Classical Dynamics Notes

Francesco Chotuck

Abstract

This is KCL undergraduate module 5CCM231, instructed by Christopher P. Herzog. The formal name for this class is "Classical Dynamics".

Contents

Basics

1 Scalars & Vectors

Definition 1.1. A **scalar** is a single number, the value of which everyone agrees on.

Definition 1.2. A **vector** in $\mathbb{R}^3$ is a triplet of numbers for example,

$$\boldsymbol{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r^1 \\ r^2 \\ r^3 \end{pmatrix}.$$

> **Note 1.1.** In these notes I shall use index notation $v^a, a = 1, 2, 3$ for the component of a vector $\boldsymbol{v}$. In particular
> $$v = x \quad v^2 = y \quad v^3 = z.$$

Theorem 1.1. For vectors, we can set up a coordinate basis

$$\boldsymbol{v} = x\boldsymbol{e}_x + y\boldsymbol{e}_y + z\boldsymbol{e}_z$$

where the values of v, v^2, v^3 depend on the choice of basis.

1.1 Scalar product

Definition 1.3. The **scalar product** (a.k.a. dot product and inner product) is a map that takes two vectors into a number and is defined as:

$$\boldsymbol{v}_1 \cdot \boldsymbol{v}_2 = x_1 x_2 + y_1 y_2 + z_1 z_2.$$

Theorem 1.2. A geometric interpretation of the scalar product is:

$$\boldsymbol{a} \cdot \boldsymbol{b} = |a| \, |b| \cos \theta.$$

Theorem 1.1

Properties of the scalar product:

1. Symmetric: $\boldsymbol{a} \cdot \boldsymbol{b} = \boldsymbol{b} \cdot \boldsymbol{a}$.

2. Distributive: $\boldsymbol{a} \cdot (\boldsymbol{b} + \boldsymbol{c}) = \boldsymbol{a} \cdot \boldsymbol{b} + \boldsymbol{a} \cdot \boldsymbol{c}$.

3. Positive definite: $\boldsymbol{v} \cdot \boldsymbol{v} \geq 0$ with equality if and only if $\boldsymbol{v} = 0$.

4. If property (3) holds then $|\boldsymbol{v}| = \sqrt{\boldsymbol{v} \cdot \boldsymbol{v}}$.

1.2 Vector product

Theorem 1.3. A plane in three-dimensions has a **normal vector**, i.e. a vector which is orthogonal to every vector in the plane.

Definition 1.4. The **cross product** of two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ is defined (**only** in 3D space), as a vector $\boldsymbol{c}$ that is orthogonal (perpendicular) to both $\boldsymbol{a}$ and $\boldsymbol{b}$, the direction of which is defined by the right-hand rule.

Theorem 1.4. The cross product is defined by the formula

$$\boldsymbol{a} \times \boldsymbol{b} = |\boldsymbol{a}| \, |\boldsymbol{b}| \sin \theta \, \widehat{\boldsymbol{n}}.$$

The vector $\widehat{\boldsymbol{n}}$ is a unit vector perpendicular to the plane containing $\boldsymbol{a}$ and $\boldsymbol{b}$, in the direction indicated by the right-and rule.

Definition 1.5. The right-hand rule is defined as by the figure below:

Figure 1: Right-hand rule
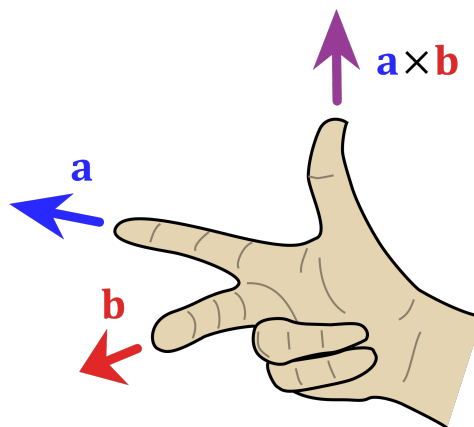
Theorem 1.2

Some fundamental properties of the vector product:

- $\boldsymbol{v} \times \boldsymbol{v} = 0$;

- anti-symmetry: $\boldsymbol{v} \times \boldsymbol{w} = -(\boldsymbol{w} \times \boldsymbol{v})$;

- orthogonality: $\boldsymbol{v} \cdot (\boldsymbol{v} \times \boldsymbol{w}) = \boldsymbol{w} \cdot (\boldsymbol{v} \times \boldsymbol{w}) = 0$.

Theorem 1.5. The value of $|\boldsymbol{a} \times \boldsymbol{b}|$ represents the area of a parallelogram with $\boldsymbol{a}$ as its base and $\boldsymbol{b}$ as the slant height.

1.3   Triple product

1.3.1   Scalar triple product

Definition 1.6. The **scalar triple product** of three vectors:

$$\boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c})$$

which yields a scalar.

Theorem 1.6. Geometrically $\boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c})$ results in the volume (more accurately the absolute value gives the volume) of a parallelepiped with sides defined by $\boldsymbol{a}, \boldsymbol{b}$ and $\boldsymbol{c}$ as labelled in the figure below.
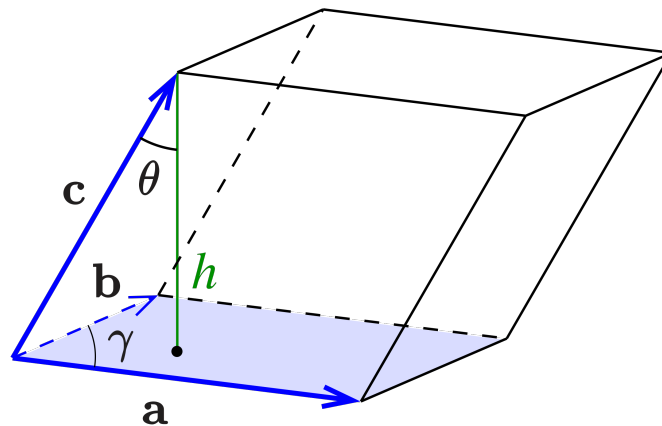
Theorem 1.3

The scalar triple product can be written

$$\boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c}) = \boldsymbol{c} \cdot (\boldsymbol{a} \times \boldsymbol{b}) = \boldsymbol{b} \cdot (\boldsymbol{c} \times \boldsymbol{a}),$$

that is, *cyclic permutations* of the vectors $\boldsymbol{a}, \boldsymbol{b}$ and $\boldsymbol{c}$ in a triple scalar product leaves its value unchanged.

1.3.2 Vector triple product

Definition 1.7. The **triple vector product** is an expression of the form

$$\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c})$$

which yields a vector.

Theorem 1.4

The vector triple vector product can be written as

$$\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = (\boldsymbol{c} \cdot \boldsymbol{a})\boldsymbol{b} - (\boldsymbol{b} \cdot \boldsymbol{a})\boldsymbol{c}.$$

Note 1.2. A helpful way to remember is that the triple vector product is equal to 'CAB-BAC'.

---

2 Matrices

Definition 2.1. (In this course) Matrices are linear maps $M : V \to V$ where V is a vector space. The linearity means that

$$M(a\boldsymbol{r}_1 + b\boldsymbol{r}_2) = aM(\boldsymbol{r}_1) + bM(\boldsymbol{r}_2).$$

> **Theorem 2.1**
>
> The kth entry of $A\boldsymbol{v}$ is the dot product of the kth row of A with $\boldsymbol{v}$:
>
> $$(A\boldsymbol{v})_k = \sum_{j=1}^{n} a_{kj}v_j = a_{k1}v_1 + a_{k2}v_2 + \cdots + a_{kn}v_n.$$

> **Theorem 2.2** (Matrix multiplication)
>
> If $A = (a_{ij})$ and $B = (b_{ij})$ then $C = AB = (c_{ij})$ where
>
> $$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$$
>
> such that $i = 1, \ldots, m$ and $j = 1, \ldots, p$.

Proof. Considering one arbitrary element on position (i, j):

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{i1} & \rightarrow & a_{in} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} b_{11} & \cdots & b_{1j} & \cdots & b_{1p} \\ \vdots & \ddots & \downarrow & \ddots & \vdots \\ b_{n1} & \cdots & b_{nj} & \cdots & b_{np} \end{pmatrix} = \begin{pmatrix} c_{11} & \cdots & c_{1j} & \cdots & c_{1p} \\ \vdots & \ddots & & & \vdots \\ c_{i1} & & c_{ij} & & c_{ip} \\ \vdots & & & \ddots & \vdots \\ c_{m1} & \cdots & c_{mj} & \cdots & c_{mp} \end{pmatrix},$$

with element c_{ij} equal to:

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}.$$

In the sum above, the left index is always i (the ith row of A) and the right index is always j (the jth column of B). The inner indices run from 1 to n so, we can use a summation index l and write this sum as

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}.$$

This gives us the element on position (i, j) in the product matrix $C = AB$ therefore, C is defined by letting $i = 1, \ldots, m$ and $j = 1, \ldots, p$. $\qquad\square$

2.1 Orthogonal matrices

Definition 2.2. An **orthogonal matrix** is one which is equal to its transpose i.e.

$$O^\top = O.$$

> **Theorem 2.3**
>
> Properties of orthogonal matrices:
>
> - $O^\top O = I$;
>
> - Orthogonal matrices leave the scalar product between two vectors invariant i.e. $(O\boldsymbol{v}) \cdot (O\boldsymbol{w}) = \boldsymbol{v} \cdot \boldsymbol{w}$.

Example 2.1

An example of orthogonal matrices are **rotation matrices**:

$$O = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

3 Tensors

3.1 ε-tensor

Definition 3.1. In three dimensions, the ε-tensor (or the Levi-Civita symbol) is defined by

$$\varepsilon_{ijk} = \begin{cases} 1 & \text{if } (i,j,k) = (1,2,3), (2,3,1), (3,1,2) \\ -1 & \text{if } (i,j,k) = (1,2,3), (2,3,1), (3,1,2) \\ 0 & \text{otherwise.} \end{cases}$$

Note 3.1. That is $\varepsilon_{ijk} = 1$ if (i,j,k) are a clockwise (or even) permutation of $(1,2,3)$ whereas, $\varepsilon_{ijk} = -1$ if (i,j,k) are an anti-clockwise (or odd) permutation of $(1,2,3)$ and 0 if any of the indices are the same.

3.2 Kronecker δ function

Definition 3.2. The Kronecker delta function is defined as

$$\delta_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j. \end{cases}$$

Note 3.2. The function, δ_{ij}, can also be thought as the identity tensor.

3.3 Vector identities

Theorem 3.1

The ith component of the resultant vector of the cross product is given by

$$(\boldsymbol{a} \times \boldsymbol{b})_i = \sum_{j=1}^{3} \sum_{k=1}^{3} \varepsilon_{ijk} a_j b_k.$$

Note 3.3. By the ith component we mean

$$(\boldsymbol{a} \times \boldsymbol{b}) = \begin{pmatrix} (\boldsymbol{a} \times \boldsymbol{b})_1 \\ (\boldsymbol{a} \times \boldsymbol{b})_2 \\ (\boldsymbol{a} \times \boldsymbol{b})_3 \end{pmatrix}.$$

4 Derivatives

4.1 The chain rule

Definition 4.1. Given a function $z = f(u_1, u_2, \ldots, u_n)$ the **differential**, dz or df is given by

$$df = \frac{\partial f}{\partial u_1} du_1 + \frac{\partial f}{\partial u_2} du_2 + \cdots + \frac{\partial f}{\partial u_n} du_n.$$

Note 4.1. We can use δ instead of d in differentials.

Theorem 4.1 (Multivariable functions)

Suppose $z = f(u_1(t), u_2(t), \ldots, u_n(t))$ then

$$\frac{\mathrm{d}z}{\mathrm{d}t} = \frac{\partial z}{\partial u_1}\frac{\mathrm{d}u_1}{\mathrm{d}t} + \frac{\partial z}{\partial u_2}\frac{\mathrm{d}u_2}{\mathrm{d}t} + \cdots + \frac{\partial z}{\partial u_n}\frac{\mathrm{d}u_n}{\mathrm{d}t}.$$

Theorem 4.2 (Paths)

Let $f : \mathbb{R}^m \to \mathbb{R}$ and $\boldsymbol{r} : \mathbb{R} \to \mathbb{R}^m$. Then

$$\frac{\mathrm{d}}{\mathrm{d}t} f[\boldsymbol{r}(t)] = \nabla f[\boldsymbol{r}(t)] \cdot \dot{\boldsymbol{r}}(t).$$

5 Coordinate systems

5.1 Polar coordinates
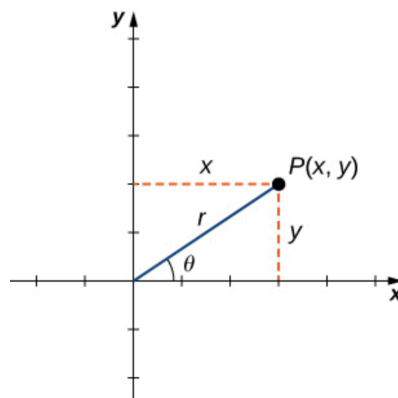
Consider a point (x, y) in the Cartesian coordinate system, in the **polar coordinate system** this point can be represented by an ordered pair (r, θ), where r is the length of distance from the origin to the point and θ is the angle between the positive x-axis and the line segment from the origin to the point; this is better illustrated in the figure above.

Theorem 5.1

Given a point P in the plane with Cartesian coordinates (x, y) and polar coordinates (r, θ), the following hold true

$$x = r\cos\theta \quad \text{and} \quad y = r\sin\theta$$
$$r^2 = x^2 + y^2 \quad \text{and} \quad \tan\theta = \frac{y}{x}.$$

5.2 Cylindrical coordinates
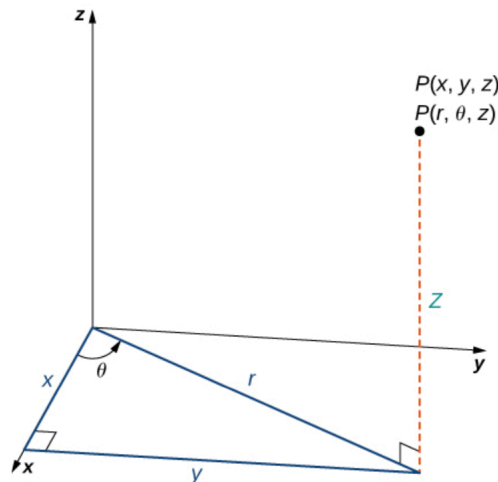
Definition 5.1. In the **cylindrical coordinate system** a point in space is represented by the ordered triple (r, θ, z) where

- (r, θ) are the polar coordinates of the point's projection in the xy-plane;

- z is the usual z-coordinate in the Cartesian coordinate system.

Theorem 5.2

The Cartesian coordinates (x, y, z) and the cylindrical coordinates (r, θ, z) of a point are related as follows:
$$x = r \cos \theta$$
$$y = r \sin \theta$$
$$z = z$$
and
$$r^2 = x^2 + y^2$$
$$\tan \theta = \frac{y}{x}$$
$$z = z.$$

5.3   Spherical coordinates
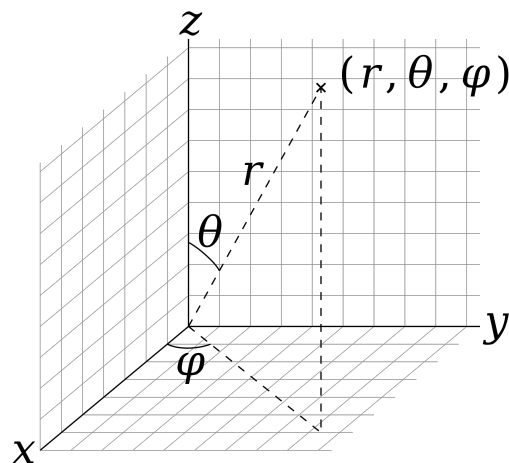
Definition 5.2. In the **spherical coordinate system**, a point in space is represented by the ordered triple (r, θ, ϕ) where

- r is the distance between the point and the origin;

- θ is the angle formed by the positive z-axis and the line segment from the origin to the point;

- ϕ is the same angle used to describe the location in cylindrical coordinates.

> **Theorem 5.3**
> The coordinates (r, θ, ϕ) for
>
> $$r \in (0, \infty), \quad \theta \in (0, \pi) \quad \phi \in (0, 2\pi),$$
>
> are expressed as
> $$x = r \sin \theta \cos \phi$$
> $$y = r \sin \theta \sin \phi$$
> $$z = r \cos \theta.$$
>
> in Cartesian coordinates.

