta$:

$$\Rightarrow \int d\phi = \int \left(J/ml^2 \sin^2\theta \right) dt$$

$$= \int \left(J/ml^2 \sin^2\theta \right) \left(\frac{d\theta}{dt} \right)^{-1} d\theta.$$

$$= J/ml^2 \int d\theta \Big/ \left[\left(\frac{2h}{ml^2} - 2V_{eff} \right)^{1/2} \sin^2\theta \right].$$

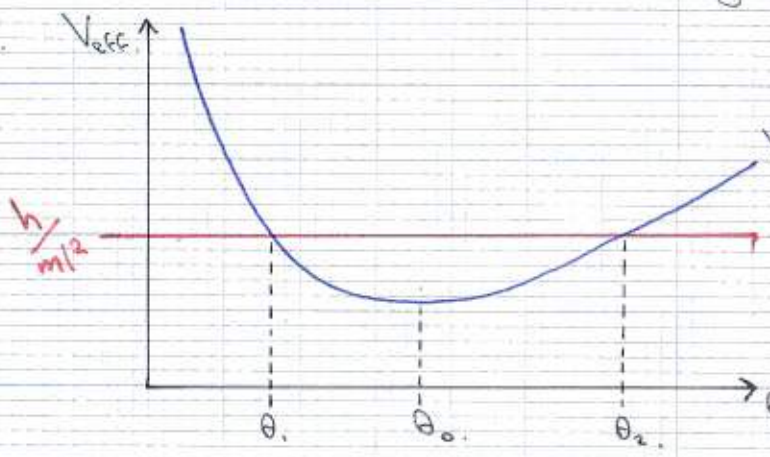
However we can use our expression for h to make some qualitative remarks about the motion. Using,

$$h/ml^2 = \tfrac{1}{2}\dot{\theta}^2 + V_{eff}.$$

And observing that $\tfrac{1}{2}\dot{\theta}^2$ is a positive definite term we have that

$$h/ml^2 \geq V_{eff}.$$

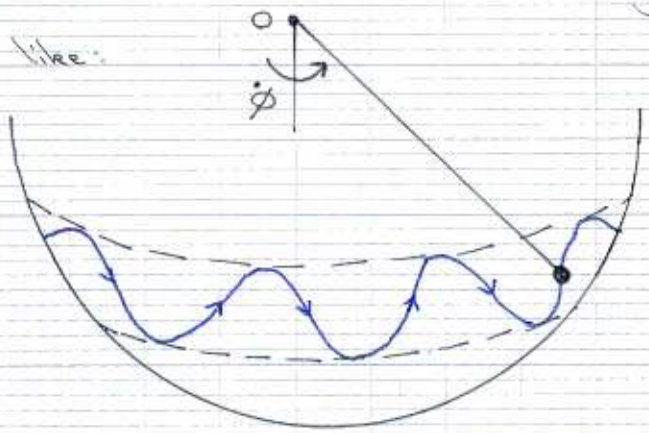
Now we can plot the effective potential using a computer on a graph to find:



N.B. $V_{eff} \to \infty$ as $\theta \to 0$ as $J \neq 0$ and angular momentum is conserved. For similar reasons $V_{eff} \to \infty$ also when $\theta \to \pi$. But this is not plotted.

The region in which $V_{eff} \leq h/m l^2$ is bounded by two angles θ_1 and θ_2 which lie somewhere between $\theta = 0$ and $\theta = \pi$.

The motion is bounded between these two angles and so we suppose it must look like:



3. Central Forces.

In this section we will focus on the dynamics resulting from central forces (i.e. $\underline{F} \propto \underline{r}$). Both Newton's gravitational force and Coulomb's force between charged particles are central forces. We will qualitatively analyse their dynamics, and prove Kepler's three laws.

The 2-Body Problem Becomes the 1-Body Problem.

Consider a system consisting of two bodies of masses m_1 and m_2 respectively whose position vectors are $\underline{r}_1$ and $\underline{r}_2$. We will consider these particles to be free with the exception that a central force acts between them. In this case the central force of interaction is proportional to the vector between particles 1 and 2; it is called central because the 2 body problem may be reduced to that of 1 body subject to a central force.

Before proceeding let us recall that a central force is conservative and preserves angular momentum:

(i.) CONSERVATIVE: A central force $\underline{F} \propto \underline{r}$, let us write $\underline{F} = k\underline{r}$ for some constant k. It is not a function of $\dot{\underline{r}}$ nor t and $\nabla \wedge \underline{F} = \underline{0}$

N.B. $$\begin{pmatrix} \partial/\partial x \\ \partial/\partial y \\ \partial/\partial z \end{pmatrix} \wedge \underline{F} = k \begin{pmatrix} \partial z/\partial y - \partial y/\partial z \\ \partial x/\partial z - \partial z/\partial x \\ \partial y/\partial x - \partial x/\partial y \end{pmatrix} = \underline{0}.$$

Hence $\underline{F}$ is conservative.

More generally a central force is derived from a potential which is a function of r only: $\underline{F} = -\nabla V(r)$, where $r = (x^2+y^2+z^2)^{1/2}$.

We can check this using the chain rule:

$$F_x = -\frac{\partial}{\partial x} V(r) = -\frac{\partial V}{\partial r} \cdot \frac{\partial r}{\partial x} = -V' x/r.$$

Similarly, $F_y = -V' y/r$ & $F_z = -V' z/r$.

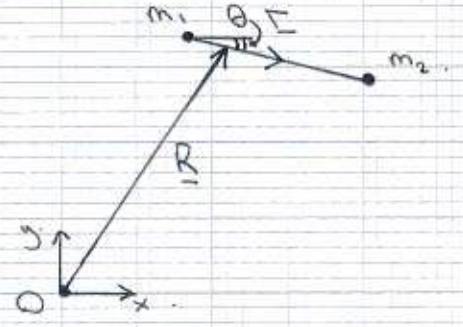
$\therefore \underline{F} = -\frac{V'}{r}\underline{r}$, and is central.

(ii) Conserves angular momentum:

If $\underline{F} = R\underline{r}$ then torque $\underline{N} \equiv \underline{r} \wedge \underline{F} = R(\underline{r} \wedge \underline{r}) = \underline{0}$

We recall that $\underline{N} = \dot{\underline{L}}$, therefore when $\underline{F}$ is central $\dot{\underline{L}} = \underline{0}$ and $\underline{L}$ is conserved.

Returning to our scenario we have:



In the sketch we have indicated the vector $\underline{r} \equiv \underline{r}_2 - \underline{r}_1$

and the centre of mass vector $\underline{R} \equiv \frac{m_1}{M}\underline{r}_1 + \frac{m_2}{M}\underline{r}_2$.

If we label the angle between $\underline{r}$ and the x-axis by θ we can write the coordinates for the particles in terms of $r \equiv |\underline{r}|$, θ and $\underline{R} \equiv \begin{pmatrix} X \\ Y \end{pmatrix}$. We can use:

$MX = m_1 x_1 + m_2 x_2$ (recall that $M \equiv m_1 + m_2$).

$MY = m_1 y_1 + m_2 y_2$

$x_2 - x_1 = r\cos\theta$.

$y_2 - y_1 = r\sin\theta$.

Rearranging these expressions gives:

$x_1 = X - \frac{m_2}{M} r\cos\theta$. $\qquad x_2 = X + \frac{m_1}{M} r\cos\theta$.

$y_1 = Y - \frac{m_2}{M} r\sin\theta$. $\qquad y_2 = Y + \frac{m_1}{M} r\sin\theta$.

As usual we can use these to find the kinetic energy:

$$T = \tfrac{1}{2}m_1(\dot{x}_1^2 + \dot{y}_1^2) + \tfrac{1}{2}m_2(\dot{x}_2^2 + \dot{y}_2^2)$$

$$= \tfrac{1}{2}M(\dot{X}^2 + \dot{Y}^2) + \tfrac{\mu}{2}(\dot{r}^2 + r^2\dot{\theta}^2)$$

$$= \tfrac{1}{2}M\dot{\underline{R}}^2 + \tfrac{\mu}{2}\dot{\underline{r}}^2.$$

Where $\mu \equiv \frac{m_1 m_2}{m_1 + m_2}$ is called the reduced mass.

The potential energy is a function of r only for a conservative central force. We may write the Lagrangian for 2 bodies interacting through a central force as:

$$L = T - V$$
$$= \tfrac{1}{2} M \dot{\underline{R}}^2 + \tfrac{\mu}{2} \dot{\underline{r}}^2 - V(r).$$

We indicated the derivation of the Lagrangian for the two-dimensional case but it takes the same form for three dimensions.

Our first observation is that $\underline{R}$ is an ignorable coordinate - there is a conserved quantity associated with this fact, namely, the rate of change of $M\dot{\underline{R}}$ (the momentum of the two-body system). From Lagrange's equation for $\underline{R}$:

$$\frac{d}{dt}\left(M\dot{\underline{R}}\right) = \underline{0}.$$

The centre of mass either moves at constant velocity or is at rest.

The equation of motion for $\underline{r}$ is:

$$\mu\ddot{\underline{r}} + \frac{1}{r} V'(r)\,\underline{r} = \underline{0}.$$

Which is independent of $\underline{R}$ and $\dot{\underline{R}}$. Consequently we may eliminate $\underline{R}$ and $\dot{\underline{R}}$ from our considerations of the two-body problem: we ignore the first term in the Lagrangian herein and the two-body interaction problem has been reduced to that of a single particle at r subject to a central force whose mass is $\mu = m_1 m_2 / (m_1 + m_2)$, the reduced mass.

Kepler's Laws of Planetary Motion.

"When a planet, or a comet, or any other body arcs through space under the influence of gravity it traces out one of a very special set of mathematical curves - either a circle or an ellipse or a parabola or a hyperbola.... Why in the world does nature choose to trace out in the sky those, and only those, elegant geometrical constructions?" David L. Goodstein & Judith R. Goodstein From the introduction to Feynman's Lost Lecture.

Gravity is an example of a central force. We know this thanks to Newton. But before Newton, Johannes Kepler (1571-1630) codified the patterns in astronomical data into three laws of planetary motion. They are:

KI. The planets travel along elliptic paths with the Sun at one focus.

KII. The line joining a planet to the Sun sweeps out equal areas in equal times.

KIII. The square of a planet's orbital period is proportional to the cube of its distance from the Sun.

As we continue our investigation of central forces we will also prove these three laws for the special case of gravity.

The Equations of Motion

Recall the Lagrangian for the two-body problem with a conservative, central force:

$$L = \frac{1}{2} M \dot{\underline{R}}^2 + \frac{\mu}{2}\left(\dot{r}^2 + r^2\dot{\theta}^2\right) - V(r)$$

The equation of motion for the centre of mass, $\underline{R}$, is trivial as $\underline{R}$ is ignorable and the problem is essentially that of one body moving under a central force (centred at the second body).

θ is also an ignorable coordinate and it's associated conserved quantity is angular momentum, $|\underline{L}| = p_\theta$:

$$p_\theta = \frac{\partial L}{\partial \dot{\theta}} = \mu r^2 \dot{\theta}.$$

The equation of motion for θ is:

$$\dot{p}_\theta = 0.$$

Integrating this gives a constant of the system which we denote L:

$$\mu r^2 \dot{\theta} = L.$$

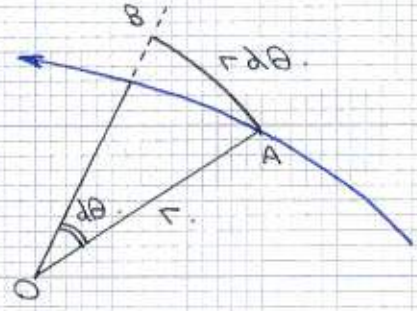
Since,

$$\frac{d}{dt}\left(\mu r^2 \dot{\theta}\right) = 0$$

then it is also true that: (as μ is a constant).

$$\frac{d}{dt}\left(\tfrac{1}{2} r^2 \dot{\theta}\right) = 0.$$

Note that $\frac{1}{2} r^2 \dot{\theta}$ is the area swept out by the radius vector in unit time.