6 Ordinary differential equations

6.1 First order

Theorem 6.1. Explicit equations are solved by direct integration i.e. the solution to

$$\frac{\mathrm{d}x}{\mathrm{d}t} = f(t),$$

is

$$x(t) = \int f(t)\, dt.$$

Theorem 6.2. Separable equations are of the form

$$\frac{\mathrm{d}x}{\mathrm{d}t} = f(t)g(x),$$

they are solved by evaluating the integrals

$$\int \frac{1}{g(x)}\, dx = \int f(t)\, dt.$$

Theorem 6.3. Linear ODEs are of the form

$$\frac{\mathrm{d}x}{\mathrm{d}t} + xf(t) = g(t).$$

They are solved by using an **integrating factor**

$$\mu(t) = e^{\int f(t)\, dt}.$$

6.2 Second order

Definition 6.1. ODEs of the form

$$\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} + a\frac{\mathrm{d}x}{\mathrm{d}t} + bx = \phi(t)$$

where a and b are constant are called second order linear ODEs.

Definition 6.2. If the function $\phi(t)$ is zero the equation is said to be **homogeneous linear**, otherwise it is called **inhomogeneous linear**.

The homogeneous equation is solved with the help of solutions of the auxiliary equation. This method is based on the observation that the equation $\ddot{x} + a\dot{x} + bx = 0$ is solved by functions of the form

$$x(t) = x = e^{\lambda t}$$

provided λ is properly chosen. The choice of λ is obtained by the roots of the substituting the solution of the ODE, using $\dot{x} = \lambda e^{\lambda t}$ and $\ddot{x} = \lambda^2 e^{\lambda t}$, so we get

$$\ddot{x} + a\dot{x} + bx = (\lambda^2 + a\lambda + b)e^{\lambda t} = 0.$$

Since $e^{\lambda t} \neq 0$, we see that $e^{\lambda t}$ is a solution of the ODE provided that

$$p(\lambda) = \lambda^2 + a\lambda + b = 0.$$

Since the auxiliary equation is a quadratic in λ we have three possibilities

(i) $a^2 - 4b \geq 0$, the auxiliary equation has two distinct real solutions λ_1, λ_2. So, it follows that $x_1 = e^{\lambda_1 t}$ and $x_2 = e^{\lambda_2 t}$ therefore, the general solution is a linear combination of these two solutions (x_1 and x_2 are complementary functions so $L(x_1) = L(x_2) = 0$) i.e.
$$x = x(t) = C_1 e^{\lambda_1 t} + C_2 e^{\lambda_2 t}.$$

(ii) $a^2 - 4b = 0$, the auxiliary equation has two repeated roots $\lambda_{1,2} = \lambda$; in this case we that the general solution is

$$x = C_1 t e^{\lambda t} + C_2 e^{\lambda t}.$$

(iii) $a^2 - 4b = 0$, the auxiliary equation has two complex conjugate roots $\lambda_\pm = \mu \pm i\nu$. This means that $x_\pm(t) = e^{(\mu \pm i\nu)t}$ are the solutions and the general solution using case (i) is
$$C_1 e^{(\mu + i\nu)t} + C_2 e^{(\mu - i\nu)t},$$

where C_1 and C_2 are free parameters. This solution is fine but, we often want the general solution in terms of a *real* function, so we have

$$x = x(t) = C_1 e^{\mu t} \cos(\nu t) + C_2 e^{\mu t} \sin(\nu t),$$

where in this case C_1 and C_2 are free *real* parameters.

Finding a particular integral

A guess is often called a "trial function" or an *ansatz* (plural *ansätze*) to make it sound better. This is often the fastest way. The reason this can be done is that the linear operator L in the equation
$$L(x) = \phi,$$

maps polynomials to polynomials, trigonometric functions to trigonometric functions, exponential functions to exponential functions and so on.

$\mathbf{f(x)}$	Substitution
c	γ
$bx + c$	$\beta x + \gamma$
$ax^2 + bx + c$	$\alpha x^2 + \beta x + \gamma$
Qe^{px}	Qe^{px} (or Qxe^{px} or Qx^2e^{px})
$r\cos(\omega x) + s\sin(\omega x)$	$\rho\cos(\omega x) + \sigma\sin(\omega x)$

Note 6.1. IMPORTANT: If there is a multiple of $f(x)$ in the complementary function then the usual choice of substitution will not work. So, multiply the substitution by x until you find a substitution that works. (This is referring to the substitution of Qe^{px})

Remark 6.1. If the inhomogeneous term is a polynomial in this course it is usual to 'guess' a solution of the ODE to be a quadratic polynomial or higher, so that the second derivative has a non-zero term.

6.3   Coupled equations

The system we are considering are equations of the form

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

which we can also write as

$$\dot{\boldsymbol{r}} = M\boldsymbol{r}, \text{ where } \boldsymbol{r} = \begin{pmatrix} x \\ y \end{pmatrix}, M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

As mentioned in the previous section we can write two coupled first order linear ODEs as second order linear ODE; we do this by eliminating y and writing the ODE in terms of x, and we get

$$\ddot{x} - \underbrace{(a + d)}_{t=\mathrm{Tr}(M)}\dot{x} + \underbrace{(ad - bc)}_{\Delta=\det(M)} = 0.$$

We know from the previous section that we should for a solution of the $x = Ae^{\lambda t}$ where λ satisfies

$$\lambda - (a + d)\lambda + (ad - bc) = 0.$$

Substituting $x = Ae^{\lambda t}$ (and after a bit of algebraic manipulation) we have

$$M\boldsymbol{\xi} = \lambda\boldsymbol{\xi}, \text{ where } M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Therefore, to solve these types of ODEs we must find the eigenvectors and their eigenvalue obtained by the characteristic polynomial. Recall from LAG1, that for any real 2×2 matrix M, is similar to one of only three forms of Jordan form matrices so, exactly one of the following holds:

(i) M has two real eigenvalues, λ_1, λ_2, and we can find two linearly independent eigenvectors ($\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$) that span $\mathbb{R}^2$; so the general solution to the matrix differential equation is

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \boldsymbol{r}(t) = C_1\boldsymbol{\xi}_1 e^{\lambda_1 t} + C_2\boldsymbol{\xi}_2 e^{\lambda_2 t}.$$

(ii) M has only one real eigenvalue and only a one–dimensional space of eigenvectors; then one can always fine non-zero vectors $\boldsymbol{\xi}$ and $\boldsymbol{\zeta}$ such that

$$M\boldsymbol{\xi} = \lambda\boldsymbol{\xi}, \quad M\boldsymbol{\zeta} = \lambda\boldsymbol{\zeta} + \boldsymbol{\xi}.$$

That is, $\boldsymbol{\xi}$ is an eigenvector and $\boldsymbol{\zeta}$ is another vector which is not an eigenvector. In this case the general solution is

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \boldsymbol{r}(t) = C_1 \boldsymbol{\xi}_1 e^{\lambda t} + C_2 (\boldsymbol{\zeta} + t\boldsymbol{\xi}) e^{\lambda t}.$$

(iii) M has two distinct complex conjugate eigenvalues $(\alpha \pm i\beta)$ with complex conjugate eigenvectors. Then we can find a non-zero complex vector $\boldsymbol{\xi}$ such that $\boldsymbol{\xi} = \boldsymbol{\mu} + i\boldsymbol{\nu}$ and $\boldsymbol{\xi}^* = \boldsymbol{\mu} - i\boldsymbol{\nu}$ are corresponding eigenvectors (where $\alpha, \beta, \boldsymbol{\mu}$ and $\boldsymbol{\nu}$ are real) such that

$$M\boldsymbol{\mu} = \alpha\boldsymbol{\mu} - \beta\boldsymbol{\nu}, \quad M\boldsymbol{\nu} = \beta\boldsymbol{\mu} + \alpha\boldsymbol{\nu}.$$

In this case the general solution is

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \boldsymbol{r}(t) = C_1 \left(\boldsymbol{\mu} \cos\left(\beta t\right) - \boldsymbol{\nu} \sin\left(\beta t\right) \right) e^{\alpha t} + C_2 \left(\boldsymbol{\mu} \sin\left(\beta t\right) + \boldsymbol{\nu} \cos\left(\beta t\right) \right) e^{\alpha t}.$$

Another way to write this is

$$\boldsymbol{r}(t) = \Re\left(D\boldsymbol{\xi} e^{\lambda t} \right),$$

where $D = C_1 + iC_2$ is an arbitrary complex number.

Recall that in each case one can find a real matrix P such that $P^{-1}MP$ is in one of the real Jordan canonical forms. The Jordan canonical form of each case are as follows:

(i)

$$P^{-1}MP = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \quad P = (\boldsymbol{\xi}_1 : \boldsymbol{\xi}_2).$$

(ii)

$$P^{-1}MP = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}, \quad P = (\boldsymbol{\xi}_1 : \boldsymbol{\zeta}).$$

(iii)

$$P^{-1}MP = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}, \quad P = (\boldsymbol{\mu} : \boldsymbol{\nu}).$$

Newtonian mechanics of a single particle

7 Newton's Laws

7.1 Newton's Laws of motion

Definition 7.1. A **particle** is an object whose shape and structure does not affect the dynamics of the system.

Theorem 7.1 (Newtow's Laws of motion)

The laws of motion:

  (I) A particle will stay at rest or move with constant velocity along a straight line unless acted upon by external force.

 (II) The rate of change of momentum of a particle is equal to and in the direction of, the **net force** acting on it.

(III) Every action has an opposite and equal reaction.

Note 7.1. I want to note here that **velocity is a vector** therefore the particle is moving along on a straight line if the velocity is constant.

Definition 7.2. An **inertial frame** is a frame in which Newton's First law [NI] holds true.

7.2 Linear momentum

Definition 7.3. If a particle has mass m and velocity $\dot{\boldsymbol{r}}$, then $\boldsymbol{p}$, its **linear momentum** is defined to be

$$\boldsymbol{p} = m\dot{\boldsymbol{r}}$$

Theorem 7.2 (Linear momentum principle)

Mathematically, Newton's Second Law states:

$$\boldsymbol{F} = \dot{\boldsymbol{p}},$$

where $\boldsymbol{p}$ denotes linear momentum and $\boldsymbol{F}$ the total external force acting on a particle.

Remark 7.1. The more familiar $\boldsymbol{F} = m\ddot{\boldsymbol{r}}$ arises when m is constant; in this module we will also consider cases when m is not constant.

Theorem 7.1 (Conservation of linear momentum)**.** In absence of a force acting on a particle the quantity of momentum is conserved.

Proof. In absence of any force we have that $\boldsymbol{F} = \boldsymbol{0}$ therefore,

$$\boldsymbol{F} = \dot{\boldsymbol{p}} = \boldsymbol{0}.$$

$\square$

> **Note 7.2.** Generally speaking, in classical mechanics to check if a quantity is 'conserved' it suffices to check if its time derivative is 0, which implies the quantity is constant for all time i.e. it is conserved.

7.2.1 Collision theory
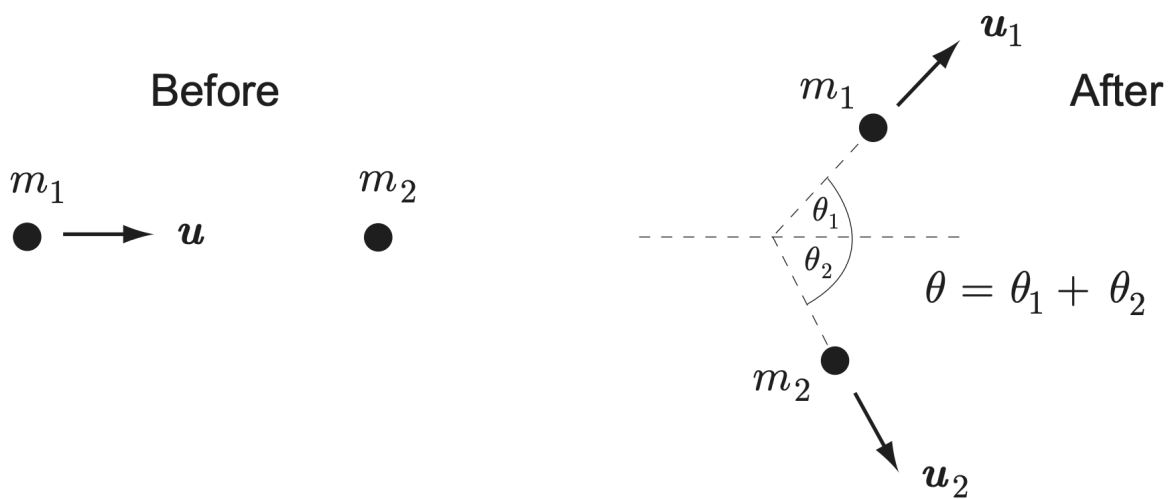
Consider the collision shown in the figure above. A particle of mass m_1 and initial velocity $\boldsymbol{u}$ is incident upon another particle, which has mass m_2 and is at rest. Suppose after the collision the masses of the particles does not change and, their new velocities are now $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ respectively. Since the two particles are in an isolated system (i.e. there are no external forces acting on the system), their total linear momentum is conserved, i.e.

$$m_1 \boldsymbol{u}_1 = m_1 \boldsymbol{u}_1 + m_2 \boldsymbol{u}_2.$$

By this linear relation it implies that these three velocities must lie in the same plane thus, scattering processes are **two-dimensional**.

> **Note 7.3.** This process is not energy preserving unless it is an elastic collision where all the energy is transferred.

7.3 Friction

Definition 7.4. Friction is the force resisting the relative motion of solid surfaces, fluid layers, and material elements sliding against each other.

Theorem 7.3

The frictional force

$$\boldsymbol{F}_F = -\nu\dot{\boldsymbol{r}}.$$

Where $\nu > 0$ is the friction coefficient. The minus sign indicates that the friction acts in the opposite direction to the velocity.

8 Angular motion

8.1 Angular momentum and torque

Definition 8.1. Suppose a particle P of mass m has position vector $\boldsymbol{r}$ and velocity $\dot{\boldsymbol{r}}$. Then $\boldsymbol{L}$, the **angular momentum** of P about a point is defined to be

$$\boldsymbol{L} = \boldsymbol{r} \times (m\dot{\boldsymbol{r}})$$
$$= \boldsymbol{r} \times \boldsymbol{p}.$$

The vector $\boldsymbol{L}$ always points in a direction orthogonal to both $\boldsymbol{r}$ and $\boldsymbol{p}$.

Remark 8.1. In angular motion the linear momentum is **not** (necessarily) conserved however, angular momentum is conserved.

Definition 8.2. The **torque**, $\boldsymbol{N}$, of a particle acting upon it is defined by

$$\boldsymbol{N} = \boldsymbol{r} \times \boldsymbol{F}.$$

Note 8.1. Torque is analogous to the force acting on a particle in linear motion.

Remark 8.2. In some texts torque is known as the **moment of a force** (about a point).

Remark 8.3. Both $\boldsymbol{L}$ and $\boldsymbol{N}$ can take different value depending on where the 'origin' is defined.

Theorem 8.1 (Angular momentum principle)

If we have a point particle with fixed mass m and momentum $\boldsymbol{p} = m\dot{\boldsymbol{r}}$ then

$$\boldsymbol{N} = \dot{\boldsymbol{L}}.$$

8.2 Conservation of angular momentum

Theorem 8.2

In any motion of a system in which there is no torque acting on the particle, the angular momentum (about any point) is conserved.

Proof. Recall that $\boldsymbol{N} = \dot{\boldsymbol{L}}$, so if $\boldsymbol{N} = \boldsymbol{0}$ then $\dot{\boldsymbol{L}} = 0$ therefore, $\boldsymbol{L} = $ constant for all time. $\square$

> **Corollary 8.1**
>
> Conservation of angular momentum implies that the motion of the particle in on a plane.

Proof. As $\boldsymbol{r} \cdot \boldsymbol{L} = \boldsymbol{0}$ the position vector is orthogonal to tha angular momentum vector. Since $\boldsymbol{L}$ is constant $\boldsymbol{r}$ lies in a plane perpendicular to $\boldsymbol{L}$ for all time. $\square$

Proof. Angular momentum is given by a vector $\boldsymbol{L} = \boldsymbol{r} \times m\dot{\boldsymbol{r}}$. As the vector cross product is used in the definition of angular momentum it implies that both $\boldsymbol{r}$ and $\dot{\boldsymbol{r}}$ should be orthogonal to $\boldsymbol{L}$. Therefore, both $\boldsymbol{r}$ and $\dot{\boldsymbol{r}}$ lie in the plane orthogonal to $\boldsymbol{L}$. As $\boldsymbol{L}$ is constant for all time so is the plane. $\square$

8.3   Circular motion

In circular motion a particle moves in a circle thus, $|\boldsymbol{r}|$ is constant (for all time). We can choose $\boldsymbol{r}$ as

$$\boldsymbol{r} = \begin{pmatrix} r\cos\theta(t) \\ r\sin\theta(t) \\ 0 \end{pmatrix}.$$

Circular motion means that $r = |\boldsymbol{r}|$ is constant but note that $\theta(t)$ is not necessarily constant for all time. If we restrict $\dot{\theta}(t) = \omega$ to be a constant we have the following definitions.