As $d\theta \to 0$ the length $r d\theta$ approaches the triangle height.

The infinitesimal area of the triangle, OAB, is:

$$d(\text{Area}) = \tfrac{1}{2} r (r d\theta).$$

$$\therefore \frac{d(\text{Area})}{dt} = \tfrac{1}{2} r^2 \frac{d\theta}{dt} = \tfrac{1}{2} r^2 \dot{\theta}.$$

The conservation of angular momentum leads to the rate of change of the area swept out by $\underline{r}$ being a constant. This is Kepler's second law: equal areas are swept out in equal times. We have not specialised to the inverse-square law of gravity yet so that the essence of KII is true for general central forces.

Since $\underline{L}$ is conserved we can also deduce that the motion is confined to a plane: $\underline{L} = \underline{r} \wedge \underline{p} \Rightarrow \underline{r} \cdot \underline{L} = 0$

As $\underline{r}$ is perpendicular to $\underline{L}$, which does not change, $\underline{r}$ is confined to a plane.

The equation of motion for r is more complicated:

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

$$\therefore \frac{d}{dt}\left(\mu \dot{r}\right) - \mu r \dot{\theta}^2 + \frac{\partial V}{\partial r} = 0.$$

$$\therefore \mu \ddot{r} = \mu r \dot{\theta}^2 - \frac{\partial V}{\partial r}.$$

As we did for the spherical pendulum, we can use the conservation of

angular momentum to write this as an equation in terms of r only. Recall that $L \equiv \mu r^2 \dot{\theta}$ is a constant of the motion. We may substitute $\dot{\theta} = L/\mu r^2$ into the equation of motion for r and eliminate $\dot{\theta}$:

$$\mu \ddot{r} = \frac{L^2}{\mu r^3} - \frac{\partial V}{\partial r} \qquad -(*)$$

The Lagrangian is not an explicit function of time, t, therefore the energy function:

$$h \equiv \sum \dot{q}_i p_i - L$$

is a constant of the motion. Explicitly,

$$h = \dot{\underline{R}} \cdot (M \dot{\underline{R}}) + \dot{r}(\mu \dot{r}) + \dot{\theta}(\mu r^2 \dot{\theta}) - L$$
$$= \tfrac{1}{2} M \dot{\underline{R}}^2 + \tfrac{\mu}{2}(\dot{r}^2 + r^2 \dot{\theta}^2) + V(r).$$

Since $M\dot{\underline{R}}$ is also a constant of the motion then $E \equiv h - \frac{1}{2} M \dot{\underline{R}}^2$ is also a constant and we have:

$$E = \tfrac{\mu}{2}(\dot{r}^2 + r^2 \dot{\theta}^2) + V(r).$$

Eliminating $\dot{\theta}$ we have:

$$E = \tfrac{\mu}{2}(\dot{r}^2) + \frac{L^2}{2\mu r^2} + V(r). \qquad -(**)$$

As we did for the spherical pendulum we can find an effective potential such that $\mu \ddot{r} = -\frac{\partial V_{eff}}{\partial r}$. We can deduce V_{eff} from either (*) or (**) above. Using (*) we find:

$$V_{eff} = \frac{L^2}{2\mu r^2} + V.$$

[Check: $-\frac{\partial V_{eff}}{\partial r} = -\left(-\frac{L^2}{\mu r^3} + \frac{\partial V}{\partial r}\right) = \mu \ddot{r}$].

And consequently,

$$E = \tfrac{\mu}{2} \dot{r}^2 + V_{eff}.$$

Noting that the first term on the right is positive or zero we deduce that:

$$V_{eff} \leq E.$$

Throughout the motion of the system.

Classification of Orbits.

Our aim in this section is to analyse the possible classes of motion qualitatively by finding the range of r for which

$$V_{eff}(r) \leq E.$$

To do so we will consider the example of the force of gravity but given a central force with a power law behaviour not that of gravity the same analysis will remain valid.

For gravity the force law is proportional to the inverse square of r:

$$F = -\frac{Gm_1m_2}{r^2} = -\frac{k}{r^2}.$$

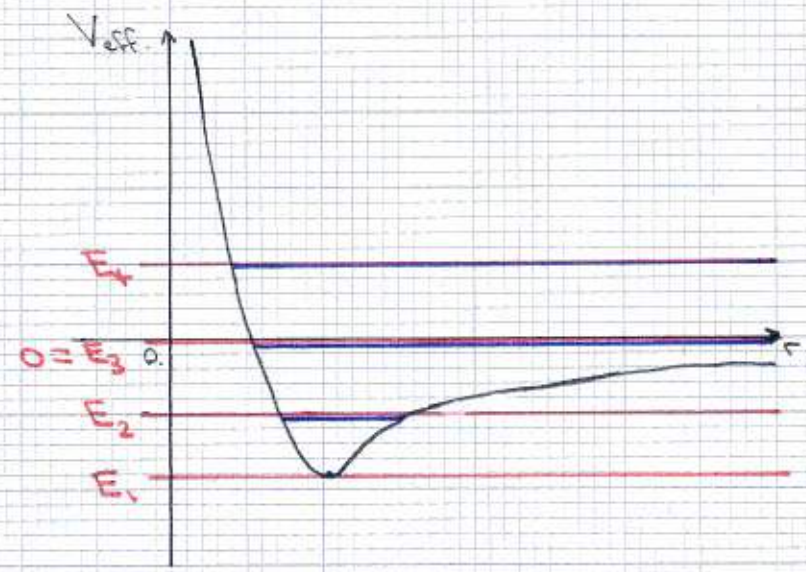
The force is derived from a potential V:

$$V = -\frac{k}{r}.$$

The effective potential becomes:

$$V_{eff} = \frac{l^2}{2\mu r^2} - k/r.$$

Let us plot this curve with respect to r:



In red are indicated four energies and in blue the range of r for which $E \geq V_{eff}$ for each energy.

By giving an energy we can now say the range of r for which $V_{eff} \leq E$. Let us consider the four main examples of orbits:

(I.) $E_1 = (V_{eff})_{minimum}$, the minimum of the effective potential.

For this energy the range of r is confined to one point, r_0.

Since r_0 is constant the motion is circular.

(N.B. $\frac{dV_{eff}}{dr} = -\frac{l^2}{\mu r^3} + \frac{k}{r^2}$, $\frac{dV_{eff}}{dr} = 0$ at $r = r_0 \Rightarrow r_0 = \frac{l^2}{\mu k}$.)

(II.) $(V_{eff})_{minimum} < E_2 < 0$.

Here the motion is bounded between two values of r: $r_1 \leq r \leq r_2$.

The radial distance is never less than r_1 and never greater than r_2 over the motion.

(III.) $E_3 = 0$.

The radius is bounded below but not above.

(IV.) $E_4 > 0$

Again the radius has a lower bound but not an upper bound.

In all cases there is a lower bound on r and r can never reach $r=0$. This is due to the $l^2/2\mu r^2$ term in the effective potential, as in the spherical pendulum it prohibits motion to $r=0$ when $l \neq 0$. It is called the angular momentum barrier and despite there being no associated real repulsive force the motion behaves as if there is one. It is a remarkable phenomenon and is the reason the Earth does not spiral rapidly into the Sun.

Motion Under Gravity.

In the previous section we have given qualitative descriptions of the orbits due to gravity in a two-body system. In this section we will find exact solutions, up to boundary conditions, and show that the planetary orbits are conic sections.

It is possible to solve the differential equation of motion for r directly following a substitution of $V = k/r$. However we will make use of a neat trick relevant only to inverse-square law central forces. In order to see why this trick is valid only for $1/r^2$ forces we will begin with a general central force:

$$\underline{F} = f(r)\frac{\underline{r}}{r} \qquad \text{where} \quad r = |\underline{r}|.$$

In this case by VII:

$$\dot{p} = \frac{r}{r} f(r)\, \mathbf{r}.$$

Writing $p = m\dot{\mathbf{r}}$ and $\mathbf{L} = \mathbf{r} \times p = m(\mathbf{r} \times \dot{\mathbf{r}})$ gives:

$$\dot{p} \times \mathbf{L} = \frac{m}{r^3} f(r) \left[\mathbf{r} \times (\mathbf{r} \times \dot{\mathbf{r}}) \right].$$

$$= \frac{m}{r^3} f(r) \left[\mathbf{r} (\mathbf{r} \cdot \dot{\mathbf{r}}) - \dot{\mathbf{r}} (\mathbf{r} \cdot \mathbf{r}) \right].$$

$$= \frac{m}{r^3} f(r) \left[\mathbf{r} (\mathbf{r} \cdot \dot{\mathbf{r}}) - \dot{\mathbf{r}}\, r^2 \right].$$

Notice that $2\, \mathbf{r} \cdot \dot{\mathbf{r}} = \frac{d}{dt}(\mathbf{r} \cdot \mathbf{r}) = \frac{d}{dt}(r^2)$.

$$\therefore \mathbf{r} \cdot \dot{\mathbf{r}} = \frac{1}{2} \frac{d}{dt}(r^2) = r \dot{r}.$$

$$\therefore \dot{p} \times \mathbf{L} = \frac{m}{r^3} f(r) \left[(r \dot{r})\, \mathbf{r} - (r^2) \dot{\mathbf{r}} \right].$$

$$= -m f(r)\, r^2 \left[\frac{\dot{r}}{r} \mathbf{r} - \dot{\mathbf{r}} \right].$$

$$= -m f(r)\, r^2 \frac{d}{dt} \left(\frac{\mathbf{r}}{r} \right)$$

Since $\mathbf{L}$ is constant $\dot{p} \times \mathbf{L} = \frac{d}{dt}(p \times \mathbf{L})$ and:

$$\frac{d}{dt}(p \times \mathbf{L}) = -m f(r)\, r^2 \frac{d}{dt} \left(\frac{\mathbf{r}}{r} \right).$$

In the case of an inverse square law such as gravity $f(r) = -k/r^2$, and we find a new conserved quantity:

$$\frac{d}{dt} \left(p \times \mathbf{L} - mk \left(\frac{\mathbf{r}}{r} \right) \right) = 0.$$

The vector quantity that is conserved,

$$\mathbf{A} = p \times \mathbf{L} - \left(\frac{mk}{r} \right) \mathbf{r}$$

is called the Runge-Lenz vector but perhaps should be called the Hermann-Bernoulli-Laplace-Hamilton-Gibbs-Runge-Lenz vector since it has been rediscovered many times. Ironically Runge and Lenz didn't rediscover the vector: Runge used Gibbs work as an example in a popular textbook and Lenz cited that text in a paper.

Challenge: can you find the coordinate transformation of the Lagrangian

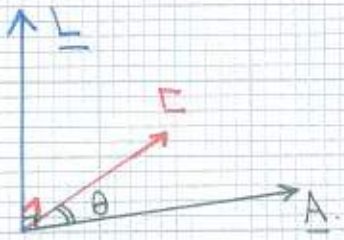
that leaves it invariant and by Noether's theorem gives that $\underline{A}$ is a conserved quantity? (See Wikipedia: "Laplace-Runge-Lenz" for the answer.)

Let us make use of $\underline{A}$ to find the shape of the orbit.

Since $\underline{L} = \underline{r} \wedge \underline{p}$ then we note that,

$$\underline{A} \cdot \underline{L} = 0.$$

$\underline{L}$ is perpendicular to both $\underline{r}$ and $(\underline{p} \wedge \underline{L})$. Therefore $\underline{A}$, which is a constant vector for the motion, is fixed in the plane of the orbit.



$\underline{r}$ and $\underline{A}$ are both perpendicular to $\underline{L}$.