Definition 8.3. The **angular velocity**, ω, is defined to be

$$\omega = \dot{\theta}$$

and $|\omega|$ is called the **angular speed**.

Definition 8.4. The **angular velocity vector** of a body is defined to be

$$\boldsymbol{\omega} = \pm\omega\hat{\boldsymbol{n}},$$

where the sign is taken to be $\pm$ on whether the sense of rotation (relative to the vector $\hat{\boldsymbol{n}}$) is right or left-handed. The direction of $\boldsymbol{\omega}$ is defined by the right-hand rule illustrated by the figure below.
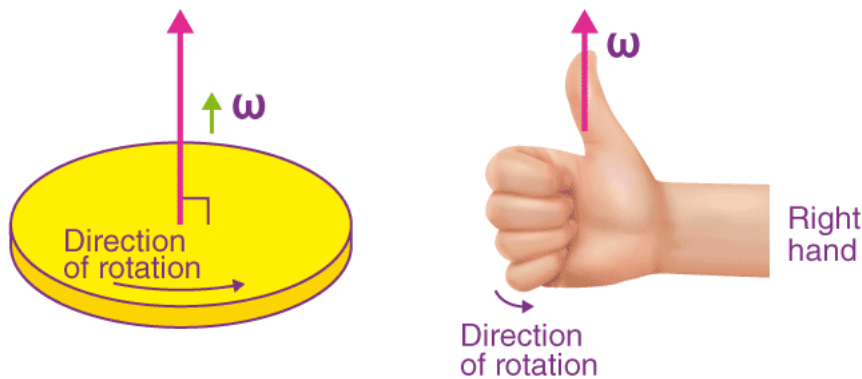
Theorem 8.3

In circular motion there is a force pointing inwards to the centre of motion defined by

$$\boldsymbol{F} = -m\omega^2 \boldsymbol{r},$$

where m is the mass of the particle.

9   Central forces

Definition 9.1. If a force, $\boldsymbol{F}(\boldsymbol{r})$, is directed towards or away from a fixed point, it is called a **central force**. Mathematically,

$$\boldsymbol{F} \propto \boldsymbol{r},$$
$$\boldsymbol{F} = \boldsymbol{r} f(\boldsymbol{r}),$$

where $f(\boldsymbol{r})$ is a scalar quantity.

Remark 9.1. Some texts may define a central force as a force that is directed towards or away from a point; it has the form

$$\boldsymbol{F} = F(|\boldsymbol{r}|)\hat{\boldsymbol{r}}.$$

Note 9.1. Since $\boldsymbol{F} = k\boldsymbol{r}$ for some scalar k we can deduce that central forces act on the line connecting the two bodies.

Theorem 9.1. For a *conservative force* that is derived from a $V(\boldsymbol{r})$ potential, then it being a *central force* implies that V is only a function of $|\boldsymbol{r}|$ (for some choice of the origin) i.e. $V = V(|\boldsymbol{r}|)$.

Theorem 9.1

The angular momentum is conserved for a central force (i.e. the torque vanishes).

Proof.

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{L} = \frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{r} \times \boldsymbol{p})$$

$$= \dot{\boldsymbol{r}} \times \boldsymbol{p} + \boldsymbol{r} \times \dot{\boldsymbol{p}}$$

$$= \dot{\boldsymbol{r}} \times m\dot{\boldsymbol{r}} + \boldsymbol{r} \times \boldsymbol{F}$$

$$= \boldsymbol{0} + \boldsymbol{r} \times \boldsymbol{r}f(\boldsymbol{r})$$

$$= \boldsymbol{r} \times k\boldsymbol{r}$$

$$= \boldsymbol{0}.$$

We take $k = f(\boldsymbol{r})$ to emphasise that it is a scalar quantity. $\square$

10   Energy

10.1   Defining energy

Definition 10.1. The **kinetic energy** T of a non-rotating body of mass m moving with speed $\dot{\boldsymbol{r}}$ is the scalar quantity

$$T = \frac{1}{2}m(\dot{\boldsymbol{r}} \cdot \dot{\boldsymbol{r}}) = \frac{1}{2}m\left|\dot{\boldsymbol{r}}\right|^2.$$

Definition 10.2. The scalar quantity

$$W = \int_{\boldsymbol{r}(t_1)}^{\boldsymbol{r}(t_2)} \boldsymbol{F} \cdot d\boldsymbol{r}$$

$$= \int_{t_1}^{t_2} \boldsymbol{F} \cdot \dot{\boldsymbol{r}}\, dt$$

is called the **work done** by the force $\boldsymbol{F}$ during the time interval $[t_1, t_2]$.

Theorem 10.1 (Energy principle for a particle)

In any motion of a particle, the increase in kinetic energy of the particle in a given time interval is **equal** to the total work done by the applied forces during this time interval. That is, over the time interval $[t_1, t_2]$ we have

$$T_2 - T_1 = \int_{t_1}^{t_2} \boldsymbol{F} \cdot \dot{\boldsymbol{r}}\, dt$$

where T_1 and T_2 are the kinetic energy of a particle at times t_1 and t_2 respectively.

Note 10.1. Mathematically we can interpret the theorem as $\Delta T = \Delta W$.

Proof. Suppose a particle P of mass m moves under the influence of a force $\boldsymbol{F}$. Then its equation of motion is

$$m\ddot{\boldsymbol{r}} = \boldsymbol{F}.$$

We place no restriction on $\boldsymbol{F}$ i.e. it may depend on the position of P, the velocity of P, the time etcetera; if more than one force is acting on P, then $\boldsymbol{F}$ means the *resultant*

vector of these forces. On taking the scalar dot product of both sides of the equation of motion with $\dot{\boldsymbol{r}}$, we obtain

$$m\dot{\boldsymbol{r}} \cdot \ddot{\boldsymbol{r}} = \boldsymbol{F} \cdot \dot{\boldsymbol{r}}.$$

Recall that by [NII] we can express $\boldsymbol{F} = \dot{\boldsymbol{p}}$. Thus, we can write

$$
\begin{aligned}
\boldsymbol{F} \cdot \dot{\boldsymbol{r}} &= \dot{\boldsymbol{p}} \cdot \dot{\boldsymbol{r}} \\
&= m\ddot{\boldsymbol{r}} \cdot \dot{\boldsymbol{r}} \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}m\,|\dot{\boldsymbol{r}}|^2\right) \\
&= \frac{\mathrm{d}T}{\mathrm{d}t}.
\end{aligned}
$$

Therefore,

$$\frac{\mathrm{d}T}{\mathrm{d}t} = \boldsymbol{F} \cdot \dot{\boldsymbol{r}};$$

integrating this equation over the time interval $[t_1, t_2]$, we obtain

$$T_2 - T_1 = \int_{t_1}^{t_2} \boldsymbol{F} \cdot \dot{\boldsymbol{r}}\, dt$$

where T_1 and T_2 are the kinetic energies of P at times t_1 and t_2 respectively. $\qquad\square$

10.2   Conservative forces

Definition 10.3. If the *work done* by a force is path independent, such force is called **conservative**.

Theorem 10.2

If $\boldsymbol{F}$ is a conservative force then it can be expressed in the form

$$\boldsymbol{F} = -\nabla V(\boldsymbol{r}),$$

where $V(\boldsymbol{r})$ is a scalar function of position. The function V is said to be the **potential energy** function for $\boldsymbol{F}$.

Proof. Suppose $\boldsymbol{F} = -\nabla V(\boldsymbol{r})$ is a conservative force acting on a particle over the path $\mathcal{C} : \boldsymbol{r}(t)$ for $t \in [t_1, t_2]$. We prove the work done by such $\boldsymbol{F}$ is *path independent*. The work done by $\boldsymbol{F}$ is

$$
\begin{aligned}
W &= \int_{\mathcal{C}} \boldsymbol{F} \cdot d\boldsymbol{r} \\
&= \int_{\mathcal{C}} -\nabla V(\boldsymbol{r}) \cdot d\boldsymbol{r} \\
&= -\int_{\boldsymbol{r}(t_1)}^{\boldsymbol{r}(t_2)} \nabla V(\boldsymbol{r}) \cdot d\boldsymbol{r}.
\end{aligned}
$$

Since $\dot{\boldsymbol{r}} = \dfrac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t}$ we have that

$$
\begin{aligned}
W &= -\int_{t_1}^{t_2} \nabla V(\boldsymbol{r}) \cdot \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t}\, dt \\
&= -\int_{t_1}^{t_2} \nabla V(\boldsymbol{r}) \cdot \dot{\boldsymbol{r}}\, dt.
\end{aligned}
$$

By the chain rule (for paths), we know that $\dfrac{\mathrm{d}}{\mathrm{d}t}V(\boldsymbol{r}) = \nabla V(\boldsymbol{r}) \cdot \dot{\boldsymbol{r}}$ so,

$$
\begin{aligned}
W &= -\int_{t_1}^{t_2} \frac{\mathrm{d}}{\mathrm{d}t}V(\boldsymbol{r})\, dt \\
&= V(\boldsymbol{r}(t_1)) - V(\boldsymbol{r}(t_2)) \\
&= V(\boldsymbol{r}_1) - V(\boldsymbol{r}_2).
\end{aligned}
$$

Therefore, the work done by a particle only depends on its endpoints and not the path taken. □

Theorem 10.3

If $\boldsymbol{F}$ is conservative then $\nabla \times \boldsymbol{F} = \boldsymbol{0}$.

> **Note 10.2.** This is because 'curl(grad)$= 0$'.

10.3　Conservation of energy

Definition 10.4. The sum of the kinetic and potential energy of a particle is called the **total energy**, i.e.

$$
E = T + V.
$$

Theorem 10.4

The total energy of a conserved system is conserved i.e. the total energy of a system for which $\boldsymbol{F} = -\nabla V(\boldsymbol{r})$ is constant for all time.

Proof. We need to show E is constant for all time hence, it suffices to show that $\dot{E} = 0$.

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}E &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}m\,|\dot{\boldsymbol{r}}|^2 + V(\boldsymbol{r}) \right) \\
&= m\dot{\boldsymbol{r}} \cdot \ddot{\boldsymbol{r}} + \nabla V(\boldsymbol{r}) \cdot \dot{\boldsymbol{r}} \\
&= (\boldsymbol{F} + \nabla V(\boldsymbol{r})) \cdot \dot{\boldsymbol{r}} \\
&= 0.
\end{aligned}
$$

Since, $\boldsymbol{F} = -\nabla V(\boldsymbol{r})$. □

10.4   Effective potentials

Definition 10.5. If a particle has conserved angular momentum and conserved energy then the **effective potential** is defined as the quantity

$$V_{\text{eff}}(r) = \frac{\ell^2}{2mr^2} + V(r),$$

where

- l is the angular momentum of the particle;

- m is the mass of the particle and

- r is the distance from the (defined) origin.

Proof. The angular momentum of the particle in motion is conserved thus, the motion of the particle is restricted to a plane, i.e. the plane orthogonal to $\boldsymbol{L}$. We choose coordinates where the plane is

$$\boldsymbol{r} = \begin{pmatrix} r\cos\theta(t) \\ r\sin\theta(t) \\ 0 \end{pmatrix}.$$

Therefore, the angular momentum is

$$\boldsymbol{L} = \boldsymbol{r} \times m\dot{\boldsymbol{r}}$$
$$= \begin{pmatrix} 0 \\ 0 \\ mr^2\dot{\theta} \end{pmatrix}.$$

Since, $\boldsymbol{L}$ is conserved we can fix

$$\ell = mr^2\dot{\theta} \quad \Rightarrow \quad \dot{\theta} = \frac{\ell}{mr^2}$$

where l is a constant. The energy is also conserved thus, we can write

$$E_0 = \frac{1}{2}m\,|\dot{\boldsymbol{r}}|^2 + V(r)$$
$$= \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2\right) + V(r)$$
$$= \frac{1}{2}m\dot{r}^2 + \frac{\ell^2}{2mr^2} + V(r)$$
$$= \frac{1}{2}m\dot{r}^2 + V_{\text{eff}}(r),$$

where

$$V_{\text{eff}}(r) = \frac{\ell^2}{2mr^2} + V(r).$$

$\square$

11 Planetary motion

11.1 The Law of Gravitation

Theorem 11.1 (Newton's universal law of gravitation)

There exists a force of attraction between two bodies given by

$$\boldsymbol{F} = -\frac{G_N M m}{|\boldsymbol{r}|^2}\hat{\boldsymbol{r}},$$

where $\boldsymbol{r} = \boldsymbol{r}_2 - \boldsymbol{r}_1$ i.e. the distance between the two bodies, the vector $\hat{\boldsymbol{r}} = \frac{\boldsymbol{r}}{|\boldsymbol{r}|}$ and G_N is Newton's constant.

Note 11.1. The unit vector $\hat{\boldsymbol{r}}$ indicates the direction of the force, $\boldsymbol{F}$.

Remark 11.1. The force illustrated by the law of gravitation is a central force.

Theorem 11.2

A planet orbiting the sun has potential

$$V(|\boldsymbol{r}_1 - \boldsymbol{r}_2|) = -\frac{G_N M m}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|}$$

and effective potential

$$V_{\text{eff}} = \frac{\ell^2}{2mr^2} - \frac{G_N M m}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|}.$$

11.2 Kepler's Laws

Theorem 11.3

The laws of planetary motion:

(KI) The planets move in an ellipse with the sun at one focus.

(KII) The line joining a given planet to the sun sweeps out equal areas over equal times.

(KIII) The square of a planet's orbital period is proportional to the cube of the semi-major axis.

Note 11.2. From Kepler's Second Law, [KII], we deduce a planet moves slower when it is close to the Sun and faster when it is away from the Sun.

> **Note 11.3.** Mathematically (KIII) can be interpreted as $T^2 \propto r^3$.

Proof. The proofs of Kepler's laws

- Proof of [KI].

- Proof of [KII].

Restrict the motion to a plane

$$\boldsymbol{r} = \begin{pmatrix} r\cos\theta \\ r\sin\theta \\ 0 \end{pmatrix}$$

then, after an infinitesimal time step dt the plane sweetps out a triangle (as shown in the figure above)

$$dA = \frac{1}{2}(r)(rd\theta) = \frac{1}{2}r^2 d\theta.$$

Therefore,

$$\frac{\mathrm{d}A}{\mathrm{d}t} = \frac{1}{2}r^2\dot{\theta}.$$

On the other hand,

$$\boldsymbol{L} = m \begin{pmatrix} r\cos\theta \\ r\sin\theta \\ 0 \end{pmatrix} \times \begin{pmatrix} \dot{r}\cos\theta - r\dot{\theta}\sin\theta \\ \dot{r}\sin\theta + r\dot{\theta}\cos\theta \\ 0 \end{pmatrix}$$

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

So, we have

$$\frac{\mathrm{d}A}{\mathrm{d}t} = \frac{1}{2m}\left|\boldsymbol{L}\right|.$$

- Proof of [KIII]. The area of the ellipse is also given by

$$\text{Area of ellipse} = \int_0^T \frac{\mathrm{d}A}{\mathrm{d}t}\,dt$$

$$= \int_0^T \frac{1}{2m}\left|\boldsymbol{L}\right|\,dt$$

$$= \frac{T}{2m}\left|\boldsymbol{L}\right|$$

since $|\boldsymbol{L}|$ is constant. Given that the area of the ellipse is proportional to $R_{\text{smaj}}^{\frac{3}{2}}$ where R_{smaj} is the semi-major axis of the ellipse, we can write

$$\text{Area of ellipse} \propto R_{\text{smaj}}^{\frac{3}{2}}$$
$$\Rightarrow T^2 \propto R_{\text{smaj}}^3.$$

$\square$

Theorem 11.4
ADD PROOFS

---

12 Solving one dimensional dynamics

Conservation of energy is enough information for us to solve for the motion of a particle in one-dimension with conservative forces. In this case the position vector $\boldsymbol{r}$ is just a scalar $r \in \mathbb{R}$. Let the potential be $V(r)$ so that the total energy is

$$E = \frac{1}{2}m \left(\frac{\mathrm{d}r}{\mathrm{d}t} \right)^2 + V(r).$$

Consider the reformulation of the energy equation:

$$\left(\frac{\mathrm{d}r}{\mathrm{d}t} \right)^2 = \frac{2}{m}(E - V(r)).$$

From this reformulation we gather that $E \geq V(r)$ as there is no such thing as negative kinetic energy. Furthermore, $E = V(r)$ if and only if $\dfrac{\mathrm{d}r}{\mathrm{d}t} = 0$ meaning the particle is at rest. Conceptually, regardless of the origin of the system, we can think of it as a particle moving in a potential $V(r)$ (we can imagine $V(r)$ being the height of a hill) then we can think of the particle moving along the hill.
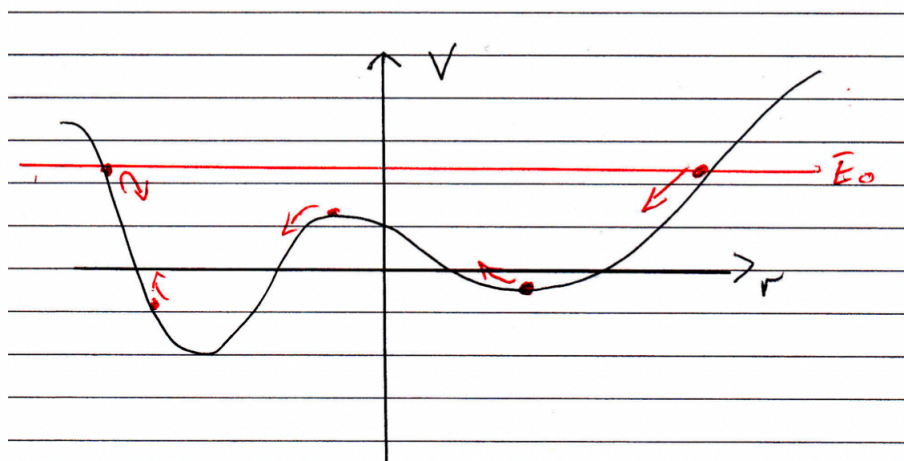
Figure 2: Potential Graph

Given that the total energy is fixed along the motion we can fix $E = E_0$, a constant. We can rewrite the energy equation as

$$\frac{\mathrm{d}r}{\mathrm{d}t} = \pm\sqrt{\frac{2}{m}(E_0 - V(r))}$$

$$dt = \pm\frac{dr}{\sqrt{\frac{2}{m}(E_0 - V(r))}}$$

$$t - t_0 = \pm\int_{r(t)}^{r(t_0)} \frac{1}{\sqrt{\frac{2}{m}(E_0 - V(r'))}} \, dr'$$

Definition 12.1. A dynamical system reached a **turning point** precisely when $E_0 = V(r)$.

At a turning point all the energy is potential energy, i.e. **no** kinetic energy. Therefore, the particle is at rest at that point; such point can be illustrated by the left-most point in the figure above.

Multi-particle systems

13 Configuration of a multi-particle system

Definition 13.1. A **multi-particle system** may consist of any number of particles $P_1, P_2, \ldots, P_n$, with masses $m_1, m_2, \ldots, m_n$ respectively.

Definition 13.2. External forces are those originating from outside the system, denoted by $\boldsymbol{F}_i^{\text{ext}}$ i.e. the external force acting on the ith particle.

> **Remark 13.1.** In the case of a single particle system, these are the only forces that act on the particle.

Definition 13.3. Internal forces (or inter particle forces) are interactions between the particles in the system. These forces are denoted by $\boldsymbol{F}_{ij}^{\text{int}}$ i.e. the force induced by the jth particle on the ith particle.

Theorem 13.1. Newton's Third Law, [NIII], implies that $\boldsymbol{F}_{ij}^{\text{int}} = -\boldsymbol{F}_{ji}^{\text{int}}$ and also that $\boldsymbol{F}_{ii}^{\text{int}} = \boldsymbol{0}$.

13.1 Centre of mass

Definition 13.4. The **centre of mass** of this system of particles is the **point of space** whose position vector $\boldsymbol{R}$ is defined by

$$\boldsymbol{R} = \frac{m_1\boldsymbol{r}_1 + m_2\boldsymbol{r}_2 + \cdots + m_N\boldsymbol{r}_N}{m_1 + m_2 + \cdots m_N} = \frac{\sum_{i=1}^{N} m_i\boldsymbol{r}_i}{\sum_{i=1}^{N} m_i} = \frac{\sum_{i=1}^{N} m_i\boldsymbol{r}_i}{M},$$

where M is the **total mass** of the system.

14 Energy

14.1 Defining energy

Definition 14.1. The **total kinetic energy** of a multi-particle system is defined as the sum of the kinetic energy of the individual particles:

$$T = \sum_{i=1}^{N} \frac{1}{2} m_i \left| \dot{\boldsymbol{r}}_i \right|^2$$

Definition 14.2. The **total work done** is defined as the sum of the of individual work done by each particle:

$$W = \sum_{i=1}^{N} \int_{t_1}^{t_2} \boldsymbol{F}_i \cdot \dot{\boldsymbol{r}}_i \, dt$$

> **Theorem 14.1** (Energy principle for a multi-particle system)
>
> In any motion of a system, the increase in the total kinetic energy of the system in a given time interval is equal to the total work done by all the external and internal forces during this time interval. That is, over the time interval $[t_1, t_2]$ we have
>
> $$\sum_{i=1}^{N} [T_i(t_2) - T_i(t_1)] = \sum_{i=1}^{N} \int_{t_1}^{t_2} \boldsymbol{F}_i \cdot \dot{\boldsymbol{r}}_i \, dt,$$
>
> where $T_i(t_1)$ and $T_i(t_2)$ are the kinetic energies of the ith particle at time t_1 and t_2 respectively.

Proof.

$$\begin{aligned}
W &= \sum_{i=1}^{N} \int_{t_1}^{t_2} \boldsymbol{F}_i \cdot \dot{\boldsymbol{r}}_i \, dt \\
&= \sum_{i=1}^{N} \int_{t_1}^{t_2} m_i \ddot{\boldsymbol{r}}_i \cdot \dot{\boldsymbol{r}}_i \, dt \\
&= \sum_{i=1}^{N} \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{1}{2} m_i \left| \dot{\boldsymbol{r}}_i \right|^2 \right) dt \\
&= \sum_{i=1}^{N} [T_i(t_2) - T_i(t_1)].
\end{aligned}$$

$\square$

> **Corollary 14.1**
>
> If the system is under the influence of conservative forces then the work done is path independent.

14.2 Conservation of energy

Definition 14.3. In a multi-particle systems **conservative forces** are defined as

- external forces: $\boldsymbol{F}_i^{\text{ext}} = -\nabla_i V_i(\boldsymbol{r}_1, \dots, \boldsymbol{r}_N)$;

- internal forces: $\boldsymbol{F}_{ij}^{\text{int}} = -\nabla_i V_{ij}(\boldsymbol{r}_1, \dots, \boldsymbol{r}_N)$,

where $\nabla_i \equiv \dfrac{\partial}{\partial \boldsymbol{r}_i}$.

> **Theorem 14.2**
>
> If the internal and external forces are conservative then the **total energy** is defined as
>
> $$E = \sum_{i=1}^{N} \frac{1}{2} m_i \left| \dot{\boldsymbol{r}}_i \right|^2 + \sum_{i=1}^{N} V_{\text{ext}} + \sum_{j=1}^{N} \sum_{i=1}^{N} V_{ij}^{\text{int}}.$$

> **Theorem 14.3**
>
> If a multi-particle system is under the influence of *conservative forces* then the total energy of the system is conserved.

Note 14.1. The total energy is constant for all time as, $\dot{E} = 0$.

Proof. Recall that

$$E = \sum_{i=1}^{N} \frac{1}{2} m_i \left| \dot{\boldsymbol{r}}_i \right|^2 + \sum_{i=1}^{N} V_{\text{ext}} + \sum_{j=1}^{N} \sum_{i=1}^{N} V_{ij}^{\text{int}}.$$

Therefore,

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} E &= \sum_{i=1}^{N} m_i \dot{\boldsymbol{r}}_i \cdot \ddot{\boldsymbol{r}}_i + \sum_{i=1}^{N} \nabla_i V^{\text{ext}} \cdot \dot{\boldsymbol{r}}_i + \sum_{j=1}^{N} \sum_{i=1}^{N} \nabla_i V_{ij}^{\text{int}} \cdot \dot{\boldsymbol{r}}_i \\
&= \sum_{i=1}^{N} \dot{\boldsymbol{r}}_i \cdot \left(m_i \ddot{\boldsymbol{r}} + \nabla_i V^{\text{ext}} + \sum_{j=1}^{N} \nabla_i V_{ij}^{\text{int}} \right) \\
&= \sum_{i=1}^{N} \dot{\boldsymbol{r}}_i \cdot \left(m_i \ddot{\boldsymbol{r}} - \boldsymbol{F}_i^{\text{ext}} - \sum_{j=1}^{N} \boldsymbol{F}_{ij}^{\text{int}} \right) \\
&= 0.
\end{aligned}$$

$\square$

15 Angular motion

15.1 Linear momentum

Definition 15.1. For a multi-particle system consisting of particles $P_1, P_2, \ldots, P_n$, with masses $m_1, m_2, \ldots, m_n$ and velocities $\dot{\boldsymbol{r}}_1, \dot{\boldsymbol{r}}_2, \ldots, \dot{\boldsymbol{r}}_N$ respectively. The **total linear momentum**, $\boldsymbol{P}$, of the system to be the vector sum of the linear momenta of the individual particles, that is,

$$\boldsymbol{P} = \sum_{i=1}^{N} \boldsymbol{p}_i = \sum_{i=1}^{N} m_i \dot{\boldsymbol{r}}_i.$$

Theorem 15.1. Newton's Second Law, [NII], implies that the **total external force** of a system can be written in terms of linear momentum in the form

$$\boldsymbol{F}^{\text{ext}} = \frac{\mathrm{d}}{\mathrm{d}t} \boldsymbol{P} = \dot{\boldsymbol{P}}.$$

Proof. The momentum of the ith particle is given by

$$\frac{\mathrm{d}}{\mathrm{d}t} \boldsymbol{p}_i = \boldsymbol{F}_i^{\text{ext}} + \sum_{j=1}^{N} \boldsymbol{F}_{ij}^{\text{int}}.$$

Therefore, summing the individual momenta to obtain the total momentum, we have

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} \sum_{i=1}^{N} \boldsymbol{p}_i &= \sum_{i=1}^{N} \boldsymbol{F}_i^{\text{ext}} + \sum_{i=1}^{N} \sum_{j=1}^{N} \boldsymbol{F}_{ij}^{\text{int}} \\
&= \sum_{i=1}^{N} \boldsymbol{F}_i^{\text{ext}} - \sum_{i=1}^{N} \sum_{j=1}^{N} \boldsymbol{F}_{ji}^{\text{int}} \quad \left(\text{using } \boldsymbol{F}_{ij} = -\boldsymbol{F}_{ji}\right) \\
&= \sum_{i=1}^{N} \boldsymbol{F}_i^{\text{ext}} - \sum_{j=1}^{N} \sum_{i=1}^{N} \boldsymbol{F}_{ij}^{\text{int}} \quad \left(\text{switching } i \to j \text{ and } j \to i\right) \\
&= \sum_{i=1}^{N} \boldsymbol{F}_i^{\text{ext}} - \sum_{i=1}^{N} \sum_{j=1}^{N} \boldsymbol{F}_{ij}^{\text{int}} \quad \left(\text{permuting the sums}\right) \\
&= \sum_{i=1}^{N} \boldsymbol{F}_i^{\text{ext}}.
\end{aligned}
$$

We have used the fact that

$$
\sum_{i=1}^{N} \sum_{j=1}^{N} \boldsymbol{F}_{ij}^{\text{int}} = - \sum_{i=1}^{N} \sum_{j=1}^{N} \boldsymbol{F}_{ij}^{\text{int}}
$$

i.e. a positive number is equal to its negative thus, it must be zero. □

Theorem 15.2. We can express linear momentum of a system in terms of its centre of mass:

$$
\boldsymbol{P} = M\dot{\boldsymbol{R}},
$$

where M is the total mass.

Theorem 15.1 (Linear momentum principle)

In any motion of a system, the rate of increase of its linear momentum is equal to the total external force acting upon it. By Newton's Second Law:

$$
\boldsymbol{F}^{\text{ext}} = \dot{\boldsymbol{P}} = M\ddot{\boldsymbol{R}},
$$

where M is the total mass.

Note 15.1. Thus, the centre of mass is only dependent on the external forces acting on the system.

15.2 Angular momentum and torque

Definition 15.2. For a multi-particle system consisting of particles $P_1, P_2, \ldots, P_n$, with masses $m_1, m_2, \ldots, m_n$ and velocities $\dot{\boldsymbol{r}}_1, \dot{\boldsymbol{r}}_2, \ldots, \dot{\boldsymbol{r}}_N$ respectively. Then the **angular momentum** of the system (about a point) is defined to be the vector sum of the angular

momenta of the individual particles, i.e.

$$\boldsymbol{L} = \sum_{i=1}^{N} \boldsymbol{r}_i \times (m_i \dot{\boldsymbol{r}}_i)$$

$$= \sum_{i=1}^{N} \boldsymbol{r}_i \times \boldsymbol{p}_i.$$

Definition 15.3. For a multi-particle system consisting of particles $P_1, P_2, \ldots, P_n$, with position vectors $\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N$ are acted upon by the system of forces $\boldsymbol{F}_1, \boldsymbol{F}_2, \ldots, \boldsymbol{F}_N$ respectively. Then $\boldsymbol{N}$, the **torque** of the system (about a point) is defined to be the vector sum of the torques of the individual forces, i.e.

$$\boldsymbol{N} = \sum_{i=1}^{N} \boldsymbol{r}_i \times \boldsymbol{F}_i$$

Theorem 15.3 (Angular momentum principle). Similar to the single particle case, in a multi particle system we have

$$\boldsymbol{N} = \dot{\boldsymbol{L}}.$$

15.3   Motion relative to the centre of mass

Sometimes it is easier to consider motion relative to the centre of mass, as such we consider the following formulation of the position vector, $\boldsymbol{r}_i$ of each particle:

$$\boldsymbol{r}_i = \boldsymbol{r}'_i + \boldsymbol{R},$$

where $\boldsymbol{R}$ is the position vector of the centre of mass and $\boldsymbol{r}'_i$ is the vector denoting the distance from the centre of mass to the ith particle.
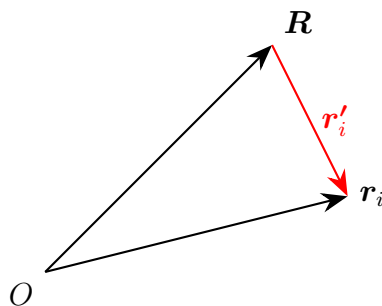
Figure 3: Illustration of relative vectors

Theorem 15.2

The centre of mass of any system moves as if it were a particle of mass the total mass, and all the **external forces** acted upon it.

15.3.1   Angular motion

Theorem 15.3

Suppose the system is influenced by central forces then the torque is given by

$$\boldsymbol{N} = \dot{\boldsymbol{L}} = \sum_{i=1}^{N} \boldsymbol{r}_i \times \boldsymbol{F}_i^{\text{ext}}.$$

Remark 15.1. By central force we mean $\boldsymbol{F}_{ij} = f_{ij}(|\boldsymbol{r}_i - \boldsymbol{r}_j|)(\boldsymbol{r}_i - \boldsymbol{r}_j)$, where f_{ij} is a scalar function of the distance between the particles.

Corollary 15.1

If the system is influenced by central force the total torque vanishes thus, angular momentum is conserved.

Note 15.2. This is similar to the one particle system case.

Theorem 15.4

The total angular momentum (relative to the centre of mass) is given by

$$\boldsymbol{L} = \sum_{i=1}^{N} m_i \boldsymbol{r}_i' \times \dot{\boldsymbol{r}}_i' + M\boldsymbol{R} \times \dot{\boldsymbol{R}}$$

$$= \boldsymbol{L}_{\text{spin}} + \boldsymbol{L}_{\text{orbital}}$$

Remark 15.2. The total angular momentum under this definition is still conserved if the system is induced by central forces.