We denote the angle between $\underline{r}$ and $\underline{A}$ by θ (in otherwords our predefined coordinate θ takes the value $\theta = 0$ in the $\underline{A}$ direction).

$$\therefore \underline{A} \cdot \underline{r} = Ar\cos\theta \qquad \text{where} \quad A \equiv |\underline{A}|, \quad r \equiv |\underline{r}|.$$

From the definition of $\underline{A}$ we find,

$$\begin{aligned} \underline{A} \cdot \underline{r} &= (\underline{p} \wedge \underline{L}) \cdot \underline{r} - mkr \\ &= (\underline{r} \wedge \underline{p}) \cdot \underline{L} - mkr \\ &= L^2 - mkr. \end{aligned}$$

Recall that $L \equiv p_\theta \equiv |\underline{L}|$.

Combining our two expressions for $\underline{A} \cdot \underline{r}$ we have:

$$\begin{aligned} Ar\cos\theta &= L^2 - mkr \\ \Rightarrow \quad r &= L^2 / (mk + A\cos\theta) \\ &= (L^2/mk) / (1 + (A/mk)\cos\theta) \\ &\equiv p / (1 + \varepsilon\cos\theta) \end{aligned}$$

Which is the equation of a conic section, $p \equiv L^2/mk$ is a constant and $\varepsilon \equiv A/mk$

is the "eccentricity" of the orbit.

Conic Sections.

$$r = g/(1 + \varepsilon\cos\theta).$$

Using circular coordinates:

$$x = r\cos\theta, \qquad y = r\sin\theta$$

We have: $r = g/(1 + (\varepsilon x/r))$.

$$\Rightarrow r = g - \varepsilon x.$$

This squares to:

$$r^2 = g^2 + \varepsilon^2 x^2 - 2g\varepsilon x.$$

$$\therefore x^2 + y^2 = g^2 + \varepsilon^2 x^2 - 2g\varepsilon x.$$

$$\Rightarrow g^2 = (1-\varepsilon^2)x^2 + 2g\varepsilon x + y^2. \qquad [1.]$$

Let us consider various values for ε.

Case (I.) $\varepsilon = 0 \Rightarrow g^2 = x^2 + y^2$

A circle of radius $g = l^2/mk$.

Case (II.) $0 < \varepsilon < 1$

We can complete the square for the coordinate x to find:

$$g^2 = \left((1-\varepsilon^2)^{1/2}x + g\varepsilon/(1-\varepsilon^2)^{1/2}\right)^2 - g^2\varepsilon^2/(1-\varepsilon^2) + y^2.$$

$$\therefore g^2/(1-\varepsilon^2) = \left((1-\varepsilon^2)^{1/2}x + g\varepsilon/(1-\varepsilon^2)^{1/2}\right)^2 + y^2.$$

$$\therefore g^2 = \left((1-\varepsilon^2)x + g\varepsilon\right)^2 + (1-\varepsilon^2)y^2. \qquad [2.]$$

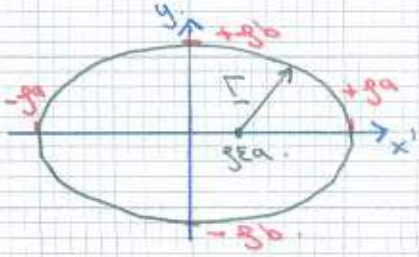
$$= x'^2/a^2 + y^2/b^2.$$

Where $x' = x + g\varepsilon/(1-\varepsilon^2)$ $\Rightarrow$ $x' = x + g\varepsilon a$.

$a^2 = 1/(1-\varepsilon^2)^2$.

$b^2 = 1/(1-\varepsilon^2)$ $\Rightarrow$ $b = (1-\varepsilon^2)^{1/2}\, a$.

With this rewriting we now recognise this as the equation of a conic section. When $0 < \varepsilon < 1$ then $a > 0$, $b > 0$, it gives the equation of an ellipse.



N.B. when $y = 0$, $x' = \pm g/(1-\varepsilon^2) = \pm ga$.

$x' = 0$, $y = \pm g/(1-\varepsilon^2)^{1/2} = \pm gb$.

Case (III.) $\varepsilon = 1$.

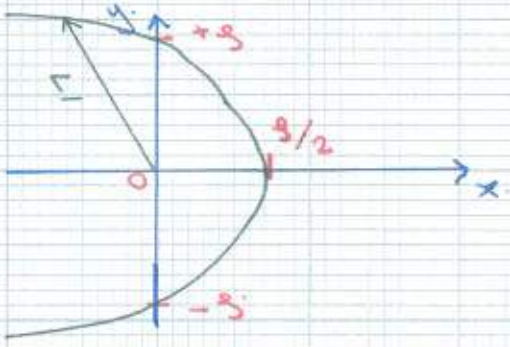
Substituting this into [2.] gives no new information:

$$\Rightarrow p^2 = p^2.$$

But substituting $\varepsilon = 1$ into [1.] gives the equation of a parabola:

$$p^2 = 2px + y^2.$$

$$\therefore x = -\frac{1}{2p}y^2 + \frac{p}{2}.$$



$y=0 \Rightarrow x = p/2.$

$x=0 \Rightarrow y^2 = p^2.$

$\therefore y = \pm p.$

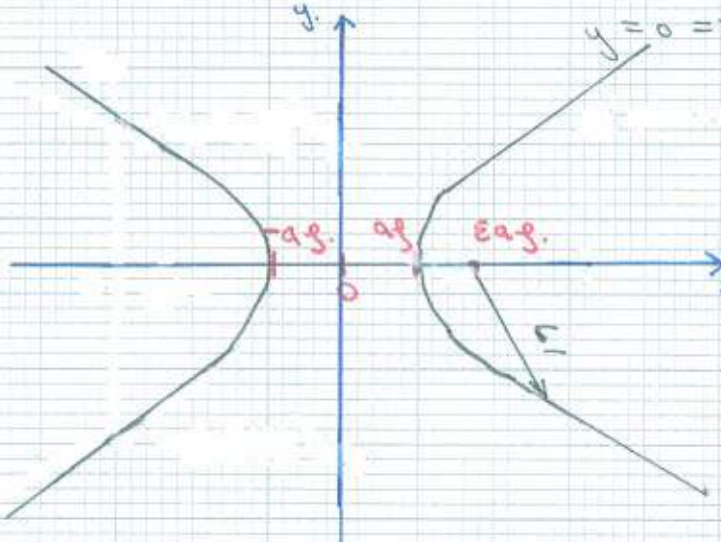
Case (IV.) $\varepsilon > 1$.

Looking at equation [2] for this case, $a^2 > 0$, $b^2 < 0$.

Writing $b'^2 = -b^2$ we find the equation of the orbit is:

$$\frac{x'^2}{a^2} - \frac{y^2}{b'^2} = p^2.$$

This equation describes a hyperbola.



$y=0 \Rightarrow x' = \pm ap.$

$x=0 \Rightarrow -y^2 = b'^2 p^2.$

as both b'^2 and y^2 are positive there are no real solutions.

N.B. we have used: $a^2 = 1/(\varepsilon^2 - 1)^2$

Compare and contrast these orbits with the four cases E_1, E_2, E_3 and E_4

in the previous section.

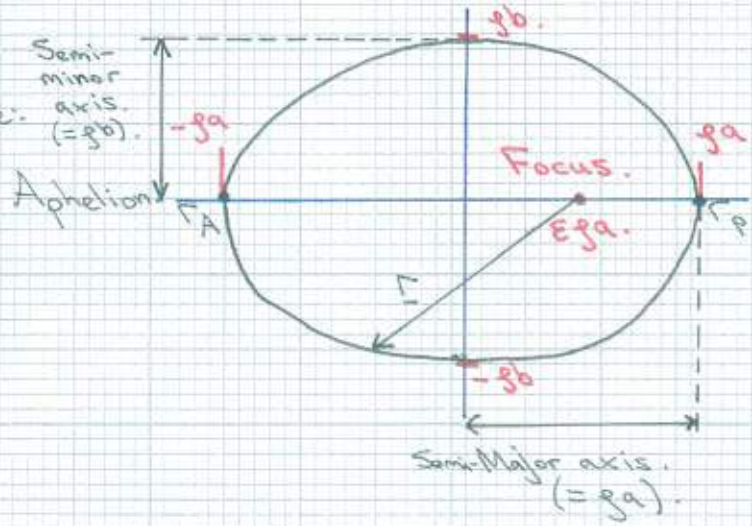
We have shown that the possible gravitational orbits are conic sections when approximated to the two-body problem. The planets move in bounded and periodic orbits therefore we have proved Kepler's First Law of planetary motion (KI): the planets move in elliptical orbits.

Orbital Periods.

Now we will use our knowledge of elliptic orbits to prove KIII that the square of the orbital period of a planet is proportional to its distance from the Sun cubed. Since we have ellipses this distance is varying so we will consider the major axis.

For our conventions, the ellipse $y^2 = x'^2/a^2 + y^2/b^2$ has semi-major and semiminor axes of lengths ag and bg respectively.

An ellipse:



Recall that $a = 1/(1-\varepsilon^2)$, $b = (1-\varepsilon^2)^{1/2} a$.

The distance from the focus to the perihelion, $r_p = ga - \varepsilon ga = ga(1-\varepsilon)$.

The distance from the focus to the aphelion, $r_A = ga + \varepsilon ga = ga(1+\varepsilon)$.

At the perihelion and aphelion the radial vector is perpendicular to the momentum vector:

$$\therefore L = |\underline{L}| = |\underline{r} \wedge \underline{p}| = |\underline{r}||\underline{p}| \quad \text{as } \sin\theta = 1.$$

Angular momentum, L, is constant and takes the same value at the perihelion and aphelion:

$$L = |\underline{r}_p||\underline{p}_p| = ga\mu(1-\varepsilon)v_p. \qquad \text{N.B. } v_p = |\dot{\underline{r}}_p|$$

And, $L = |\underline{r}_A||\underline{p}_A| = g_a \mu (1+\varepsilon) v_A$. N.B. $v_A \equiv |\dot{\underline{r}}_A|$

We can write v_A and v_p in terms of L:

$$v_p = L/g_a \mu (1-\varepsilon)$$

$$v_A = L/g_a \mu (1+\varepsilon).$$

The energy can also be equated at the perihelion and aphelion since it too is a constant of the motion:

$$E = (\mu/2) v_a^2 + V(r_a) = (\mu/2) v_p^2 + V(r_p) \quad *$$

Substituting v_a and v_p gives an expression for L^2:

$$(\mu/2)\left(L^2/\mu^2 g^2 a^2\right)\left(1/(1+\varepsilon)^2 - 1/(1-\varepsilon)^2\right) = (k/g_a)\left(-1/(1-\varepsilon) + 1/(1+\varepsilon)\right).$$

$$\therefore \quad \left(L^2/2\mu g^2 a^2\right)\left(-4\varepsilon/(1+\varepsilon)^2(1-\varepsilon)^2\right) = (k/g_a)\left(-2\varepsilon/(1-\varepsilon^2)\right).$$

$$\therefore \quad -2\varepsilon L^2/\mu g^2 a^2 (1-\varepsilon^2) = -2\varepsilon k/g_a.$$

$$\Rightarrow \quad L^2 = k\mu(g_a)(1-\varepsilon^2).$$

The angular momentum depends on g_a, the semi-major axis, and the eccentricity (ε) of the orbit.

We can also evaluate the energy for the orbit by substituting our expressions for v_p and L^2 into E:

$$E = \tfrac{\mu}{2} v_p^2 - k/r_p.$$

$$= (\mu/2)\frac{k\mu(g_a)(1-\varepsilon^2)}{g^2 a^2 \mu^2 (1-\varepsilon)^2} - \frac{k}{g_a(1-\varepsilon)}.$$

$$= \frac{k(1-\varepsilon^2)}{2(1-\varepsilon)^2 g_a} - \frac{k}{g_a(1-\varepsilon)}.$$

$$= \frac{k(1+\varepsilon) - 2k}{2g_a(1-\varepsilon)}.$$

$$= -k/2g_a.$$

The energy is dependent only on the major axis of the orbit.