Proof. First we note:

$$M\boldsymbol{R} = \sum_{i=1}^{N} m_i \boldsymbol{r}_i$$

$$= \sum_{i=1}^{N} m_i \left(\boldsymbol{r}_i' + \boldsymbol{R} \right)$$

$$= \sum_{i=1}^{N} m_i \boldsymbol{r}_i' + M\boldsymbol{R}.$$

Which implies that

$$\sum_{i=1}^{N} m_i \boldsymbol{r}_i' = 0 \quad \Rightarrow \quad \sum_{i=1}^{N} m_i \dot{\boldsymbol{r}}_i' = 0.$$

Recall that angular momentum is

$$\sum_{i=1}^{N} \boldsymbol{r}_i \times (m_i \dot{\boldsymbol{r}}_i).$$

Therefore, using $\boldsymbol{r}_i = \boldsymbol{r}_i' + \boldsymbol{R}$ we have that

$$\boldsymbol{L} = \sum_{i=1}^{N} (\boldsymbol{r}_i' + \boldsymbol{R}) \times m(\dot{\boldsymbol{r}}_i' + \dot{\boldsymbol{R}})$$

$$= \sum_{i=1}^{N} m_i \left(\boldsymbol{r}_i' \times \dot{\boldsymbol{r}}_i'\right) + \underbrace{\sum_{i=1}^{N} m_i \left(\boldsymbol{r}_i' \times \dot{\boldsymbol{R}}\right)}_{0} + \underbrace{\sum_{i=1}^{N} m_i \left(\boldsymbol{R} \times \dot{\boldsymbol{r}}_i'\right)}_{0} + \sum_{i=1}^{N} m_i \left(\boldsymbol{R} \times \dot{\boldsymbol{R}}\right)$$

$$= \sum_{i=1}^{N} m_i \left(\boldsymbol{r}_i' \times \dot{\boldsymbol{r}}_i'\right) + M\boldsymbol{R} \times \dot{\boldsymbol{R}}$$

$$= \boldsymbol{L}_{\text{spin}} + \boldsymbol{L}_{\text{orbital}}.$$

$\square$

Definition 15.4. The sum of the individual angular momenta of each particle about the centre of mass is called the **spin angular momentum**, denoted by $\boldsymbol{L}_{\text{spin}}$.

Definition 15.5. The angular momentum of the centre of mass is called the **orbital angular momentum**, denoted by $\boldsymbol{L}_{\text{orbital}}$.

16   Rigid body motion

Definition 16.1. A body is said to be a **rigid body** if the distance between any pair of its particles remains constant. That is, if P_i and P_j are typical particles of the body with position vectors $\boldsymbol{r}_i(t)$ and $\boldsymbol{r}_j(t)$ at time t, then

$$|\boldsymbol{r}_i(t) - \boldsymbol{r}_j(t)| = \text{constant}.$$

> **Note 16.1.** We can think of a rigid body as an idealization of a body that does not deform or change shape.

> **Remark 16.1.** In rigid body motion we can **ignore** all the internal forces.

Theorem 16.1. External forces on a rigid body act on the centre of mass, i.e.

$$\boldsymbol{F} = M\ddot{\boldsymbol{R}},$$

where M is the total mass and $\boldsymbol{R}$ is the position vector of the centre of mass.

16.1   Rotation of rigid bodies

Theorem 16.2. The total angular momentum can be decomposed into an orbital and spin part, i.e.

$$\boldsymbol{L} = \boldsymbol{L}_{\text{spin}} + \boldsymbol{L}_{\text{orbital}}.$$

Theorem 16.3. A rotating rigid body spins about the axis given by the direction of the angular velocity vector, $\boldsymbol{\omega}$. The magnitude indicates at which rate the body is spinning; it has units of $\mathrm{rad\,s}^{-1}$.

Theorem 16.1

The velocity of the ith particle relative to the centre of mass is given by

$$\dot{\boldsymbol{r}}'_i = \boldsymbol{v}'_i = \boldsymbol{\omega} \times \boldsymbol{r}'_i,$$

where $\boldsymbol{r}'_i$ is the distance from the ith particle to the centre of mass i.e. $\boldsymbol{r}_i = \boldsymbol{r}'_i + \boldsymbol{R}$.

Proof. As shown by the figure below, let the angle between $\boldsymbol{\omega}$ and $\boldsymbol{r}'_i$ be θ. Then the distance of the ith particle from the axis of rotation is $|\boldsymbol{r}'_i|\sin\theta$. If the object is rotating at $|\boldsymbol{\omega}|\,\mathrm{rad\,s}^{-1}$, then the ith particle must be travelling at a speed of $|\boldsymbol{r}'_i|\,|\boldsymbol{\omega}|\sin\theta$ in a direction orthogonal to the plane defined by ω and $\boldsymbol{r}'_i$ i.e. $\boldsymbol{v}'_i = \boldsymbol{\omega} \times \boldsymbol{r}'_i$.

$\square$
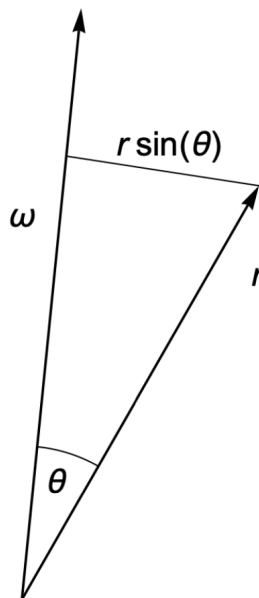
16.2   The moment of inertia tensor

Definition 16.2. The **moment of inertia** tensor (about a point) is defined by

$$\mathcal{I}_{ij} = \sum m \left(\left(\sum_{k=1}^{3} x_k^2 \right) \delta_{ij} - x_i x_j \right),$$

where δ_{ij} is the Kronecker delta function.

Remark 16.2. The moment of inertia tensor is **symmetric**, i.e. $\mathcal{I}_{ij} = \mathcal{I}_{ji}$.

Theorem 16.4. The moment of inertia tensor can be written as a matrix:

$$\mathcal{I} = \begin{pmatrix} I_{11} & I_{12} & I_{13} \\ I_{21} & I_{22} & I_{23} \\ I_{31} & I_{32} & I_{33} \end{pmatrix} = \begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix}.$$

Corollary 16.1. Since $\mathcal{I}$ is a symmetric matrix it can be diagonalised as

$$\mathcal{I} = \begin{pmatrix} I_{xx} & 0 & 0 \\ 0 & I_{yy} & 0 \\ 0 & 0 & I_{zz} \end{pmatrix},$$

where, $I_{xx}, I_{yy}, I_{zz} \in \mathbb{R}$.

Definition 16.3. The diagonal elements of $\mathcal{I}$ are called the **principal moments of inertia**.

The elements of the inertia tensor

Proposition 16.1. The **diagonal** elements of $\mathcal{I}$ are given by

$$\mathcal{I}_{11} = \sum m \left(x_2^2 + x_3^2 \right)$$
$$\mathcal{I}_{22} = \sum m \left(x_1^2 + x_3^2 \right)$$
$$\mathcal{I}_{33} = \sum m \left(x_1^2 + x_2^2 \right),$$

where the sum is taken over all the particles in the body.

Proposition 16.2. The **off-diagonal** element of $\mathcal{I}$ are given by

$$\mathcal{I}_{12} = \mathcal{I}_{21} = -\sum m \, x_1 x_2$$
$$\mathcal{I}_{23} = \mathcal{I}_{32} = -\sum m \, x_2 x_3$$
$$\mathcal{I}_{31} = \mathcal{I}_{13} = -\sum m \, x_2 x_1,$$

where the sum is taken over all the particles of the body.

In this course we will think of rigid bodies having a continuous mass distribution; in which case there exists a density function $\rho(x, y, z)$ and an infinitesimal mass element $\rho(x, y, z)dxdydz$ that replaces the point mass m_i.

Proposition 16.3. Suppose the rigid body $\mathcal{B}$ has a continuous mass distribution with density function $\rho(x, y, z)$. Then the **off-diagonal** element of $\mathcal{I}$ are given by

$$\mathcal{I}_{xx} = \int_{\mathcal{B}} \left(y^2 + z^2 \right) \rho(x, y, z) \, dx \, dy \, dz$$

$$\mathcal{I}_{yy} = \int_{\mathcal{B}} \left(x^2 + z^2 \right) \rho(x, y, z) \, dx \, dy \, dz$$

$$\mathcal{I}_{zz} = \int_{\mathcal{B}} \left(x^2 + y^2 \right) \rho(x, y, z) \, dx \, dy \, dz,$$

Proposition 16.4. Suppose the rigid body $\mathcal{B}$ has a continuous mass distribution with density function $\rho(x, y, z)$. Then the **off-diagonal** element of $\mathcal{I}$ are given by

$$\mathcal{I}_{xy} = \mathcal{I}_{yx} = -\int_{\mathcal{B}} [xy]\rho(x, y, z)\, dx\, dy\, dz,$$

$$\mathcal{I}_{yz} = \mathcal{I}_{zy} = -\int_{\mathcal{B}} [yz]\rho(x, y, z)\, dx\, dy\, dz,$$

$$\mathcal{I}_{zx} = \mathcal{I}_{xz} = -\int_{\mathcal{B}} [yx]\rho(x, y, z)\, dx\, dy\, dz,,$$

Theorem 16.2

The angular momentum about the centre of mass, $\boldsymbol{L}_{\text{spin}}$, is given by

$$\boldsymbol{L}_{\text{spin}} = \mathcal{I} \cdot \boldsymbol{\omega},$$

where $\mathcal{I}$ is the moment of inertia tensor and $\boldsymbol{\omega}$ is the angular velocity.

Note 16.2. When doing problems we should choose a coordinate basis such that $\boldsymbol{\omega}$ is aligned with one of the axes, for example $\boldsymbol{\omega} = \omega_x \boldsymbol{e}_x$.

16.3   Parallel axes theorem

Theorem 16.3 (Parallel axes theorem)

Let I_R be the principal moment of inertia of a body about some axis through its centre of mass R, and let I be the moment of inertia of the body about a **parallel** axis (in respect to the original axis). Then

$$I = I_R + Md^2,$$

where M is the mass of the body and d is the distance between the two parallel axes.

Note 16.3. The theorem says we can shift the principal moments of inertia by a factor of Md^2, if the new axis is parallel to the original one.

16.4   Kinetic energy of a rotating body

Theorem 16.4

The **rotational kinetic energy** of a body is given by

$$T_{\text{rot}} = \frac{1}{2}\boldsymbol{\omega} \cdot \boldsymbol{L}_{\text{spin}}$$

$$T_{\text{rot}} = \frac{1}{2}\left(I_{xx}\omega_1^2 + I_{yy}\omega_2^2 + I_{zz}\omega_3^2\right).$$

Remark 16.3. Notice that the rotational kinetic energy can also be written as

$$T_{\text{rot}} = \frac{L_1^2}{2I_1} + \frac{L_2^2}{2I_2} + \frac{L_3^2}{2I_3}.$$

Proof. Consider the kinetic energy of a rigid body:

$$T = \frac{1}{2} \sum_{i=1}^{N} m_i \left| \dot{\boldsymbol{r}}_i \right|^2.$$

We will rewrite the kinetic energy in terms of the contributions of the centre of mass motion and rotation about the centre of mass by introducing $\boldsymbol{r}_i = \boldsymbol{R} + \boldsymbol{r}_i'$ and its derivative $\dot{\boldsymbol{r}}_i = \dot{\boldsymbol{R}} + \dot{\boldsymbol{r}}_i'$

$$T = \frac{1}{2} \sum_{i=1}^{N} m_i \left| \dot{\boldsymbol{R}} + \dot{\boldsymbol{r}}_i' \right|^2$$

$$= \frac{1}{2} \sum_{i=1}^{N} \left(m_i \left| \dot{\boldsymbol{R}} \right|^2 + m_i \dot{\boldsymbol{r}}_i' \cdot \boldsymbol{R} + M_i \left| \dot{\boldsymbol{r}}_i' \right|^2 \right).$$

As shown previously we know that $\sum_{i=1}^{N} m_i \boldsymbol{r}_i' = 0$ hence, $\sum_{i=1}^{N} m_i \dot{\boldsymbol{r}}_i' = 0$. Thus, we can separate the kinetic energy as

$$T = \frac{1}{2} M \left| \dot{\boldsymbol{R}} \right| + \frac{1}{2} \sum_{i=1}^{N} m_i \left| \dot{\boldsymbol{r}}_i' \right|^2.$$

Focusing only on the rotational part we have

$$T_{\text{rot}} = \frac{1}{2} \sum_{i=1}^{N} m_i \left| \dot{\boldsymbol{r}}_i' \right|^2$$

$$= \frac{1}{2} \sum_{i=1}^{N} m_i \left(\boldsymbol{\omega} \times \boldsymbol{r}_i' \right) \cdot \left(\boldsymbol{\omega} \times \boldsymbol{r}_i' \right).$$

Using the vector identity

$$\left(\boldsymbol{a} \times \boldsymbol{b} \right) \cdot \boldsymbol{c} = \boldsymbol{a} \cdot \left(\boldsymbol{b} \times \boldsymbol{c} \right),$$

where $\boldsymbol{a} = \boldsymbol{\omega}, \boldsymbol{b} = \boldsymbol{r}_i'$ and $\boldsymbol{c} = \boldsymbol{\omega} \times \boldsymbol{r}_i'$. We obtain

$$T_{\text{rot}} = \frac{1}{2} \sum_{i=1}^{N} m_i \boldsymbol{\omega} \cdot \left[\boldsymbol{r}_i' \times \left(\boldsymbol{\omega} \times \boldsymbol{r}_i' \right) \right]$$

$$= \frac{1}{2} \boldsymbol{\omega} \cdot \sum_{i=1}^{N} m_i \left[\boldsymbol{r}_i' \times \left(\boldsymbol{\omega} \times \boldsymbol{r}_i' \right) \right].$$

This sum is precisely the definition of $\boldsymbol{L}_{\text{spin}}$ thus,

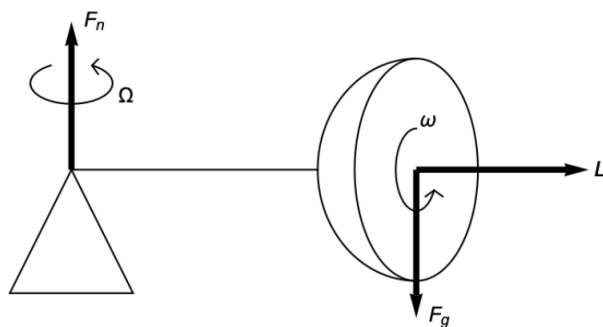$$T_{\text{rot}} = \frac{1}{2} \boldsymbol{\omega} \cdot \boldsymbol{L}_{\text{spin}}.$$

$\square$

> **Theorem 16.5**
>
> The **total kinetic energy** of a rotating body is given by
>
> $$T = \frac{1}{2}M \left| \dot{\boldsymbol{R}} \right|^2 + \frac{1}{2}\boldsymbol{\omega} \cdot \boldsymbol{L}_{\text{spin}}.$$

17   Gyroscopes

(a) Gyroscope   (b) Force diagram of a gyroscope

Suppose we have an idealised gyroscope i.e. it consists of a massless rod fixed at one end to a post. On this rod at some distance from the post, is a rapidly spinning, massive cylinder. There are two external forces acting on the rod and the cylinder: gravity pushing down on the cylinder and there is an upward normal force at the post at the end of the rod.

Imagine a situation where the downward force of gravity, $\boldsymbol{F}_g$, is exactly balanced by the upward normal force from the rod, $\boldsymbol{F}_n$, i.e. $\boldsymbol{F}_g + \boldsymbol{F}_n = 0$. The normal force does not exert any torque as it is located at fixed point, but the gravitational force will exert a torque. Suppose the centre of mass of the cylinder is located at distance $\boldsymbol{R}$ out along the rod, then the torque about the pivot (the post) will be $\boldsymbol{N} = \boldsymbol{R} \times \boldsymbol{F}_g$. The torque will point orthogonal to the plane established by $\boldsymbol{R}$ and $\boldsymbol{F}_g$ and will act to change the angular momentum.

The angular momentum, $\boldsymbol{L}$, is oriented outward (by choice) parallel to the rod. The torque, because it is orthogonal to $\boldsymbol{L}$ cannot change the magnitude of $\boldsymbol{L}$. However, the torque will change the direction of the angular momentum. The net effect is that the gyroscope will slowly spin around the post, a process called **precession**.