Finally let us prove KIII. The rate of change of area swept out

* N.B. V and not V_{eff} is used as $v_A \equiv |\dot{\underline{r}}_A|$ - see problem 7.2 in the problem sheets.

by a planet on an elliptical path is:

$$\frac{d(\text{Area})}{dt} = \tfrac{1}{2} r^2 \dot{\theta} = \frac{L}{2\mu}.$$

Denoting the orbital period by τ we can integrate this expression to find the area of the elliptical orbit:

$$\text{Area} = \int_0^{\tau} \left(\frac{L}{2\mu}\right) dt = \frac{L\tau}{2\mu}.$$

The area of an ellipse can also be expressed in terms of its major and minor axes:

$$\text{Area} = \pi (ga)(gb) = \pi g^2 a^2 (1-\varepsilon^2)^{1/2}.$$

Where we have used $b = (1-\varepsilon^2)^{1/2} a$. Using our expression for L^2 we see that:

$$(1-\varepsilon^2)^{1/2} = \frac{L}{(k\mu g a)^{1/2}}.$$

Giving,

$$\text{Area} = \pi (ga)^{3/2} \frac{L}{\sqrt{k\mu}}.$$

Equating our two expressions for the area enclosed by the orbit:

$$\frac{L\tau}{2\mu} = \pi (ga)^{3/2} \frac{L}{\sqrt{k\mu}}$$

$$\Rightarrow \quad \tau = 2\pi (ga)^{3/2} \left(\frac{\mu}{k}\right)^{1/2}.$$

and $$\tau^2 = 4\pi^2 (ga)^3 \left(\frac{\mu}{k}\right).$$

So that the orbital period squared is proportional to the semimajor axis cubed which is the essence of KIII.

4. Hamiltonian Mechanics.

Hamiltonian mechanics is a reformulation of Lagrangian mechanics that was introduced by William Rowan Hamilton in 1833, when he was 28 years old. The central function in Hamiltonian mechanics is called the Hamiltonian and it plays the role of the Lagrangian in Lagrangian mechanics. The Hamiltonian is essentially the energy function, h, of the system we have met previously but expressed in terms of the generalised coordinates, q_i, conjugate momenta, p_i, and time, t. Writing the equations of motion in these coordinates reveals the symplectic symmetry of the dynamical system and with the introduction of Poisson brackets the groundwork is complete for the study of quantum mechanics, statistical mechanics and chaos.

Hamiltonians and Lagrangians.

In Lagrangian mechanics the equations of motion of the system with n degrees of freedom are n second-order equations of the form:

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

These can be solved to find $q_i(t)$ when $2n$ initial values are specified. The system is described by a point in an n-dimensional configuration space, and L is a function of q_i, $\dot{q}_i$ and t.

The Hamiltonian formulation is different: the aim is to describe the system in terms of $2n$ first-order differential equations of motion. The dynamics will remain unchanged of course and still $2n$ initial values are needed to specify the motion of the system. However the system will now be described by $q_i(t)$ and $p_i(t)$, that is $2n$ independent variables. The system is described by a point in a $2n$-dimensional space called phase space.

The principle difference is one of style, rarely will the equations of motion be significantly simplified in one system compared to the other. The variables used to describe the Lagrangian are the generalised coordinates, q_i, their velocities, $\dot{q}_i$ and time, t, while the Hamiltonian is a function of the generalised coordinates, q_i,

the conjugate momenta, $p_i = \frac{\partial L}{\partial \dot{q}_i}$, and time, t. The method for changing the variables in this way is called the Legendre transformation.

The Legendre Transformation.

Our aim is to transform the Lagrangian, $L(q_i, \dot{q}_i, t)$, into the Hamiltonian, $H(q_i, p_i, t)$, a function of different variables which encodes the same dynamical information. This can be achieved with the Legendre transformation. Before we do so let's consider the transformation of $f(x,y)$ into $g(u,y)$ using the Legendre transformation.

Now the total derivative of $f(x,y)$ is:

$$df = \left(\frac{\partial f}{\partial x}\right) dx + \left(\frac{\partial f}{\partial y}\right) dy.$$

Let us define a new function $g(u,x,y)$ depending on three variables, then,

$$dg = \left(\frac{\partial g}{\partial u}\right) du + \left(\frac{\partial g}{\partial x}\right) dx + \left(\frac{\partial g}{\partial y}\right) dy$$

In particular let $g(u,x,y) \equiv ux - f(x,y)$ then,

$$dg = x\,du + u\,dx - \left(\frac{\partial f}{\partial x}\right) dx - \left(\frac{\partial f}{\partial y}\right) dy.$$

Now if we choose $u \equiv \frac{\partial f}{\partial x}$ we can eliminate the dx term:

$$dg = x\,du - \left(\frac{\partial f}{\partial y}\right) dy.$$

Since there is no dx term in dg then g is a function of just two variables, u and y. Or,

$$g(u,y) \equiv u\,x(u,y) - f(x(u,y), y).$$

This is the Legendre transform, where we have used $u = \frac{\partial f}{\partial x}$ to determine $x(u,y)$.

An inverse Legendre transform always exists:

$$f(x,y) = -\left(\frac{\partial g}{\partial u}\right) u - g.$$

Hence no "information" is lost in this transformation.

The Hamiltonian is the Legendre transform of the Lagrangian:

Now $$dL = \sum_i \left\{ \left(\frac{\partial L}{\partial q_i}\right) dq_i + \left(\frac{\partial L}{\partial \dot{q}_i}\right) d\dot{q}_i \right\} + \left(\frac{\partial L}{\partial t}\right) dt$$

dL plays the role of df above. Recall that $p_i \equiv \frac{\partial L}{\partial \dot{q}_i}$ and from Lagrange's equations we have:

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

$$\therefore \quad \dot{p}_i = \frac{\partial L}{\partial q_i}.$$

Consequently $$dL = \sum_i \left\{ \dot{p}_i\, dq_i + p_i\, d\dot{q}_i \right\} + \frac{\partial L}{\partial t}\, dt.$$

We aim to eliminate $d\dot{q}_i$ in dL so $\dot{q}_i$ plays the role of x in the Legendre transform, such that: (N.B. $u \to p_i$)

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

In fact we have met the Hamiltonian before, or something similar, in the energy function, $h \equiv \sum_i p_i \dot{q}_i - L$. The difference is that the energy function was given in terms of $q_i, \dot{q}_i$ and t, while the Hamiltonian is strictly a function of q_i p_i and t.

Examples of Hamiltonians.

To find a Hamiltonian we first start with the Lagrangian, then we apply the Legendre transformation and finally we express the result in terms of q_i, p_i and t. It can be a lengthy process we will consider two of the simplest examples.

1. The free particle.

The Lagrangian for a free particle in three dimensions is:

$$L = \tfrac{1}{2} m (\dot{x}^2 + \dot{y}^2 + \dot{z}^2).$$

Using $p_i = \frac{\partial L}{\partial \dot{q}_i}$ we write down:

$$p_x = m\dot{x}, \quad p_y = m\dot{y}, \quad p_z = m\dot{z}.$$

$$\therefore \quad H \equiv \sum_i p_i \dot{q}_i - L$$

$$= m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - \tfrac{1}{2} m (\dot{x}^2 + \dot{y}^2 + \dot{z}^2).$$

$$= \tfrac{1}{2} m (\dot{x}^2 + \dot{y}^2 + \dot{z}^2)$$

$$= \tfrac{1}{2} m \left(\left(\frac{p_x}{m}\right)^2 + \left(\frac{p_y}{m}\right)^2 + \left(\frac{p_z}{m}\right)^2 \right)$$

$$= \left(\frac{1}{2m}\right)\left(p_x^2 + p_y^2 + p_z^2 \right)$$

In the last step we have eliminated $\dot{x}$, $\dot{y}$ and $\dot{z}$ since the Hamiltonian is a function of q_i, p_i and t only.

2. The particle in a plane subject to a central force.

In terms of the polar coordinates, r and θ, we find the Lagrangian is:

$$L = \tfrac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) - V(r).$$

For some potential $V(r)$ associated to the central force, $\underline{F} = -\underline{\nabla} V$.

The conjugate momenta are:

$$p_r = \frac{\partial L}{\partial \dot{r}} = m\dot{r}. \qquad \Rightarrow \dot{r} = p_r/m.$$

$$p_\theta = \frac{\partial L}{\partial \dot{\theta}} = mr^2\dot{\theta}. \qquad \Rightarrow \dot{\theta} = p_\theta/mr^2.$$

Hence the Hamiltonian is:

$$\begin{aligned} H &= \sum_i p_i\dot{q}_i - L \\ &= p_r\dot{r} + p_\theta\dot{\theta} - \tfrac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) + V(r) \\ &= p_r^2/m + p_\theta^2/mr^2 - \tfrac{1}{2}m\left(\left(\frac{p_r}{m}\right)^2 + r^2\left(\frac{p_\theta}{mr^2}\right)^2\right) + V(r). \\ &= \left(\tfrac{1}{2m}\right)\left(p_r^2 + p_\theta^2/r^2\right) + V(r). \end{aligned}$$

Note that if we write $L = T - V$ then $H = T + V$ just as for the energy function with the same provisos.

Hamilton's Equations of Motion.

We have accomplished part of our aim, namely that we have a function of the variables q_i, p_i and t that encodes the dynamics. But where are the equations of motion, are there $2n$ of them and are they first-order differential equations?

Let us look at the total differential of $H(q_i, p_i, t)$:

$$dH = \sum_i\left(\frac{\partial H}{\partial q_i}dq_i + \frac{\partial H}{\partial p_i}dp_i\right) + \frac{\partial H}{\partial t}dt.$$

From the Legendre transformation that took us from L to H:

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

We have,

$$\begin{aligned} dH &= \sum_i\left(\cancel{\dot{q}_i dp_i} + \dot{q}_i dp_i - \frac{\partial L}{\partial q_i}dq_i - \cancel{\frac{\partial L}{\partial \dot{q}_i}d\dot{q}_i}\right) - \frac{\partial L}{\partial t}dt. \\ &= \sum_i\left(\dot{q}_i dp_i - \left(\frac{\partial L}{\partial q_i}\right)dq_i\right) - \frac{\partial L}{\partial t}.dt. \end{aligned}$$

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

Comparing the two expressions allows us to read off the equations of motion in the Hamiltonian formulation:

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

$$-\left(\frac{\partial L}{\partial q_i}\right) = \frac{\partial H}{\partial q_i}$$

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

Recall from Lagrange's equations that:

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

$$\therefore \quad \frac{\partial L}{\partial q_i} = \dot{p}_i$$

Which gives Hamilton's equations of motion:

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q_i} \quad \text{and} \quad \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}.$$

These are $2n+1$ equations of motion, however we had neglected the $-\frac{\partial L}{\partial t} = \frac{\partial H}{\partial t}$ equation in our counting — and they are first-order differential equations.

Examples of the Equations of Motion.

Let us write down the equations of motion for the two Hamiltonians we wrote down earlier.

1. The free particle.

$$H = \left(1/2m\right)\left(p_x^2 + p_y^2 + p_z^2\right)$$

The equations of motion are:

$$\dot{x} = \partial H/\partial p_x = p_x/m. \qquad \dot{p}_x = -\frac{\partial H}{\partial x} = 0.$$

$$\dot{y} = p_y/m. \qquad \dot{p}_y = 0.$$

$$\dot{z} = p_z/m. \qquad \dot{p}_z = 0.$$

→ constant linear momentum, as we may have expected.

2. The particle in a plane subject to a central force.

$$H = \left(1/2m\right)\left(p_r^2 + p_\theta^2/r^2\right) + V(r)$$

The equations of motion are:

$$\dot{r} = \frac{\partial H}{\partial p_r} = p_r/m. \qquad \dot{p}_r = -\frac{\partial H}{\partial r} = +\frac{p_\theta^2}{mr^3} - \frac{\partial V}{\partial r}.$$

$$\dot{\theta} = \frac{\partial H}{\partial p_\theta} = p_\theta/mr^2. \qquad \dot{p}_\theta = -\frac{\partial H}{\partial \theta} = 0.$$

Notice that p_θ is conserved and that $\dot{p}_r = -\frac{\partial V_{eff}}{\partial r}$

where $V_{eff} = \frac{p_\theta^2}{2mr^2} + V(r)$, as before in the Lagrangian formulation.

Conservation Laws.

We can use the equations of motion in the Hamiltonian formulation to rediscover some conserved quantities that we have seen already in the Lagrangian description.

Energy

If the Hamiltonian is explicitly independent of time then it is a constant of them motion:

$$\frac{dH(q_i, p_i, t)}{dt} = \frac{\partial H}{\partial q_i}\frac{\partial q_i}{\partial t} + \frac{\partial H}{\partial p_i}\frac{\partial p_i}{\partial t} + \frac{\partial H}{\partial t}.$$

$$= -(\dot{p}_i)\dot{q}_i + (\dot{q}_i)\dot{p}_i + \frac{\partial H}{\partial t}$$

$$= \partial H/\partial t.$$

Thus if $\partial H/\partial t = 0$ then H is a constant of the motion.

You may think that it is the energy (and often it is) but the condition that it is the energy relies on the functions:

$$\Gamma_1(q_1, \ldots q_{3n-k}; t)$$
$$\vdots$$
$$\Gamma_n(q_1, \ldots q_{3n-k}; t)$$

all being explicitly independent of time, t. This is best illustrated with an example.

Example: A Cart Pulled by a Harmonic Oscillator.



Consider the simple one-dimensional system drawn above which indicates a mass, m, attached by a spring (or harmonic oscillator) to a driving force causing the other end of the spring to move at $V_0 t$.

The potential of a harmonic oscillator is $\frac{1}{2}k(\Delta x)^2$ where Δx is the amount the spring/oscillator is stretched from its rest position. Hence the potential here is $V(x) = \frac{1}{2}k(v_0 t - x)^2$.

The kinetic energy is just $\frac{1}{2}m\dot{x}^2$, so the Lagrangian is:

$$L = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}k(v_0 t - x)^2.$$

To find the Hamiltonian we need: $p_x \equiv \frac{\partial L}{\partial \dot{x}} = m\dot{x}$, $q_x \equiv x$.

So,

$$\begin{aligned} H &\equiv \sum \dot{q}_i p_i - L \\ &= m\dot{x}^2 - \tfrac{1}{2}m\dot{x}^2 + \tfrac{1}{2}k(v_0 t - x)^2 \\ &= \tfrac{1}{2}m\dot{x}^2 + \tfrac{1}{2}k(v_0 t - x)^2 \\ &= \tfrac{1}{2}m\left(\frac{p_x}{m}\right)^2 + \tfrac{1}{2}k(v_0 t - x)^2 \\ &= p_x^2/2m + \tfrac{1}{2}k(v_0 t - x)^2. \end{aligned}$$

Since the Cartesian coordinates as functions of the generalised coordinates do not explicitly depend on time (here $q_x = x$) then we observe that the Hamiltonian H is the total energy.

But note here that $\partial H/\partial t = kv_0 \neq 0$, so it is not a constant.

Alternatively, suppose we had picked as our general coordinate:

$$q_x = x - v_0 t.$$

Then the Lagrangian becomes:

$$\begin{aligned} L &= \tfrac{1}{2}m\left(\frac{d}{dt}(q_x + v_0 t)\right)^2 - \tfrac{1}{2}kq_x^2 \\ &= \tfrac{1}{2}m\left(\dot{q}_x + v_0\right)^2 - \tfrac{1}{2}kq_x^2 \\ &= \tfrac{1}{2}m\dot{q}_x^2 + mv_0\dot{q}_x + \tfrac{1}{2}mv_0^2 - \tfrac{1}{2}kq_x^2. \end{aligned}$$

Using, $\frac{\partial L}{\partial \dot{q}_x} \equiv p_x = m\dot{q}_x + mv_0$. then we find:

$$\begin{aligned} H &\equiv \sum \dot{q}_i p_i - L \\ &= \dot{q}_x(m\dot{q}_x + mv_0) - L \\ &= \tfrac{1}{2}m\dot{q}_x^2 - \tfrac{1}{2}mv_0^2 + \tfrac{1}{2}kq_x^2. \\ &= (p_x - mv_0)^2/2m - \tfrac{1}{2}mv_0^2 + \tfrac{1}{2}kq_x^2. \end{aligned}$$

Here the Cartesian coordinates are explicit functions of time:

$$q_x = x - v_0 t \iff x = q_x + v_0 t.$$

So the Hamiltonian for these generalised coordinates is not the total energy of the system. On the other hand $\partial H/\partial t = 0$ now so it is a constant of the system.

Notice that with the second set of coordinates we have moved to an inertial frame which moves with the driven end of the spring at speed v_0. The change of coordinates subtracts the kinetic energy of the frame $\frac{1}{2}mv_0^2$ from the original Hamiltonian.

Conserved Conjugate Momenta.

Hamilton's equations of motion are:

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

Hence if q_i is ignorable and doesn't appear in the Hamiltonian then p_i is conserved as $\dot{p}_i = 0$.

Poisson Brackets.

Let us now introduce some useful notation which played a fundamental part in the foundations of quantum mechanics!

Let F and g be two functions on phase space,

i.e. $F \equiv F(q_1 \cdots q_n, p_1 \cdots p_n)$

$g \equiv g(q_1 \cdots q_n, p_1 \cdots p_n)$

Then we define the Poisson bracket as:

$$\{F, g\} \equiv \sum_i \left(\frac{\partial F}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial g}{\partial q_i}\frac{\partial F}{\partial p_i} \right).$$

It has the following properties:

- $\{F, g\} = -\{g, F\}$ ANTISYMMETRIC.
- $\{aF + bg, h\} = a\{F, h\} + b\{g, h\}$ BILINEAR.

$$\Rightarrow \{F, ag + bh\} = a\{F, g\} + b\{F, h\}$$

- $\{fg,h\} = f\{g,h\} + \{f,h\}g$. LEIBNITZ RULE.
- $\{f,\{g,h\}\} + \{g,\{h,f\}\} + \{h,\{f,g\}\} = 0$. JACOBI IDENTITY.

These properties can all be checked using the definition and some patience.

There are some fundamental Poisson brackets that we will use later:

$$\{q_i, q_k\} = \sum_i \left(\frac{\partial q_i}{\partial q_i}\frac{\partial q_k}{\partial p_i} - \frac{\partial q_k}{\partial q_i}\frac{\partial q_i}{\partial p_i} \right)$$
$$= \sum_i \left(\delta_{ik}(0) - \delta_{ik}(0) \right)$$
$$= 0.$$

$$\{p_i, p_k\} = \sum_i \left(\frac{\partial p_i}{\partial q_i}\frac{\partial p_k}{\partial p_i} - \frac{\partial p_k}{\partial q_i}\frac{\partial p_i}{\partial p_i} \right)$$
$$= \sum_i \left((0)\delta_{ik} - (0)\delta_{ii} \right)$$
$$= 0.$$

$$\{q_i, p_k\} = \sum_i \left(\frac{\partial q_i}{\partial q_i}\frac{\partial p_k}{\partial p_i} - \frac{\partial p_k}{\partial q_i}\frac{\partial q_i}{\partial p_i} \right)$$
$$= \sum_i \left(\delta_{ii}\delta_{ki} - 0 \right)$$
$$= \delta_{ik}.$$

When we consider $\{f,g\}$ in general we will be able to use these fundamental Poisson brackets to calculate the answer.

Let's calculate two more Poisson brackets:

$$\{q_i, g\} = \sum_i \left(\frac{\partial q_i}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial g}{\partial q_i}\frac{\partial q_i}{\partial p_i} \right)$$
$$= \sum_i \left(\delta_{ii}\frac{\partial g}{\partial p_i} - 0 \right)$$
$$= \frac{\partial g}{\partial p_i}.$$

Similarly, $\{p_i, g\} = \sum_i \left(\frac{\partial p_i}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial g}{\partial q_i}\frac{\partial p_i}{\partial p_i} \right)$
$$= \sum_i \left(0 - \delta_{ii}\frac{\partial g}{\partial q_i} \right)$$
$$= -\frac{\partial g}{\partial q_i}.$$

The last two brackets look similar to the equations of motion (namely $\dot{q}_i = \partial H/\partial p_i$, $\dot{p}_i = -\partial H/\partial q_i$) and we can write the equations of motion as Poisson brackets. Before doing so let us observe that for any function $f(q_i, p_i, t)$ then,

$$\frac{df}{dt} = \sum_i \left(\frac{\partial f}{\partial q_i}\frac{\partial q_i}{\partial t} + \frac{\partial f}{\partial p_i}\frac{\partial p_i}{\partial t} \right) + \frac{\partial f}{\partial t}$$

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

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

This is a very useful equation; we can use it to find constants of the motion. For suppose $Q \equiv Q(q_i, p_i)$ is a constant of the motion then as $\partial Q/\partial t = 0$ and $\frac{dQ}{dt} = 0$:

$$\Rightarrow \{Q, H\} = 0.$$

Any time-independent Q ($\partial Q/\partial t = 0$) which "Poisson commutes" with the Hamiltonian, H, is a constant of the motion.

e.g. if H is explicitly time-independent $\partial H/\partial t = 0$ then as $\{H, H\} = 0$ H is a constant of the motion.

e.g. if H is independent of q_i then $\{p_i, H\} = -\frac{\partial H}{\partial q_i} = 0$ and $\frac{\partial p_i}{\partial t} = 0$.

hence p_i is a constant of the motion.

We can now write the equations of motion using Poisson brackets:

$$\frac{dq_i}{dt} = \{q_i, H\} + \cancel{\frac{\partial q_i}{\partial t}}^{\,0}$$

$$\Rightarrow \quad \dot{q}_i = \{q_i, H\}.$$

$$\frac{dp_i}{dt} = \{p_i, H\} + \cancel{\frac{\partial p_i}{\partial t}}^{\,0}$$

$$\Rightarrow \quad \dot{p}_i = \{p_i, H\}.$$

N.B. q_i, p_i and t are independent variables hence $\partial q_i/\partial t = \partial p_i/\partial t = 0$.

Conservation of angular-momentum using Poisson brackets.

Suppose we have a system in which two components of angular momentum are conserved then:

$$\{L_x, H\} = 0 = \{L_y, H\}.$$

Then by the Jacobi identity:

$$\{\{L_x, L_y\}, H\} = -\{\{L_y, H\}, L_x\} - \{\{H, L_x\}, L_y\}$$

$$= 0.$$

Hence if L_x and L_y are conserved then the quantity $\{L_x, L_y\}$ is conserved too. What is $\{L_x, L_y\}$?