Suppose the gyroscope is spinning about a principal axis, then we can write $|\boldsymbol{L}| = I\omega$ where I is one of the principal components of the moment of inertia tensor. Furthermore, the magnitude of the torque is $|\boldsymbol{N}| = mgR$. Our 'ansatz' is that the gyroscope should move in a circle and that the torque always points orthogonally to the angular momentum. So, we can write

$$\boldsymbol{L} = I\omega \left[\boldsymbol{e}_x \cos(\Omega t) + \boldsymbol{e}_y \sin(\Omega t)\right],$$
$$\boldsymbol{N} = mgR \left[-\boldsymbol{e}_x \sin(\Omega t) + \boldsymbol{e}_y \cos(\Omega t)\right].$$

Therefore, we have a solution to $\boldsymbol{N} = \dfrac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{L}$, provided that

$$I\omega\Omega = mgR.$$

In other words, the **precession frequency** is

$$\Omega = \frac{mgR}{I\omega}.$$

18   The two body problem

Suppose that we have two particles with positions $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ which move subject to an external force, $\boldsymbol{F}^{\mathrm{ext}}$, as well as an internal force $\boldsymbol{F}_{12} = -\boldsymbol{F}_{21}$. We have seen that the centre of mass, $\boldsymbol{R}$, is only dependent on $\boldsymbol{F}^{\mathrm{ext}}$. We now change the variable to

$$\boldsymbol{r}_{12} = \boldsymbol{r}_1 - \boldsymbol{r}_2$$

where

$$\boldsymbol{R} = \frac{m_1\boldsymbol{r}_1 + m_2\boldsymbol{r}_2}{m_1 + m_2} \quad \text{and} \quad M = m_1 + m_2.$$

We need to invert this to find $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ as functions of $\boldsymbol{R}$ and $\boldsymbol{r}_{12}$. Furthermore, we have that

$$M\boldsymbol{R} = m_1\boldsymbol{r}_1 + m_2(\boldsymbol{r}_1 - \boldsymbol{r}_{12}).$$

Rearranging gives

$$\begin{aligned}
\boldsymbol{r}_1 &= \frac{M\boldsymbol{R} + m_2\boldsymbol{r}_{12}}{M} \\
&= \boldsymbol{R} + \frac{m_2}{M}\boldsymbol{r}_{12}
\end{aligned}$$

and

$$\begin{aligned}
\boldsymbol{r}_2 &= \boldsymbol{r}_1 - \boldsymbol{r}_{12} \\
&= \frac{M\boldsymbol{R} - m_1\boldsymbol{r}_{12}}{M} \\
&= \boldsymbol{R} - \frac{m_1}{M}\boldsymbol{r}_{12}.
\end{aligned}$$

The conserved energy is

$$\begin{aligned}
E &= \frac{1}{2}m_1\left|\dot{\boldsymbol{r}}_1\right|^2 + \frac{1}{2}m_2\left|\dot{\boldsymbol{r}}_2\right|^2 + V^{\mathrm{ext}} + V_{12} \\
&= \frac{1}{2}m_1\left|\dot{\boldsymbol{R}} + \frac{m_2}{M}\dot{\boldsymbol{r}}_{12}\right|^2 + \frac{1}{2}m_2\left|\dot{\boldsymbol{R}} - \frac{m_1}{M}\dot{\boldsymbol{r}}_{12}\right|^2 + V^{\mathrm{ext}} + V_{12} \\
&= \frac{1}{2}M\left|\dot{\boldsymbol{R}}\right|^2 + \frac{1}{2}\frac{m_1 m_2}{M}\left|\boldsymbol{r}_{12}\right|^2 + V^{\mathrm{ext}} + V_{12}
\end{aligned}$$

Definition 18.1. The quantity μ, defined by

$$\mu = \frac{m_1 m_2}{m_1 + m_2},$$

is called the **reduced mass**.

Remark 18.1. Consider when $m_1 \gg m_2$ then

$$M = m_1 + m_2 \approx m_1$$

$$\mu = \frac{m_1 m_2}{m_1 + m_2} \approx m_2.$$

Therefore, we fix a particle m_1 as the origin and treat m_2 on its own – similar to the case of planetary motion where the sun is the origin.

18.1   From two body to one body problem

Theorem 18.1

If the external potential only depends on $\boldsymbol{R}$ and the internal potential only depends on $\boldsymbol{r}_{12} = \boldsymbol{r}_1 - \boldsymbol{r}_2$ i.e.

$$V^{\text{ext}} = V^{\text{ext}}(\boldsymbol{R}) \quad \text{and} \quad V_{12}^{\text{int}} = V_{12}^{\text{int}}(\boldsymbol{r}_1 - \boldsymbol{r}_2),$$

then we can reduce the two body problem into two, one body problems: one for $\boldsymbol{R}$ and one for $\boldsymbol{r}_{12}$.

We now illustrate the process. The equation of motions for $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ are

$$m_1 \ddot{\boldsymbol{r}}_1 = -\nabla_1 V^{\text{ext}} - \nabla_1 V_{12}$$
$$m_2 \ddot{\boldsymbol{r}}_2 = -\nabla_2 V^{\text{ext}} - \nabla_2 V_{12}.$$

The equation of motion for $\boldsymbol{R}$ is obtained by summing the two equations above.

$$
\begin{aligned}
M\ddot{\boldsymbol{R}} &= \left(-\nabla_1 V^{\text{ext}} - \nabla_1 V_{12}\right) + \left(-\nabla_2 V^{\text{ext}} - \nabla_2 V_{12}\right) \\
&= \left(-\nabla_1 V^{\text{ext}} - \nabla_1 V_{12}\right) + \left(-\nabla_2 V^{\text{ext}} - [-\nabla_1 V_{12}]\right) \\
&= -\nabla_1 V^{\text{ext}} - \nabla_2 V^{\text{ext}} \\
&= -\frac{m_1}{M}\nabla_R V^{\text{ext}} - \frac{m_2}{M}\nabla_R V^{\text{ext}} \\
&= -\nabla_R V^{\text{ext}} \\
&= \sum_{i=1}^{N} \boldsymbol{F}_i^{\text{ext}}.
\end{aligned}
$$

Here we have used [NIII] to deduce

$$\nabla_1 V_{12} = -\nabla_2 V_{12},$$

as well as the chain rule to write

$$\nabla_1 V^{\text{ext}} = \frac{m_1}{M}\nabla_R V^{\text{ext}} \quad \text{and} \quad \nabla_2 V^{\text{ext}} = \frac{m_2}{M}\nabla_R V^{\text{ext}}$$

(assuming that V^{ext} only depends on $\boldsymbol{R}$). Therefore, we have

$$M\ddot{\boldsymbol{R}} = -\nabla_R V^{\text{ext}}(\boldsymbol{R}).$$

Easier way

Suppose we have

$$m_1 \ddot{\boldsymbol{r}}_1 = \boldsymbol{F}_1^{\text{ext}} + \boldsymbol{F}_{12}^{\text{int}}$$
$$m_2 \ddot{\boldsymbol{r}}_2 = \boldsymbol{F}_2^{\text{ext}} + \boldsymbol{F}_{21}^{\text{int}}.$$

We now proceed as follows, consider

$$m_1 m_2 \ddot{\boldsymbol{r}}_1 = m_2 \boldsymbol{F}_1^{\text{ext}} + m_2 \boldsymbol{F}_{12}^{\text{int}}$$
$$m_1 m_2 \ddot{\boldsymbol{r}}_2 = m_1 \boldsymbol{F}_2^{\text{ext}} + m_1 \boldsymbol{F}_{21}^{\text{int}},$$

then we have,

$$m_1 m_2 (\ddot{\boldsymbol{r}}_1 - \ddot{\boldsymbol{r}}_2) = m_2 \boldsymbol{F}_1^{\text{ext}} - m_1 \boldsymbol{F}_2^{\text{ext}} + m_2 \boldsymbol{F}_{12}^{\text{int}} + m_1 \boldsymbol{F}_{12}^{\text{int}}$$
$$= \boldsymbol{F}^{\text{ext}} + (m_1 + m_2) \boldsymbol{F}_{12}^{\text{int}}$$

where $\boldsymbol{F}^{\text{ext}} = m_2 \boldsymbol{F}_1^{\text{ext}} - m_1 \boldsymbol{F}_2^{\text{ext}}$. Therefore, we have

$$\mu \ddot{\boldsymbol{r}}_{12} = \widetilde{\boldsymbol{F}}^{\text{ext}} + \boldsymbol{F}_{12}^{\text{int}}$$

where, μ is the reduced mass and $\widetilde{\boldsymbol{F}}^{\text{ext}} = \frac{\boldsymbol{F}^{\text{ext}}}{m_1 + m_2}$.

Now that we know $\ddot{\boldsymbol{r}}_{12}$ we can solve for $\boldsymbol{r}_{12}$; from above we know that

$$\boldsymbol{r}_1 = \boldsymbol{R} + \frac{m_2}{M} \boldsymbol{r}_{12}$$

and

$$\boldsymbol{r}_{12} = \boldsymbol{r}_1 - \boldsymbol{r}_2$$

thus, we can solve for both by just knowing $\boldsymbol{r}_{12}$.

Theorem 18.2

Suppose $M \ddot{\boldsymbol{R}} = -\nabla_R V^{\text{ext}}(\boldsymbol{R})$, then the total energy of the centre of mass is given by

$$E_{\text{cm}} = \frac{1}{2} M \left| \dot{\boldsymbol{R}} \right|^2 + V^{\text{ext}}(\boldsymbol{R}).$$

The energy, E_{cm} is a conserved quantity.

Proof.

$$\dot{E}_{\text{cm}} = M \dot{\boldsymbol{R}} \cdot \ddot{\boldsymbol{R}} + \nabla_R V^{\text{ext}}(\boldsymbol{R}) \cdot \dot{\boldsymbol{R}}$$
$$= \left(M \ddot{\boldsymbol{R}} + V^{\text{ext}}(\boldsymbol{R}) \right) \cdot \dot{\boldsymbol{R}}$$
$$= 0.$$

Since $M \ddot{\boldsymbol{R}} = -\nabla_R V^{\text{ext}}(\boldsymbol{R})$ $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

To obtain an equation for the relative position we consider the equations of motions for $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$, as before but written as

$$m_1 m_2 \ddot{\boldsymbol{r}}_1 = m_2 \left(-\nabla_1 V^{\text{ext}} - \nabla_1 V_{12} \right)$$
$$m_1 m_2 \ddot{\boldsymbol{r}}_2 = m_1 \left(-\nabla_2 V^{\text{ext}} - \nabla_2 V_{12} \right).$$

Note that $M\mu = m_1 m_2$; subtracting the two equations we have

$$M\mu\ddot{\boldsymbol{r}}_{12} = -m_2\nabla_1 V^{\text{ext}} + m_1\nabla_2 V^{\text{ext}} - m_2\nabla_1 V_{12} + m_1\nabla_2 V_{12}.$$

Supposing that $V^{\text{ext}} = V^{\text{ext}}(\boldsymbol{R})$ then we can use the chain rule to express

$$-m_2\nabla_1 V^{\text{ext}} + m_1\nabla_2 V^{\text{ext}} = -\frac{m_1 m_2}{M}\nabla_R V^{\text{ext}} + \frac{m_1 m_2}{M}\nabla_R V^{\text{ext}}$$
$$= 0.$$

Furthermore, if $V_{12} = V_{12}(\boldsymbol{r}_{12})$ then we can write

$$\nabla_2 V_{12} = -\nabla_1 V_{12}$$
$$= -\nabla_{12} V_{12}.$$

Therefore, the equation of motion for $\boldsymbol{r}_{12}$ becomes,

$$\mu\ddot{\boldsymbol{r}}_{12} = -\nabla_{12} V_{12}(\boldsymbol{r}_{12}).$$

Theorem 18.3

The total energy for relative motion is given by

$$E_{12} = \frac{1}{2}\mu\left|\dot{\boldsymbol{r}}_{12}\right|^2 + V_{12}(\boldsymbol{r}_{12}).$$

The energy, E_{12}, is a conserved quantity.

Corollary 18.1

The total energy of the system, $E = E_{\text{cm}} + E_{12}$, is a conserved quantity.

Proof. Both E_{cm} and E_{12} are conserved thus, the sum of two conserved quantities is also conserved i.e. $\dot{E} = \dot{E}_{\text{cm}} + \dot{E}_{12} = 0$. $\qquad\square$

Theorem 18.4

If $V_{12} = V_{12}(|\boldsymbol{r}_{12}|)$ then the energy and angular momentum of the relative system will be conserved leading to a single one-dimensional problem:

$$E_{12} = \frac{1}{2}\mu\left|\dot{\boldsymbol{r}}_{12}\right|^2 + V_{\text{eff}}^{\text{int}} \quad\text{and}\quad V_{\text{eff}}^{\text{int}} = V_{12}(|\boldsymbol{r}_{12}|) + \frac{\ell_{12}^2}{2\mu\left|\boldsymbol{r}_{12}\right|^2}.$$

Theorem 18.5

If $V^{\text{ext}} = V^{\text{ext}}(|\boldsymbol{R}|)$ then

$$E_{\text{cm}} = \frac{1}{2}M\left|\dot{\boldsymbol{R}}\right|^2 + V_{\text{eff}}^{\text{ext}} \quad\text{and}\quad V_{\text{eff}}^{\text{ext}} = V^{\text{ext}}(|\boldsymbol{R}|) + \frac{\ell_{\text{cm}}^2}{2M\left|\boldsymbol{R}\right|^2}.$$

Analytical mechanics

19   Lagrangian mechanics

> **Theorem 19.1**
>
> REMARK!!!!
> (For our course) The Lagrangian formulation is only true when the system is influenced by **conservative forces**!!

19.1   Generalised coordinates

Consider a rigid body composed of many particles, N. The positions of all the particles may be specified by $3N$ coordinates. However, these $3N$ coordinates cannot all vary independently but are subject to constraints.

> **Remark 19.1.** Further explanation:
>
> - By saying the generalised coordinates must be **independent variables**, we mean that there must be *no functional relation connecting them*. Therefore, if one of the coordinates were to be removed, the remaining $n - 1$ coordinates would still determine the configuration of the system.
>
> - When we say the generalised coordinates $q_1, q_2, \ldots q_n$ **determine the configuration** of the system $\mathcal{S}$, we mean that, when the values of the coordinates $q_1, q_2, \ldots q_n$ are given, the position of every particle of $\mathcal{S}$ is determined. That is, the position vectors $\{\boldsymbol{r}_i\}$ of the particles must be known functions of the independent variables $q_1, q_2, \ldots q_n$ such that
> $$\boldsymbol{r}_i = \boldsymbol{r}_i(q_1, q_2, \ldots q_n) \quad \text{for } i = 1, 2, \ldots n.$$

> **Remark 19.2.** Generalised coordinates are scalars and not vectors.

Definition 19.1. Let $\mathcal{S}$ be a dynamical system defined by a set of generalised coordinates. Then the **number** of generalised coordinates needed to specify the configuration of $\mathcal{S}$ is called the number of **degree of freedom** of $\mathcal{S}$.

19.2   Constraints

A general mechanical system $\mathcal{S}$ consists of any number of particles $P_1, P_2, \ldots, P_n$. The particles of $\mathcal{S}$ may have interconnections of various kinds (light strings, springs and so on) and also be subject to external connections and constraints. These could include features such as a particle being forced to remain on a fixed surface or suspended from a fixed point by a light inextensible string.

Definition 19.2. A 'position' in which a particle in a dynamical system are at time t is called a **configuration**.

Definition 19.3. If the configuration of a system $\mathcal{S}$ is determined by the values of a set of independent variables $q_1, q_2, \ldots, q_n$ is said to be a set of **generalised coordinates** for $\mathcal{S}$.

19.2.1   Unconstrained systems

Definition 19.4. Particles in a dynamical system which are free to move anywhere in space *independently of each other* are said to be in an **unconstrained system**.

Theorem 19.1. In an unconstrained system the equations of motion for the system are given by Newton's approach i.e.
$$m_i \ddot{\boldsymbol{r}}_i = \boldsymbol{F}_i$$
(for $i = 1, \ldots, N$), where $\boldsymbol{F}_i$ is the force acting on the particle P_i.

19.2.2   Constrained systems

Definition 19.5. Systems in which the particles are which are subject to geometrical or kinematical constraints are called **constrained systems**.

Definition 19.6. Kinematical constraints are those that involve the position and velocity vectors $\{\boldsymbol{r}_i, \dot{\boldsymbol{r}}_i\}$.

Definition 19.7. Geometrical constraints are those that involve only the position vectors $\{\boldsymbol{r}_i\}$. We will call these constraints **holonomic constraints**.

> **Remark 19.3.** Constraints which are not holonomic are called **non-holonomic**.

> **Theorem 19.2**
> A holonomic constraint takes the form
> $$C(q_i, t) = 0.$$

> **Remark 19.4.** In general each holonomic constraint reduces the number of degrees of freedom of the system by one.

> **Proposition 19.1**
> TYPICAL QUESTION: when asked why a constraint is holonomic it is enough to say: "the constraints allows for algebraic elimination of one coordinate".

19.3   The Lagrangian

Definition 19.8. We define the **Lagrangian** as a (linearly) independent quantity defined by
$$\mathcal{L} = T - V$$
$$= \frac{1}{2} m |\dot{\boldsymbol{r}}|^2 - V(\boldsymbol{r}).$$

Remark 19.5. The Lagrangian is not a functional because it depends on $\boldsymbol{r}$ and $\dot{\boldsymbol{r}}$ at a single time.

19.4 The principle of least action

Definition 19.9. The quantity

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

is called the **action (functional)** corresponding to the Lagrangian $\mathcal{L}(q_i, \dot{q}_i, t)$ (for the time interval $[t_1, t_2]$).

Note 19.1. Functionals are functions with functions as their inputs; in this case the functional S depends on the whole $q_i(t)$.