Using $\underline{L} = \underline{r} \wedge \underline{p}$ we write the components out:

$$L_x = r_y p_z - r_z p_y.$$

$$L_y = r_z p_x - r_x p_z.$$

$$L_z = r_x p_y - r_y p_x.$$

Now,
$$\{L_x, L_y\} = \{r_y p_z - r_z p_y, r_z p_x - r_x p_z\}.$$

$$= \{r_y p_z, r_z p_x\} + \{r_z p_y, r_x p_z\}.$$

$$= r_y \{p_z, r_z\} p_x + \{r_z, p_z\} p_y r_x.$$

$$= -r_y p_x + p_y r_x.$$

$$= L_z.$$

In the second and third lines we have made use of the fundamental Poisson brackets: $\{r_i, p_j\} = \delta_{ij}$, $\{r_i, r_j\} = 0$, $\{p_i, p_j\} = 0$. This allows us to isolate all the possible non-zero brackets in line 2 (here $\{r_z, p_z\}$ is the only possible non-zero bracket). We have used the Leibniz rule to turn brackets of products into a sum of brackets of single phase space variables (e.g. $\{r_y p_z, r_z p_x\} = r_y \{p_z, r_z p_x\} + \{r_y, r_z p_x\} p_z$.

$= r_y \{p_z, r_z\} p_x + r_y \{p_z, p_x\}^{0} r_z$).

Therefore if L_x and L_y is conserved then L_z is conserved too.
One can also show that $\underline{L}^2$ is conserved:

$$\{L^2, H\} = \{L_x^2 + L_y^2 + L_z^2, H\}$$

$$= 2\{L_x, H\} + 2\{L_y, H\} + 2\{L_z, H\}$$

$$= 0.$$

Canonical Transformations.

Hamiltonian mechanics puts q_i and p_i on an almost equal footing: if we interchange q_i and p_i the equations of motion are changed only by a minus sign. We can write the equations of motion more compactly using the vector, $\underline{x} \equiv \begin{pmatrix} q_1 \\ \vdots \\ q_n \\ p_1 \\ \vdots \\ p_n \end{pmatrix}$.

and the matrix $\underline{\underline{J}} = \begin{pmatrix} 0 & & & 1 & & 0 \\ & \ddots & & & \ddots & \\ 0 & & 0 & 0 & & 1 \\ -1 & & 0 & 0 & & 0 \\ & \ddots & & & \ddots & \\ 0 & & -1 & 0 & & 0 \end{pmatrix} = \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix}$

Where $\underline{\underline{J}}$ is a $2n \times 2n$ square matrix, then,

$$\dot{q}_i = \partial H/\partial p_i, \quad \dot{p}_i = -\partial H/\partial q_i$$

are equivalent to:

$$\dot{\underline{x}} = \underline{\underline{J}} \cdot \frac{\partial H}{\partial \underline{x}}. \qquad \left(\text{In component notation: } \dot{x}_i = \sum_j J_{ij} \frac{\partial H}{\partial x_j} \right)$$

Suppose we make a transformation of the coordinates $x_i \to y_i(x)$ we may wonder under what conditions the equations of motion remain unchanged i.e. under what conditions do we find:

$$\dot{y}_i = \underline{\underline{J}} \cdot \frac{\partial H}{\partial y_i}.$$

Let us work out the left-hand-side:

$$\dot{y}_i = \sum_j \frac{\partial y_i}{\partial x_j} \frac{\partial x_j}{\partial t} \qquad \text{but } \dot{x}_j = \sum_k J_{jk} \frac{\partial H}{\partial x_k}$$

$$= \sum_{j,k} \frac{\partial y_i}{\partial x_j} J_{jk} \frac{\partial H}{\partial x_k}.$$

$$= \sum_{j,k,l} \frac{\partial y_i}{\partial x_j} J_{jk} \frac{\partial H}{\partial y_l} \frac{\partial y_l}{\partial x_k}.$$

$$= \sum_{i,l} \sum_{j,k} \left(\frac{\partial y_i}{\partial x_j} \right) J_{jk} \left(\frac{\partial y_l}{\partial x_k} \right) \frac{\partial H}{\partial y_l}.$$

Let us write $\mathcal{J}_{ij} \equiv \partial y_i / \partial x_j$ it is the Jacobian. Then if

$$\sum_{j,k} \mathcal{J}_{ij} J_{jk} \mathcal{J}_{lk} = J_{il} = (\mathcal{J} \cdot J \cdot \mathcal{J}^T)_{il}.$$

the form of the equations of motion is unaltered.

In matrix notation we have:

$$\mathbf{M} \cdot \mathbf{J} \cdot \mathbf{M}^T = \mathbf{J}$$

If this is true the Jacobian is "symplectic" and a change of coordinates and momenta that has a symplectic Jacobian is said to be a "canonical" transformation.

The Poisson bracket is invariant under canonical transformations:

$$\{f, g\} = \sum_{i,j} \left(\frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial g}{\partial q_i}\frac{\partial f}{\partial p_i} \right)$$

$$= \sum_{i,j} \frac{\partial f}{\partial x_i} J_{ij} \frac{\partial g}{\partial x_j}.$$

If we make a canonical transformation: $x_i \to y_i(x)$ such that $\mathbf{M}\mathbf{J}\mathbf{M}^T = \mathbf{J}$

$$\{f, g\} = \sum_{i,j,k,l} \frac{\partial f}{\partial y_k}\left(\frac{\partial y_k}{\partial x_i}\right) J_{ij} \left(\frac{\partial y_l}{\partial x_j}\right)\frac{\partial g}{\partial y_l}$$

$$= \sum_{k,l} \frac{\partial f}{\partial y_k} J_{kl} \frac{\partial g}{\partial y_l}.$$

We can also prove the reverse statement: any transformation of phase space which preserves the Poisson bracket structure is canonical.

Suppose we transform phase space to arrive at:

$$Q_i = Q_i(q,p) \quad \text{and} \quad P_i = P_i(q,p)$$

The Poisson bracket structure is preserved if:

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

as for q_i and p_i.

The Jacobian is: $$M_{ij} = \begin{pmatrix} \partial Q_i/\partial q_j & \partial Q_i/\partial p_j \\ \partial P_i/\partial q_j & \partial P_i/\partial p_j \end{pmatrix}$$

Now, $$\left(\mathbf{M}\cdot\mathbf{J}\cdot\mathbf{M}^T\right)_{ij} = \begin{pmatrix} \{Q_i, Q_j\} & \{Q_i, P_j\} \\ -\{Q_i, P_j\} & \{P_i, P_j\} \end{pmatrix} = \mathbf{J}$$

If the Poisson bracket structure is preserved.

Generating Functions.

In this section we will use functions on the phase space to generate infinitesimal transformations of the form:

$$q_i \to q_i' = q_i + \alpha f_i(q,p)$$

$$p_i \to p_i' = p_i + \alpha g_i(q,p)$$

Where α is an infinitesimal parameter. Let us ask the question for which functions f and g this is a canonical transformation, preserving the form of the equations of motion.

Writing the Jacobian:

$$\mathcal{J}_{ij} = \begin{pmatrix} \delta_{ij} + \alpha\, \partial f_i/\partial q_j & \alpha\, \partial f_i/\partial p_j \\ \alpha\, \partial g_i/\partial q_j & \delta_{ij} + \alpha\, \partial g_i/\partial p_j \end{pmatrix}$$

Requiring that $\mathcal{J}_{ij} \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix} \mathcal{J}_{kl} = \mathcal{J} \cdot \mathcal{J} \cdot \mathcal{J}^T = \mathcal{J}$ gives:

$$\begin{pmatrix} \delta_{ij} + \alpha\, \partial f_i/\partial q_j & \alpha\, \partial f_i/\partial p_j \\ \alpha\, \partial g_i/\partial q_j & \delta_{ij} + \alpha\, \partial g_i/\partial p_j \end{pmatrix} \begin{pmatrix} \alpha\, \partial f_j/\partial q_l & \delta_{jl} + \alpha\, \partial g_j/\partial p_l \\ -\delta_{jl} - \alpha\, \partial f_j/\partial q_l & -\alpha\, \partial g_j/\partial q_l \end{pmatrix}$$

$$= \begin{pmatrix} \alpha\left(\partial f_i/\partial p_l - \partial f_i/\partial q_l\right) + O(\alpha^2) & \delta_{il} + \alpha\left(\partial g_i/\partial p_l + \partial f_i/\partial q_l\right) + O(\alpha^2) \\ -\delta_{il} - \alpha\left(\partial f_i/\partial q_l + \partial g_i/\partial p_l\right) + O(\alpha^2) & \alpha\left(\partial g_i/\partial q_l - \partial g_i/\partial q_l\right) + O(\alpha^2) \end{pmatrix}$$

$$= \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix}.$$

Neglecting terms of $O(\alpha^2)$ then we see that a canonical transformation satisfies:

$$\frac{\partial f_i}{\partial q_l} = -\frac{\partial g_i}{\partial p_l}.$$

Without doing the matrix multiplication above we could have evaluated the Poisson brackets and checked the condition that the Poisson structure is preserved, it is much easier. Observe the power of the Poisson bracket:

$$\{q_i', q_j'\} = \{q_i + \alpha f_i, q_j + \alpha f_j\} = \alpha\{f_i, q_j\} + \alpha\{q_i, f_j\} + O(\alpha^2).$$
$$= O(\alpha^2).$$

$$\{p_i', p_j'\} = \{p_i + \alpha g_i, p_j + \alpha g_j\} = \alpha\{g_i, p_j\} + \alpha\{p_i, g_j\} + O(\alpha^2).$$
$$= O(\alpha^2).$$

$$\{q_i', p_j'\} = \{q_i + \alpha f_i, p_j + \alpha g_j\} = \delta_{ij} + \alpha\{f_i, p_j\} + \alpha\{q_i, q_j\} + O(\alpha^2).$$
$$= \delta_{ij} + \alpha\left(\frac{\partial f_i}{\partial q_j} + \frac{\partial g_i}{\partial p_j}\right) + O(\alpha^2).$$

We can satisfy $\frac{\partial f_i}{\partial q_i} = -\frac{\partial g_i}{\partial p_i}$ if there exists some function $G(q,p)$ which generates the transformation as:

$$f_i \equiv \frac{\partial G}{\partial p_i}, \quad g_i \equiv -\frac{\partial G}{\partial q_i}$$

This automatically implies $\partial f_i/\partial q_i = \partial^2 G/\partial q_i \partial p_i = -\partial g_i/\partial p_i$.

N.B. $q_i' = q_i + \alpha f_i = q_i + \alpha \frac{\partial G}{\partial p_i} = q_i + \alpha\{q_i, G\}$, $p_i' = p_i + \alpha\{p_i, G\}$.

Examples. These are the canonical transformations generated by G.

$G = H \Rightarrow f_i = \frac{\partial G}{\partial p_i} = \frac{\partial H}{\partial p_i} = \dot{q}_i$

$\Rightarrow g_i = -\frac{\partial G}{\partial q_i} = -\frac{\partial H}{\partial q_i} = \dot{p}_i$

The coordinate transformations are:

$$q_i \to q_i + \alpha \dot{q}_i = q_i + \delta t \frac{\partial q_i}{\partial t}.$$

$$p_i \to p_i + \alpha \dot{p}_i = p_i + \delta t \frac{\partial p_i}{\partial t}.$$

Writing $\alpha = \delta t$ we see that the Hamiltonian generates time translations.

$G = p_k \Rightarrow f_i = \frac{\partial p_k}{\partial p_i} = \delta_{ik}$.

$\Rightarrow g_i = -\frac{\partial p_k}{\partial q_i} = 0$.

The coordinate transformations are:

$$q_i \to q_i + \alpha \delta_{ik}.$$

$$p_i \to p_i.$$

So p_k generates a translation in the direction q_k.

$G = -q_k \Rightarrow f_i = 0$.

$\Rightarrow g_i = \delta_{ik}$

The coordinate transformations are:

$$q_i \to q_i$$

$$p_i \to p_i + \alpha \delta_{ik}.$$

Hence $-q_k$ generates momentum boosts in the k'th coordinate.

$G = L_z = q_x p_y - q_y p_x$

$\Rightarrow f_i = \frac{\partial L_z}{\partial p_i} = q_x \delta_{yi} - q_y \delta_{xi}$

$\Rightarrow g_i = -\frac{\partial L_z}{\partial q_i} = -p_y \delta_{xi} + p_x \delta_{yi}$.