Theorem 19.3 (Principle of least action)

Particles move to extremise the action S as a functional of all possible paths between $\boldsymbol{r}_1(t)$ and $\boldsymbol{r}_2(t)$.

Note 19.2. The principle of least action is in other words saying: a particle with a fixed energy moving from point A to B (these endpoints are fixed) during a time interval $[t_1, t_2]$ will choose the path that uses the **least** amount of energy. That is, the functional S is always a minimum; mathematically we can determine when S is a minimum with derivatives. Therefore, mathematically the principle of least action says

$$\frac{\partial S}{\partial q(t)} = 0.$$

Theorem 19.4 (Euler-Lagrange equations)

The vector function q_i makes the action functional minimised if and only if q_i satisfies the **Euler-Lagrange** differential equations

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

Note 19.3. The E-L equations are equivalent to $F = ma$.

Proof. This proof will be for a single particle, but it can be simply extended to multiple particles. The principle of least action requires the action to be minimised thus, we need $\delta S = 0$. Recall the total differential

$$dF = \frac{\partial F}{\partial x}dx + \frac{\partial F}{\partial y}dy + \cdots$$

We want to set $\delta S = 0$ so, we need to evaluate δS first:

$$\delta S = \int_{t_1}^{t_2} \delta\mathcal{L}(q, \dot{q}) \, dt.$$

We have

$$\delta\mathcal{L}(q, \dot{q}) = \frac{\partial\mathcal{L}}{\partial q}\delta q + \frac{\partial\mathcal{L}}{\partial \dot{q}}\delta\dot{q}$$

therefore,

$$
\begin{aligned}
\delta S &= \int_{t_1}^{t_2} \frac{\partial\mathcal{L}}{\partial q}\delta q + \frac{\partial\mathcal{L}}{\partial \dot{q}}\delta\dot{q} \\
&= \int_{t_1}^{t_2} \left(\frac{\partial\mathcal{L}}{\partial q}\delta q\right) dt + \int_{t_1}^{t_2} \left(\frac{\partial\mathcal{L}}{\partial \dot{q}}\delta\dot{q}\right) dt \\
&= \underbrace{\int_{t_1}^{t_2} \left(\frac{\partial\mathcal{L}}{\partial q}\delta q\right) dt}_{I_1} + \underbrace{\int_{t_1}^{t_2} \left(\frac{\partial\mathcal{L}}{\partial \dot{q}}\frac{\mathrm{d}(\delta q)}{\mathrm{d}t}\right) dt}_{I_2}.
\end{aligned}
$$

Recall 'integration by parts':

$$\int_a^b u\frac{\mathrm{d}v}{\mathrm{d}x}\, dx = [uv]_a^b - \int_a^b v\frac{\mathrm{d}u}{\mathrm{d}x}\, dx,$$

evaluating I_2 we have

$$= -\int_{t_1}^{t_2} \left(\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial \dot{q}}\right) \delta q \, dt.$$

The first term is 0 since we take the endpoints to be fixed we have that $\delta q(t_1) = \delta q(t_2) = 0$. Finally, we have

$$
\begin{aligned}
\delta S &= I_1 + I_2 \\
&= \int_{t_1}^{t_2} \left(\frac{\partial\mathcal{L}}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial \dot{q}}\right) \delta q \, dt
\end{aligned}
$$

which implies $\delta S = 0$ if and only if

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\mathcal{L}}{\partial \dot{q}}\right) - \frac{\partial\mathcal{L}}{\partial q} = 0.$$

$\qquad\square$

> **Note 19.4.** In this course to denote small changed in variables we use δ and d interchangeably.

19.5 Conjugate momenta

Definition 19.10. Consider a holonomic mechanical system with Lagrangian $\mathcal{L} = \mathcal{L}(q_i, \dot{q}_i, t)$. Then the scalar quantity p_j, defined by

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

is called the **conjugate momentum** corresponding to the coordinate q_j.

Remark 19.6. We can think of

$$\boldsymbol{F}_i = \frac{\partial \mathcal{L}}{\partial q_i}$$

then the Euler-Lagrange equation reads

$$\frac{\mathrm{d}}{\mathrm{d}t} p_i = \boldsymbol{F}_i.$$

Example 19.1

Consider a system with Lagrangian

$$\mathcal{L} = \frac{1}{2} M \dot{x}^2 + \frac{1}{2} m \left(\dot{x}^2 + \dot{y}^2 + 2\dot{x}\dot{y}\cos(\alpha) \right) + mgy\sin(\alpha).$$

Find the conjugate momenta.

Solution. With this Lagrangian, the momenta p_x and p_y are given by

$$p_x = \frac{\partial \mathcal{L}}{\partial \dot{x}} = M\dot{x} + m\left(\dot{x} + \dot{y}\cos\alpha\right),$$

$$p_y = \frac{\partial \mathcal{L}}{\partial \dot{y}} = m\left(\dot{y} + \dot{x}\cos\alpha\right).$$

19.5.1 Conservation of conjugate momenta

Definition 19.11. If $\dfrac{\partial \mathcal{L}}{\partial q_j} = 0$ (i.e. if the coordinate q_j is absent from the Lagrangian) then q_j is said to be an **ignorable** (or **cyclic**) coordinate.

Remark 19.7. Definition from the lecture notes:
If

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

then, q_i is said to be ignorable.

Theorem 19.2. If q_i is an ignorable coordinate then,

$$\frac{\mathrm{d}}{\mathrm{d}t} p_i = 0.$$

Corollary 19.1 (Conservation of momentum)

If q_i is an *ignorable* coordinate (in the sense that it does not appear in the Lagrangian) then p_j, the conjugate momentum to q_j, is constant in any motion.

Remark 19.8. That is, to say the conjugate momentum is conserved.

Example 19.2

Consider a Lagrangian given by

$$\mathcal{L} = \frac{1}{2}ma^2\left[\dot{\theta}^2 + \left(\dot{\phi}\sin\theta\right)^2\right] + mga\cos\theta.$$

Verify that ϕ is an ignorable coordinate and find the corresponding conjugate momentum.

Solution. Since $\dfrac{\partial\mathcal{L}}{\partial\phi} = 0$, the coordinate ϕ is ignorable. It follows that the conjugate momentum, p_ϕ, is conserved, where

$$p_\phi = \frac{\partial\mathcal{L}}{\partial\dot{\phi}} = ma^2\sin^2(\theta\dot{\phi}).$$

19.6   Symmetries

Definition 19.12. A **symmetry** is a transformation

$$q_i \to q_i'$$
$$\dot{q}_i \to \dot{q}_i'$$
$$t \to t'$$

such that

1. the Lagrangian is invariant i.e.

$$\mathcal{L}(q_i', \dot{q}_i', t') = \mathcal{L}(q_i, \dot{q}_i, t);$$

2. it is a symmetry of the action

$$\mathcal{L}(q_i', \dot{q}_i', t') = \mathcal{L}(q_i, \dot{q}_i, t) + \frac{\mathrm{d}J}{\mathrm{d}t}.$$

Remark 19.9. We allow for a shift by a total derivative of J as after taking the integral for the action, the term will disappear, i.e. it leaves the action invariant.

Proposition 19.1. For a continuous symmetry it is enough (and simpler) to study them for infinitesimal transformations:

$$q_i \to q_i + \varepsilon T_i(q, t) = q_i'$$
$$\dot{q}_i \to \dot{q}_i + \varepsilon \dot{T}(q, t) = \dot{q}_i'$$

for $\varepsilon \ll 1$.

Proposition 19.2

If q_* is an ignorable coordinate of the Lagrangian then, there exists a symmetry of the Lagrangian

$$\mathcal{L}(q_i', \dot{q}_i', t) = \mathcal{L}(q_i, \dot{q}_i, t)$$

where

$$q_i \to q_i' = q_i \quad \text{for } i \neq *$$
$$\dot{q}_i \to \dot{q}_i' = \dot{q}_i \quad \text{for } i \neq *$$
$$q_* \to q_*' = q_* + \varepsilon$$
$$\dot{q}_* \to \dot{q}_*' = \dot{q}_*.$$

19.7   Noether's theorem

Theorem 19.5 (Noether's theorem – Lagrangian)

For each symmetry of the Lagrangian,

$$q_i \to q_i' = q_i + \varepsilon T_i(q, t)$$
$$\dot{q}_i \to \dot{q}_i' = \dot{q}_i + \varepsilon \dot{T}_i(q, t)$$

there is a conserved quantity:

$$Q = \sum_{i=1}^{N} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} T_i(q, t).$$

The quantity Q, we call **Noether's charge**.

Proof. We want to show that Q is a conserved quantity i.e. $\dfrac{\mathrm{d}Q}{\mathrm{d}t} = 0$, we now compute

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = \sum_{n=1}^{N} \left[\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) T_i + \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \frac{\mathrm{d}T_i}{\mathrm{d}t} \right]$$

$$= \sum_{n=1}^{N} \left[\frac{\partial \mathcal{L}}{\partial q_i} T_i + \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \dot{T}_i \right]$$

where in the second equality we have used to Euler-Lagrange equations. For each symmetry

$$q_i \to q_i + \varepsilon T_i$$
$$\dot{q}_i \to \dot{q}_i + \varepsilon \dot{T}_i$$

we have, $\mathcal{L}(q_i + \varepsilon T_i, \dot{q}_i + \varepsilon \dot{T}_i, t) = \mathcal{L}(q_i, \dot{q}_i, t)$. Expanding with the Taylor series we must have

$$\mathcal{L}(q_i, \dot{q}_i, t) = \mathcal{L}(q_i + \varepsilon T_i, \dot{q}_i + \varepsilon \dot{T}_i, t)$$

$$= \mathcal{L}(q_i, \dot{q}_i, t) + \varepsilon \sum_{n=1}^{N} \left[\frac{\partial \mathcal{L}}{\partial q_i} T_i + \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \dot{T}_i \right] + O(\varepsilon^2)$$

which implies

$$\sum_{n=1}^{N} \left[\frac{\partial \mathcal{L}}{\partial q_i} T_i + \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \dot{T}_i \right] = 0$$

hence, $\dfrac{\mathrm{d}Q}{\mathrm{d}t} = 0$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

> **Theorem 19.6** (Noether's theorem – Action)
> For each symmetry of the action
>
> $$q_i \to q_i' = q_i + \varepsilon T_i(q, t)$$
> $$\dot{q}_i \to \dot{q}_i' = \dot{q}_i + \varepsilon \dot{T}_i(q, t)$$
>
> such that
>
> $$\mathcal{L}(q_i', \dot{q}_i', t) = \mathcal{L}(q_i, \dot{q}_i, t) + \varepsilon \frac{\mathrm{d}J}{\mathrm{d}t},$$
>
> there is a conserved quantity:
>
> $$Q = \sum_{i=1}^{N} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} T_i(q, t) - J.$$

> **Remark 19.10.** This keeps the action invariant i.e. $S[q_i'] = S[q_i]$ but not the Lagrangian.

Proof. We compute

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t} \left[\sum_{i=1}^{N} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} T_i - J \right]$$

$$= \sum_{i=1}^{N} \left[\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) T_i + \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \dot{T}_i \right] - \frac{\mathrm{d}J}{\mathrm{d}t}$$

$$= \sum_{n=1}^{N} \left[\frac{\partial \mathcal{L}}{\partial q_i} T_i + \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \dot{T}_i \right] - \frac{\mathrm{d}J}{\mathrm{d}t}.$$

For each symmetry we have, $\mathcal{L}(q_i, \dot{q}_i, t) + \varepsilon \dfrac{\mathrm{d}J}{\mathrm{d}t} = \mathcal{L}(q_i + \varepsilon T_i, \dot{q}_i + \varepsilon \dot{T}_i, t)$ therefore, using a Taylor expansion we have

$$\mathcal{L}(q_i, \dot{q}_i, t) + \varepsilon \frac{\mathrm{d}J}{\mathrm{d}t} = \mathcal{L}(q_i + \varepsilon T_i, \dot{q}_i + \varepsilon \dot{T}_i, t)$$

$$= \mathcal{L}(q_i, \dot{q}_i, t) + \varepsilon \sum_{n=1}^{N} \left[\frac{\partial \mathcal{L}}{\partial q_i} T_i + \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \dot{T}_i \right] + O(\varepsilon^2),$$

from which we imply

$$\sum_{n=1}^{N} \left[\frac{\partial \mathcal{L}}{\partial q_i} T_i + \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \dot{T}_i \right] = \frac{\mathrm{d}J}{\mathrm{d}t}.$$

In conclusion, $\dfrac{\mathrm{d}Q}{\mathrm{d}t} = 0$. $\qquad\qquad\square$

19.7.1   Symmetry and conservation principles

> **Theorem 19.7** (Invariance under spatial translation)
>
> Suppose that the kinetic and potential energy only depends on the seperation between any two pairs of particle (e.g. $T = T(\boldsymbol{r}_i - \boldsymbol{r}_j)$ and $V = V(\boldsymbol{r}_i - \boldsymbol{r}_j)$). Then, there is an overall **translational symmetry**:
>
> $$\boldsymbol{r}_i \to \boldsymbol{r}_i + \varepsilon\boldsymbol{a} \quad \text{and} \quad \dot{\boldsymbol{r}}_i \to \dot{\boldsymbol{r}}_i,$$
>
> where $\boldsymbol{a}$ is a fixed vector. Translational symmetry implies there is conservation of the **total momentum** of the system.

> **Remark 19.11.** This symmetry reflects the **homogeneity** of space, namelly that there is no preferred location.

Proof. Suppose
$$\boldsymbol{r}_i \to \boldsymbol{r}_i + \varepsilon\boldsymbol{a} \quad \text{and} \quad \dot{\boldsymbol{r}}_i \to \dot{\boldsymbol{r}}_i,$$
where $\boldsymbol{a}$ is a constant. Then, $\boldsymbol{r}_i - \boldsymbol{r}_j$ remains invariant and $\dot{\boldsymbol{a}} = \boldsymbol{0}$. Therefore, the Lagrangian will be invaraint under the transformation and by Noether's theorem the quantity

$$
\begin{aligned}
Q &= \sum_{i=1}^{N} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} T_i \\
&= \sum_{i=1}^{N} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \boldsymbol{a} \\
&= \sum_{i=1}^{N} \boldsymbol{p}_i \cdot \boldsymbol{a} \\
&= \boldsymbol{P} \cdot \boldsymbol{a},
\end{aligned}
$$

is conserved. (Note that we take $T_i = \boldsymbol{a}$). $\qquad\square$

> **Theorem 19.8** (Invariance under rotation)
>
> Suppose that the kinetic and potential energy only depends on the distance $|\boldsymbol{r}_i - \boldsymbol{r}_j|$ between any two pairs of particles (and not the direction). Then we can cosider a rotation of all the particles:
>
> $$\boldsymbol{r}_i \to \boldsymbol{r}_i + \varepsilon T\boldsymbol{r}_i \quad \text{and} \quad \dot{\boldsymbol{r}}_i \to \dot{\boldsymbol{r}}_i + \varepsilon\dot{T}\dot{\boldsymbol{r}}_i,$$
>
> where T is a constant anty-symmetric matrix (i.e. $T^\top = -T$). Rotational symmetry implies there is conservation of the **total angular momentum** of the system.

> **Remark 19.12.** This symmetry reflects the **isotropy** of space, namely that it looks the same in all directions.