The coordinate transformations are:

$$q_i \to q_i + \alpha(q_x\delta_{yi} - q_y\delta_{xi})$$

i.e. $q_x \to q_x + \alpha(-q_y)$. $\quad(\{q_x, L_z\} = -\{q_x, q_y p_x\} = -q_y)$

$q_y \to q_y + \alpha(q_x)$. $\quad(\{q_y, L_z\} = \{q_y, q_x p_y\} = q_x)$

$q_z \to q_z$. $\quad(\{q_z, L_z\} = 0)$

$$p_i \to p_i + \alpha(-p_y\delta_{xi} + p_x\delta_{yi})$$

i.e. $p_x \to p_x - \alpha p_y$. $\quad(\{p_x, L_z\} = \{p_x, q_x p_y\} = -p_y)$

$p_y \to p_y + \alpha p_x$. $\quad(\{p_y, L_z\} = \{p_y, -q_y p_x\} = p_x)$

$p_z \to p_z$. $\quad(\{p_z, L_z\} = 0.)$

This is a rotation about the z-axis. In general L_i generates rotations about the q_i direction.

Noether's Theorem Again!

Let us take some time to put our generating functions for canonical coordinate transformations on a clear footing.

We have been thinking of canonical transformations as a change of coordinates:

$$q_i \to q_i' = q_i + \alpha f_i(q_i, p_i) \qquad f_i \equiv \partial G/\partial p_i$$

$$\equiv q_i + \alpha \partial G/\partial p_i$$

$$\equiv q_i + \alpha\{q_i, G\}$$

$$p_i \to p_i' = p_i + \alpha g_i(q_i, p_i) \qquad g_i \equiv -\partial G/\partial q_i.$$

$$\equiv p_i - \alpha \partial G/\partial q_i$$

$$\equiv p_i + \alpha\{p_i, G\}$$

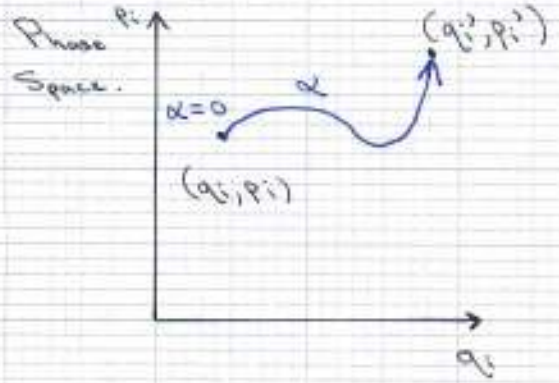
Such that when $\alpha = 0$, $(q_i', p_i') = (q_i, p_i)$.

We have two ways of labelling the same point in phase space when $\alpha = 0$. As α changes we have a map between two different points in phase space, corresponding to two different states of the system. We may think of α as parameterising a change in the system such that:

$$\partial q_i/\partial \alpha = \partial G/\partial p_i, \qquad \partial p_i/\partial \alpha = -\partial G/\partial q_i.$$

If G were the Hamiltonian these would be Hamilton's equations

and α would be time, t. This was the case in our first example. These equations describe how q_i and p_i change as α varies they are said to describe a "Hamiltonian Flow" on phase space (For some "Hamiltonian", G)



So, as described, the generating function G generates a transformation of the system paramaterised by α. For the Lagrangian formulation we saw that every continuous symmetry gave rise to a conserved quantity. This was Noether's theorem and we may wonder how it appears in the Hamiltonian Formulation.

Consider the infinitesimal canonical transformation generated by G:

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

$$= \sum_i \alpha \frac{\partial H}{\partial q_i} \frac{\partial G}{\partial p_i} + \alpha \left(\frac{\partial H}{\partial p_i} \right) \left(-\frac{\partial G}{\partial q_i} \right)$$

$$= \alpha \{H, G\}$$

When $\delta H = 0$ this is a symmetry of the Hamiltonian. $\Rightarrow \{H, G\} = 0$ which we already know means that:

$$\dot{G} \equiv \{G, H\} = 0.$$

Hence if G generates a symmetry of the Hamiltonian then G is a conserved quantity of the dynamics.

Conversely if we have a conserved quantity we can use it to generate a canonical transformation of the Hamiltonian, which is what we did earlier in the examples and may be worth reviewing at this point. This is a simple proof of Noether's theorem that we saw previously.

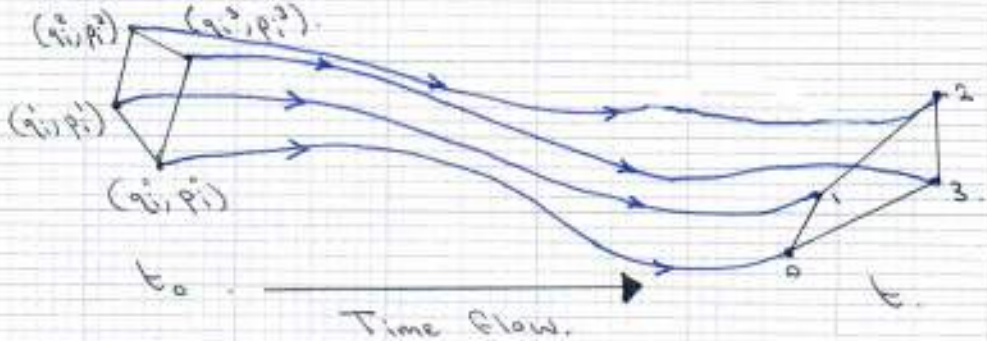
Liouville's Theorem.

Joseph Liouville (1809-1882) was a French mathematician who worked in a great number of different mathematical subjects. He was the first to prove the existence of transcendental numbers and promoted the overlooked work by Évariste Galois*. He invented a transcendental number $0.1100010000000000000000100\ldots$ (0 everywhere apart from the positions $n!$ where there is a 1, $n = 1, 2, 3 \ldots$) and has a crater on the moon named after him. There are two Liouville Theorems in Mathematics one concerns complex analysis and the other Hamiltonian dynamics, we will discuss the latter.

Liouville's theorem is concerned with the properties of phase space. Given Hamilton's equations:

$$\dot{q}_i = \partial H / \partial p_i \quad , \quad \dot{p}_i = -\partial H / \partial q_i .$$

We can describe the evolution of a dynamical system as a line in phase space starting at $(q_1(t_0), \ldots q_n(t_0), p_1(t_0) \ldots p_n(t_0))$ and moving to $(q_1(t), \ldots q_n(t), p_1(t) \ldots p_n(t))$ at time t. If we had started at a neighbouring point we would have evolved to a different point in phase space at time t:



Using such concepts we can think of Hamilton's equations as describing the flow of phase space and morphing volumes of phase space at time t_0 into volumes of phase space at later times.

Liouville's theorem:

The volume of a region of phase space is invariant under the Hamiltonian flow.

---

* the founder of group theory was died in a duel at the age of 21.

The Hamiltonian flow is the change in phase space brought about by the time evolution of Hamilton's equations:

$$\dot{q}_i = \partial H/\partial p_i \quad , \quad \dot{p}_i = -\partial H/\partial q_i .$$

The idea is simply that the volume, or integration measure,

$$V = dq_1\, dq_2 \ldots dq_n\, dp_1\, dp_2 \ldots dp_n$$

is constant for all times, t.

Consider the evolution of the infinitesimal volume V, by Hamilton's equations in time dt:

$$q_i \to q_i + \dot{q}_i\, dt = q_i + \frac{\partial H}{\partial p_i} dt \equiv \hat{q}_i$$

$$p_i \to p_i + \dot{p}_i\, dt = p_i - \frac{\partial H}{\partial q_i} dt \equiv \hat{p}_i$$

The new infinitesimal volume in phase space is:

$$\hat{V} = d\hat{q}_1\, d\hat{q}_2 \ldots d\hat{q}_n\, d\hat{p}_1\, d\hat{p}_2 \ldots d\hat{p}_n .$$

$$\equiv \det\left(\underline{\underline{\mathcal{J}}}\right) V$$

Where $\underline{\underline{\mathcal{J}}}$ is the Jacobian matrix: $\underline{\underline{\mathcal{J}}} \equiv \begin{pmatrix} \partial\hat{q}_i/\partial q_j & \partial\hat{q}_i/\partial p_j \\ \partial\hat{p}_i/\partial q_j & \partial\hat{p}_i/\partial p_j \end{pmatrix}$.

(N.B. the transformation is:

$$\begin{pmatrix} d\hat{q}_1 \\ d\hat{q}_2 \\ \vdots \\ d\hat{p}_n \end{pmatrix} \equiv \underline{\underline{\mathcal{J}}} \cdot \begin{pmatrix} dq_1 \\ dq_2 \\ \vdots \\ dp_n \end{pmatrix}$$

hence V is scaled by a factor $\det\left(\underline{\underline{\mathcal{J}}}\right)$.

Liouville's theorem says that $\det\left(\underline{\underline{\mathcal{J}}}\right) = 1$. Using the transformations $q_i \to \hat{q}_i$ and $p_i \to \hat{p}_i$ we can check if this is true.

We need:

$$\partial\hat{q}_i/\partial q_j = \partial/\partial q_j\left(q_i + \frac{\partial H}{\partial p_i} dt\right) = \delta_{ij} + \frac{\partial^2 H}{\partial q_j \partial p_i} dt .$$

$$\partial\hat{q}_i/\partial p_j = \partial/\partial p_j\left(q_i + \frac{\partial H}{\partial p_i} dt\right) = \frac{\partial^2 H}{\partial p_j \partial p_i} dt .$$

$$\partial\hat{p}_i/\partial q_j = \partial/\partial q_j\left(p_i - \frac{\partial H}{\partial q_i} dt\right) = -\partial^2 H/\partial q_j \partial q_i$$

$$\partial\hat{p}_i/\partial p_j = \partial/\partial p_j\left(p_i - \partial H/\partial q_i\, dt\right) = \delta_{ij} - \frac{\partial^2 H}{\partial p_j \partial q_i} dt .$$

To evaluate $\det \underline{\underline{\mathcal{J}}}$ we will use the formula:

$$\det\left(\underline{\underline{\mathbb{1}}} + \varepsilon \underline{\underline{M}}\right) = 1 + \varepsilon \mathrm{Tr}(\underline{\underline{M}}) + O(\varepsilon^2) .$$

Let us convince ourselves of this formula, consider a 2D matrix,

$\underline{\underline{M}} = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}$ then $\underline{\underline{\mathbb{1}}} + \varepsilon\underline{\underline{M}} = \begin{pmatrix} 1+\varepsilon m_{11} & \varepsilon m_{12} \\ \varepsilon m_{21} & 1+\varepsilon m_{22} \end{pmatrix}$.

And,

$$\det(\mathbb{1} + \varepsilon \underline{\underline{M}}) = (1+\varepsilon m_{11})(1+\varepsilon m_{22}) - \varepsilon^2 m_{21} m_{12}$$
$$= 1 + \varepsilon(m_{11}+m_{22}) + \varepsilon^2(m_{11}m_{22} - m_{21}m_{12})$$
$$= 1 + \varepsilon \operatorname{Tr} \underline{\underline{M}} + O(\varepsilon^2).$$

For an $n \times n$ matrix $(\varepsilon \underline{\underline{M}} + \mathbb{1})$ only the diagonal elements give rise to the ε coefficient as with the 2D case, this is something you can check for yourself. We have:

$$\underline{\underline{\mathcal{T}}} = \begin{pmatrix} \delta_{ij} + \left(\partial^2 H / \partial q_i \partial p_j\right) dt & \left(\partial^2 H / \partial p_i \partial p_j\right) dt \\ -\left(\partial^2 H / \partial q_i \partial q_j\right) dt & \delta_{ij} - \left(\partial^2 H / \partial p_i \partial q_j\right) dt \end{pmatrix}$$

(N.B. $\delta_{ij} = \mathbb{1}$).

So that,

$$\det(\underline{\underline{\mathcal{T}}}) = \det(\mathbb{1} + dt\, \underline{\underline{M}}) \quad \text{for some } \underline{\underline{M}} \text{ which we can read off.}$$
$$= 1 + (dt)\operatorname{Tr}(\underline{\underline{M}}) + O(dt^2).$$
$$= 1 + (dt)\left(\frac{\partial^2 H}{\partial q_i \partial p_i} - \frac{\partial^2 H}{\partial p_i \partial q_i}\right) + O(dt^2)$$
$$= 1 + O(dt^2).$$

Hence $\det(\underline{\underline{\mathcal{T}}}) = 1$ for infinitesimal dt and so volumes (both infinitesimal and integrated) are constant in phase space under Hamiltonian flow.

(For reference above: $\underline{\underline{M}} = \begin{pmatrix} \partial^2 H / \partial q_i \partial p_j & \partial^2 H / \partial p_i \partial p_j \\ -\partial^2 H / \partial q_i \partial q_j & -\partial^2 H / \partial p_i \partial q_j \end{pmatrix}$).

Comments

- Liouville's theorem is useful for statistical mechanics, which deals with multiparticle systems where the number of particles is very large ($\sim 10^{23}$). Consider a gas: each gas molecule's motion will be described by a curve in phase space and for each molecule the curve will start at a different point in phase space given by its initial conditions. All the initial condition phase space points will be bounded by some volume in phase space and Liouville's theorem tells us that this minimal bounding volume is a constant of this very complicated system.
- Liouville's theorem is reminiscent of Heisenberg's uncertainty relation in quantum mechanics.

The principle of uncertainty says that:

$$\Delta q \Delta p \geq \hbar/2 \qquad (\hbar \text{ is Planck's constant})$$

Where Δq is the uncertainty in position

Δp is the uncertainty in momentum.

Liouville's theorem can also be expressed as:

$$dq_i\, dp_i = C, \text{ a constant.}$$

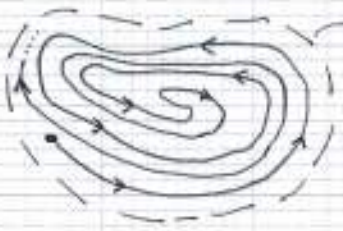
This does not appear very different from Heisenberg's uncertainty principle, but since dq_i and dp_i can be continuous quantities in classical mechanics but Δq and Δp are discrete in quantum mechanics it leads to big differences.

The Poincaré Recurrence Theorem.

(Jules) Henri Poincaré (1854 - 1912) is frequently described as the last great polymath or the last universalist of mathematics since he worked on so many different mathematical areas in his lifetime. He was a founder of topology and, effectively, chaos theory; he made the Poincaré conjecture recently proved by Perelman (gaining him a Field's medal that he rejected) and some argue that it was Poincaré's work on relativity that Einstein based his own, later, work of relativity on - although not Einstein. The Poincaré recurrence theorem was published in 1890.

The idea and observation behind the proof are relatively simple, we will state the theorem momentarily but will first make some crucial observations to motivate it. Picture the evolution of a dynamical system as a curve moving through phase space. It can only move through a finite subvolume of phase space (this subvolume corresponds to points where $V(q) \leq E$ and where E remains constant). The curve is paramaterised by t from zero to infinity and could have infinite length. If the curve intersects itself at any point then its motion is cyclic and repeats following the same

curve as before (this would give a closed loop in phase space). But what about a non-cyclic system whose motion never repeats? In this case we draw a curve of infinite length that never intersects itself inside a finite volume:

e.g.  A finite volume of phase space.

If the finite volume of phase space does not evolve we can convince ourselves that the non-intersecting curve returns to some point near the original starting point an infinite number of times in infinite time. Poincaré's theorem states just this, although it does not specialise to the Hamiltonian flow of phase space but simply to a volume preserving flow generating a curve restricted to bounded volumes of phase space. This is the case for Hamiltonian dynamics: Liouville's theorem states that the volume of phase space is constant under a Hamiltonian flow.

Theorem (Poincaré): A system with a bounded phase space will return to the neighbourhood of a previous point of its phase space in a finite time.

Proof. (For the case of a Hamiltonian flow.)

We will use the following notation: a neighbourhood of a point on the dynamical curve in phase space at time $\lambda\Delta t$ is denoted B_λ:



$\lambda \in \mathbb{Z}$.

N.B. a neighbourhood of P is an open set around P.

The theorem states that the dynamical curve will return to B_0 in a finite time. Notice that by Liouville's theorem we can choose an open set around P such that the volume of this neighbourhood is preserved by the Hamiltonian flow. As H generates the flow from B_0 to B_λ the

shape of the neighbourhood may vary but its volume is constant.

Because the phase space is finite there must exist integers λ and λ' such that the neighbourhoods B_λ and $B_{\lambda'}$ overlap:



i.e. $B_\lambda \cap B_{\lambda'} \neq \emptyset$ (the empty set).

If this was never true then the volume of phase space would be infinite $\bigcup_{\lambda=0}^{\infty} B_\lambda \to \infty$ as the B_λ could never overlap. But by our assumption of finite phase space volume we must have overlapping neighbourhoods.

Since the (Hamiltonian) flow preserves the phase space volume, the transformation has a non-zero determinant of its Jacobian, $\frac{\partial y}{\partial x}$, and so is invertible, i.e. we can invert the map: $B_\lambda \to B_{\lambda'}$. Applying the inverse of the map from $f(B_0) \to B_\lambda$ to the non-zero intersection $(B_\lambda \cap B_{\lambda'})$ we find:

$$\begin{aligned} f^{-1}(B_\lambda \cap B_{\lambda'}) &= f^{-1}(B_\lambda) \cap f^{-1}(B_{\lambda'}) \\ &= B_{(\lambda-\lambda)} \cap B_{(\lambda'-\lambda)} \\ &= B_0 \cap B_{(\lambda'-\lambda)} \\ &\neq \emptyset. \end{aligned}$$

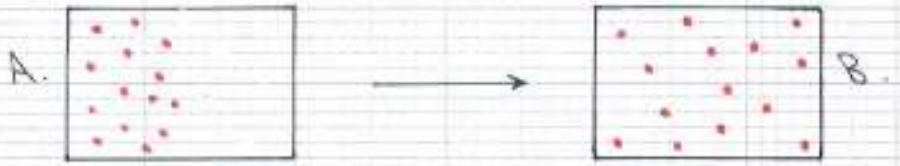
After a time $(\lambda'-\lambda)\Delta t$ the neighbourhood of P, B_0, has been returned to: after finite time the system returns arbitrarily close to an initial point.*

The Poincaré theorem gives rise to a seeming paradox in thermodynamics. Consider a gas contained on one side of a divide in a closed volume:



* There is a theorem of Poker that the addicted gambler believes in which is similar: a good hand will always be dealt again in finite time.

When the divide is removed the molecules will fill out the box:



Poincaré's recurrence theorem tells us that in finite time the system will return arbitrarily close to situation A, with all the gas on one side of the box. It doesn't matter how small the initial containing volume of the gas is the theorem tells us that eventually the system will return arbitrarily close to this configuration!. However the theorem does not indicate how long it will take to do so.

The paradox is given in terms of entropy. The 2nd Law of thermodynamics states that entropy always increases, while Poincaré's theorem indicates that systems will return to very near an earlier entropy after a finite time. The disappointing resolution is that such a recurrence would take longer than the expected lifetime of the universe.

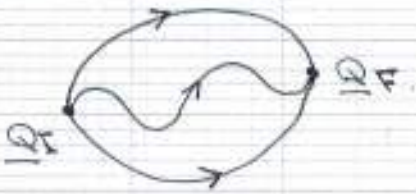
5. The Principle of Least Action.

We consider the notion of an action in this section and demonstrate that both Lagrange's and Hamilton's equations of motion may be derived from a universal concept.

Lagrange's Equations.

Recall that the dynamics of a Lagrangian system are described by a curve in configuration space (an n-dimensional vector space whose vectors are the generalised coordinates, $\begin{pmatrix} q_1 \\ \vdots \\ q_n \end{pmatrix}$).

Suppose we know only the initial, starting point, $\underline{Q}_I$, and the final point, $\underline{Q}_F$, of the system in configuration space:



How do we find the path taken from $\underline{Q}_I$ to $\underline{Q}_F$? We would solve Lagrange's equations and find the correct path $\underline{Q}(t)$. However we can also find the path by assigning a number, called the action, S, to each path, defined as:

$$S[\underline{Q}(t)] = \int_{t_I}^{t_F} L(q_i, \dot{q}_i)\, dt.$$

Then we use the following principle:

Principle of Least Action: The dynamical path of the system is that which extremises the action.

This is identical to Lagrange's equations which we now show.

Consider a slight variation from the path $\underline{Q}(t)$:

$$q_i \to q_i + \delta q_i \quad \Rightarrow \dot{q}_i \to \dot{q}_i + \delta \dot{q}_i$$

Such that the end-points of the path coincide with those of $\underline{Q}(t)$:

i.e. $\delta q_i(t_I) = \delta q_i(t_F) = 0.$

The change in the action due to this variation is:

$$\begin{aligned} \delta S &= \delta\left[\int_{t_I}^{t_F} L(q_i, \dot{q}_i)\, dt\right] \\ &= \int_{t_I}^{t_F} \delta L(q_i, \dot{q}_i)\, dt \end{aligned}$$

$$\therefore \quad \delta S = \int_{t_I}^{t_F} \sum_i \left(\frac{\partial L}{\partial q_i} \delta q_i + \frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i \right) dt.$$

If we integrate the second term by parts we have:

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

Since we fixed the endpoints of the path $\delta q_i(t_I) = \delta q_i(t_F) = 0$ the last term is zero.

The principle of least action states that the dynamical path occurs when the action is extremised i.e. when $\delta S = 0$. This must be true for all variations from the path (i.e. we cannot set $\delta q_i = 0$ to find $\delta S = 0$ we must be more general). It means that $\delta S = 0$ if

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

For any δq_i this is true if $\frac{\partial L}{\partial q_i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = 0$ which we recognise as Lagrange's equations.

We can also derive Hamilton's equations from the principle of least action.

<u>Hamilton's Equations.</u>

As $$H \equiv \sum_i \dot{q}_i p_i - L.$$

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

And the action is:

$$S = \int_{t_I}^{t_F} \left(\sum_i \dot{q}_i p_i - H \right) dt$$

The variation of the action is:

$$\delta S = \int_{t_I}^{t_F} \delta \left(\sum_i \dot{q}_i p_i - H \right) dt.$$

$$= \int_{t_I}^{t_F} \sum_i \left(\delta \dot{q}_i p_i + \dot{q}_i \delta p_i - \frac{\partial H}{\partial q_i} \delta q_i - \frac{\partial H}{\partial p_i} \delta p_i \right) dt.$$

$$= \int_{t_I}^{t_F} \sum_i \left(-\frac{d}{dt}(p_i) \delta q_i + \dot{q}_i \delta p_i - \frac{\partial H}{\partial q_i} \delta q_i - \frac{\partial H}{\partial p_i} \delta p_i \right) dt + \sum_i \left[p_i \delta q_i \right]_{t_I}^{t_F}$$

Where we have integrated by parts all the $\delta \dot{q}_i$ terms. We have:

$$\delta S = \int_{t_I}^{t_F} \sum_i \left(\left(-\dot{p}_i - \frac{\partial H}{\partial q_i} \right) \delta q_i + \left(\dot{q}_i - \frac{\partial H}{\partial p_i} \right) \delta p_i \right) dt + \sum_i \left[p_i \delta q_i \right]_{t_I}^{t_F}$$