Proof. We must show that under the symmetry transformation the length is preserved i.e. $\delta(|\boldsymbol{r}_i - \boldsymbol{r}_j|^2)$. We now compute this:

$$
\begin{aligned}
\delta(|\boldsymbol{r}_i - \boldsymbol{r}_j|^2) &= \delta((\boldsymbol{r}_i - \boldsymbol{r}_j) \cdot (\boldsymbol{r}_i - \boldsymbol{r}_j)) \\
&= 2(\boldsymbol{r}_i - \boldsymbol{r}_j) \cdot \delta(\boldsymbol{r}_i - \boldsymbol{r}_j) \\
&= 2\varepsilon(\boldsymbol{r}_i - \boldsymbol{r}_j) \cdot \mathbb{T} \cdot (\boldsymbol{r}_i - \boldsymbol{r}_j) \\
&= 2\varepsilon(\boldsymbol{r}_i - \boldsymbol{r}_j)^\top \cdot \mathbb{T} \cdot (\boldsymbol{r}_i - \boldsymbol{r}_j),
\end{aligned}
$$

where $\mathbb{T}$ is an anti-symmetric matric. Denote $\Delta\boldsymbol{r} = \boldsymbol{r}_i - \boldsymbol{r}_j$ then,

$$
\begin{aligned}
(\Delta\boldsymbol{r})^\top \mathbb{T} \Delta\boldsymbol{r} &= \left((\Delta\boldsymbol{r})^\top \mathbb{T} \Delta\boldsymbol{r}\right) \\
&= (\Delta\boldsymbol{r})^\top \mathbb{T}^\top \Delta\boldsymbol{r} \\
&= -(\Delta\boldsymbol{r})^\top \mathbb{T} \Delta\boldsymbol{r}.
\end{aligned}
$$

Therefore, we must have that $(\Delta\boldsymbol{r})^\top \mathbb{T} \Delta\boldsymbol{r} = 0$ which implies $\delta(|\boldsymbol{r}_i - \boldsymbol{r}_j|^2) = 0$; by similar reasoning we can show $\delta(|\dot{\boldsymbol{r}}_i - \dot{\boldsymbol{r}}_j|) = 0$. We now show the conserved quantity is indeed the total angular momentum. Parametrise the matrix such that

$$
\mathbb{T} = \begin{pmatrix} 0 & -T_3 & T_2 \\ T_3 & 0 & -T_1 \\ -T_2 & T_1 & 0 \end{pmatrix}
$$

so, that we can write

$$
\mathbb{T}_{ab} = -\sum_{c=1}^{3} \varepsilon_{abc} T_c.
$$

Applying Noether's theorem we know the conserved quantity is

$$
\begin{aligned}
Q &= \sum_{i=1}^{N} \frac{\partial \mathcal{L}}{\partial \dot{\boldsymbol{r}}_i} \cdot \mathbb{T} \boldsymbol{r}_i \\
&= \sum_{i=1}^{N} \boldsymbol{p}_i \mathbb{T} \boldsymbol{r}_i \\
&= \sum_{i=1}^{N} \sum_{ab} \boldsymbol{p}_i^a \cdot \mathbb{T}_{ab} \boldsymbol{r}_i^b \\
&= -\sum_{i=1}^{N} \sum_{abc} p_i^a \varepsilon_{abc} r_i^b T_c \\
&= \mathbb{T} \cdot \sum_{i=1}^{N} \boldsymbol{r}_i \times \boldsymbol{p}_i \\
&= \mathbb{T} \cdot \boldsymbol{L},
\end{aligned}
$$

which is the total angular momentum along the direction of $\mathbb{T}$. $\qquad\square$

Example 19.3

If there is a rotational symmetry then, the continuous infinitesimal symmetry is given by

$$x \to x - \varepsilon y$$
$$y \to y + \varepsilon x.$$

That is the infinitesimal version of

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

where $\theta = \varepsilon$.

Theorem 19.9 (Invariance under time translations)

Suppose the Lagrangian does not have any explicit time dependence:

$$\frac{\partial \mathcal{L}}{\partial t} = 0.$$

Then, the **total energy** of the system is conserved.

Proof. First we notice that this is a symmetry of the action. Since the Lagrangian has no explicit time dependence we have that $\mathcal{L} = \mathcal{L}(q_i, \dot{q}_i)$. A translation in time means a shift

$$t_1 \to t_1 + \varepsilon$$
$$t_2 \to t_2 + \varepsilon.$$

Therefore, we have that

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

$$= \int_{t_1}^{t_2} \mathcal{L}(q_i(t+\varepsilon), \dot{q}_i(t+\varepsilon)) - \mathcal{L}(q_i(t), \dot{q}_i(t)) \, dt,$$

where we have used a change of variable for the first term $t' = t + \varepsilon$. As such the symmetry is given by

$$q_i(t) \to q_i(t+\varepsilon) = q_i(t) + \varepsilon \dot{q}_i + \cdots \qquad \Rightarrow \delta q_i = \varepsilon \dot{q}_i$$
$$\dot{q}_i(t) \to \dot{q}_i(t+\varepsilon) + \dot{q}_i(t) + \varepsilon \ddot{q}_i + \cdots \qquad \Rightarrow \delta \dot{q}_i = \varepsilon \ddot{q}_i$$

so, $T_i = \dot{q}_i$ (since $\delta q_i = \varepsilon T_i$). Next we have that (by the Taylor expansion):

$$\mathcal{L}(q_i(t+\varepsilon), \dot{q}_i(t+\varepsilon)) = \mathcal{L}(q_i(t), \dot{q}_i(t)) + \varepsilon \frac{d\mathcal{L}}{dt}$$

where

$$\frac{d\mathcal{L}}{dt} = \sum_{i=1}^{N} \frac{\partial \mathcal{L}}{\partial q_i} \dot{q}_i + \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \ddot{q}_i.$$

This implies that

$$\delta S = \varepsilon \int_{t_1}^{t_2} \frac{d\mathcal{L}}{dt} \, dt.$$

Applying Noether's theorem we know the conserved quantity, i.e. the charge Q, is given by

$$Q = \sum_{i=1}^{N} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} T_i - J$$

$$= \sum_{i=1}^{N} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \dot{q}_i - \mathcal{L}$$

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

$$= E.$$

$\square$

19.8   Symmetry in the potential

19.8.1   2-dimensional

In this section we explore in which type of potential we have symmetries for.

> **Theorem 19.10**
>
> If the potential takes the form of
>
> - $V(x,y) = f(x)$ then, y is ignorable and all translations in y are preserved;
>
> - $V(x,y) = f(y)$ then, x is ignorable and all translation in (x) are preserved;
>
> - $V(x,y) = f(\sqrt{x^2 + y^2}) = f(r)$ then, θ is ignorable, and all rotations are preuserved;
>
> - V without any explicit time dependence implies conservation of energy.

19.8.2   3-dimensional

> **Theorem 19.11**
>
> If the potential takes the form of
>
> - $V = V(\boldsymbol{r}_i - \boldsymbol{r}_j)$ the symmetry is $\boldsymbol{r}_i \to \boldsymbol{r}_i + \varepsilon\boldsymbol{a}$, which implies conservation of **total momentum**;
>
> - $V = (|\boldsymbol{r}_i - \boldsymbol{r}_j|)$ the symmetry is $\boldsymbol{r}_i \to \boldsymbol{r}_i + \varepsilon T\boldsymbol{r}_i$ (where T is an anti-symmetric matrix), which implies conservation of **total angular momentum**;
>
> - V without any explicit time dependence the symmetry is $t \to t + \varepsilon$, which implies conservation of energy.

20   Hamiltonian Mechanics

In Hamiltonian mechanics, the conjugate momenta $p_i \equiv \dfrac{\partial \mathcal{L}}{\partial \dot{q}_i}$ are 'promoted' to coordinates on equal footing with the generalised coordinates q_i.

Definition 20.1. The coordinates (q, p) are **canonical variables.**

20.1   Legendre transforms

Definition 20.2. The process to convert Lagrangian formulation to a Hamiltonian formulation is called **Legendre transforms**.

Consider a function $F(x, y)$ which we want to swap with a new function $\widetilde{F}(x, u)$ without losing any information. To do this we note that the total differential of F is given by

$$dF = \frac{\partial F}{\partial x}dx + \frac{\partial F}{\partial y}dy.$$

Now, we introduce a new function

$$\widetilde{F}(x, y, u) = uy - F(x, y),$$

which is initially a function of (x, y, u) so that

$$d\widetilde{F} = ydu + udy - \frac{\partial F}{\partial x}dx - \frac{\partial F}{\partial y}dy.$$

However, if we let

$$u = \frac{\partial F}{\partial y}$$

then the dy term in $d\widetilde{F}$ disappears, as such $\widetilde{F}$ is now only a function of (x, u):

$$\widetilde{F}(x, u) = uy(x, u) - F(x, y(x, u)).$$

Theorem 20.1 (Legendre transform)

Suppose that

$$v_i = \nabla_{u_i} F(u_i, w_i)$$

and

$$u_i = \nabla_{v_i} G(v_i, w_i)$$

where the function $G(v_i, w_i)$ is related to the function $F(u_i, w_i)$ by the formula

$$G(v_i, w_i) = \sum_{i=1}^{N} u_i v_i - F(u_i, w_i).$$

Furthermore, the derivatives of F and G with respect to $\{w_i\}$ are related by

$$\nabla_{w_i} F(u_i, w_i) = -\nabla_{w_i} G(v_i, w_i).$$

The relationship between the functions F and G is symmetrical and each is said to be the Legendre transform of the other.

Note 20.1. By $F(u_i, w_i)$ we mean $F = F(u_1, u_2, \ldots, u_n, w_1, \ldots, w_n)$. Similarly, for $G(v_i, w_i)$.

Definition 20.3. The function $H(q_i, p_i, t)$ which is the Legendre transform of the Lagrangian function $\mathcal{L}(q_i, \dot{q}_i, t)$ is called the **Hamiltonian function**.

Theorem 20.2

The Hamiltonian function is related to the Lagrangian function by

$$H(q_i, p_i, t) = \sum_{i=1}^{N} p_i \dot{q}_i - \mathcal{L}(q_i, \dot{q}_i, t).$$

The derivatives are related as such

$$\nabla_{q_i} \mathcal{L}(q_i, \dot{q}_i, t) = -\nabla_{q_i} H(q_i, p_i, t).$$

20.2 Hamilton's equations

Theorem 20.3 (Hamilton's equations)

The n Lagrange equations (from Euler-Lagrange) are equivalent to the system of $2n$ first order ODEs

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

Proof. We evaluate the total differential of the Hamiltonian and then examine the results. Firstly, as $H \equiv H(q_i, p_i, t)$ then

$$dH = \sum_{i=1}^{N} \left(\frac{\partial H}{\partial q_i} dq_i + \frac{\partial H}{\partial p_i} dp_i + \frac{\partial H}{\partial t} dt \right).$$

Secondly, we also know $H = \sum_{i=1}^{N} \dot{q}_i p_i - \mathcal{L}$ so, we have

$$
\begin{aligned}
dH &= \sum_{i=1}^{N} \left(p_i d\dot{q}_i + \dot{q}_i dp_i - \frac{\partial \mathcal{L}}{\partial q_i} dq_i - \frac{\partial \mathcal{L}}{\partial \dot{q}_i} d\dot{q}_i - \frac{\partial \mathcal{L}}{\partial t} dt \right) \\
&= \sum_{i=1}^{N} \left(\dot{q}_i dp_i - \frac{\partial \mathcal{L}}{\partial q_i} dq_i - \frac{\partial \mathcal{L}}{\partial t} dt \right)
\end{aligned}
$$

where we have used the definition of conjugate momentum, $p_i = \dfrac{\partial \mathcal{L}}{\partial \dot{q}_i}$ to eliminate the first and fourth term of the first line. By comparing coefficients in the rwo expressions for dH we find

$$\dot{q}_i = \frac{\partial H}{\partial p_i},$$
$$-\frac{\partial \mathcal{L}}{\partial q_i} = \frac{\partial H}{\partial q_i},$$
$$\frac{\partial H}{\partial t} = -\frac{\partial \mathcal{L}}{\partial t}.$$

Note that by the Euler-Lagrange equations we have that $\dot{p}_i = \dfrac{\partial \mathcal{L}}{\partial q_i}$ so, the first two equations give

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

$\square$

20.3   Properties of the Hamiltonian

Theorem 20.4

When H has no explicit time dependence then H is a constant of the motion.

Proof. Suppose that $H = H(q_i, p_i)$ and that $\{q_i(t), p_i(t)\}$ is a motion of the system. Then, in this motion

$$\begin{aligned}
\frac{\mathrm{d}H}{\mathrm{d}t} &= \sum_{i=1}^{N} \frac{\partial H}{\partial q_i} \dot{q}_i + \sum_{i=1}^{N} \frac{\partial H}{\partial p_i} \dot{p}_i \\
&= \sum_{i=1}^{N} \frac{\partial H}{\partial q_i} \left(\frac{\partial H}{\partial p_i} \right) + \sum_{i=1}^{N} \frac{\partial H}{\partial p_i} \left(-\frac{\partial H}{\partial q_i} \right) \\
&= 0.
\end{aligned}$$

$\square$

Corollary 20.1

For a conservative system with a Hamiltonian function, $H = H(q_i, p_i)$ the total energy is given by H.

Theorem 20.5

If q_i is an ignorable coordinate (does not appear in the Hamiltonian) then p_i, the generalised momentum conjugate to q_i, is constant in any motion i.e. momentum is conserved.

20.4   Hamiltonian phase space

Suppose a mechanical system $\mathcal{S}$ has generalised coordinates q_i, conjugate momenta p_i and Hamiltonian $H(q_i, p_i, t)$. If the initial values of q_i and p_i are known then the subsequent motion of $\mathcal{S}$, described by the functions $\{q_i(t), p_i(t)\}$, is uniquely determined by Hamilton's equations. This motion can be represented geometrically by the motion of a 'point' (called a **phase point**) in Hamiltonian **phase space**. Hamiltonian phase space is a real space of $2n$ dimensions in which a 'point' is a set of values $(q_1, q_2, \ldots, q_n, p_1, p_2, \ldots, p_n)$ of the independent variables $\{q_i, p_i\}$. Each motion of the system then corresponds to the motion of a phase point through the phase space.

Definition 20.4. The **phase space** is a $2N$ dimensional space with a basis made from the N generalised coords $\{q_i\}$ and their conjugate momenta $\{p_i\}$.

20.4.1   The phase fluid

The paths of phase points have a simpler structure when the Hamiltonian is not explicitly dependent on t. In this case, H is a constant of the motion, so that each phase path must lie on a 'surface' of constant energy within the phase space. Thus, the phase space is filled with the non-intersecting level surfaces of H i.e. each phase path is restricted to one of these level surfaces.

> **Note 20.2.** If the phase space has dimension six, then a 'surface' of constant H has dimension five.

Theorem 20.6

For a system with Hamiltonian, $H = H(q_i, p_i)$ there can only be one phase path passing through any point of the phase space.

Proof. Suppose there is one phase point situated at $(q_i, p_i)_0$ at time t_1, and another phase point is at $(q_i, p_i)_0$ at time t_2. Then, since H is independent of t, the second motion can be obtained from the first by making the substitution $t \to t + t_1 - t_2$, a shift in the origin of time. Therefore, the two phase points travel along the same path with the second point delayed relative to the first by the constant time $t_2 - t_1$. Hence, **phase paths cannot intersect**. $\qquad\square$

This means that the phase space is filled with non-intersecting phase paths like the streamliens oof a fluid in steady flow. Each motion of the system corresponds to a phase point moving along on eof these paths, just as the real particels of a fluid move along the fluid streamline. Because of this analogy with fluid mechanics, the motion of phase points in phase space is called the **phase flow**.

20.5   Poisson brackets

We constructed the phase space to be always even-dimensional, as a result there exists a skew-symmetric structure known as a **symplectic structure** which is determined by Poisson brackets.

Definition 20.5. Suppose that $f(q_i, p_i)$ and $g(q_i, p_i)$ are any two functions of position in the phase space of a mechanical system. Then the **Poisson bracket**, $\{f, g\}$ of f and g is defined by

$$\{f, g\} = \nabla_{q_i} f \cdot \nabla_{p_i} g - \nabla_{p_i} f \cdot \nabla_{q_i} g$$
$$= \sum_{i=1}^{N} \left(\frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} - \frac{\partial g}{\partial q_i} \frac{\partial f}{\partial p_i} \right).$$

> **Theorem 20.7**
>
> Properties of Poisson brackets:
>
> - $\{f, f\} = 0$;
>
> - $\{f, g\} = -\{g, f\}$;
>
> - $\{af + bg, h\} = a\{f, h\} + b\{g, h\}$, for $a, b \in \mathbb{R}$;
>
> - $\{h, af + bg\} = a\{h, f\} + b\{h, g\}$, for $a, b \in \mathbb{R}$;
>
> - $\{fg, h\} = \{f, h\}g + f\{g, h\}$;
>
> - $\{f, gh\} = \{f, g\}h + g\{f, h\}$.

Theorem 20.1. The equations of motion can be written with Poisson brackets as

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

Corollary 20.1. For any function $f(q_i, p_i)$ on phase space we have that $\dot{f} = \{f, H\}$.

Proof.

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

$$= \sum_{i=1}^{N} \left(\frac{\partial f}{\partial q_i} \frac{\mathrm{d}q_i}{\mathrm{d}t} + \frac{\mathrm{d}p_i}{\mathrm{d}t} \frac{\partial f}{\partial p_1} \right)$$

$$= \frac{\mathrm{d}f}{\mathrm{d}t},$$

if $f = f(q_i, p_i)$. $\square$

20.6   Canonical transformations and symmetries

Definition 20.6. The following set of Poisson brackets are known as **fundamental** (or **canonical**) **Poisson brackets**.

$$\{q_i, p_j\} = \delta_{ij}$$
$$\{q_i, q_j\} = 0$$
$$\{p_i, p_j\} = 0.$$

Definition 20.7. The transformations

$$q_i \to q_i' \quad \text{and} \quad p_i \to p_i'$$

are known as **canonical transformations** if they preserve fundamental Poisson brackets i.e. if

$$\{q_i', p_j'\} = \delta_{ij}$$
$$\{q_i', q_j'\} = 0$$
$$\{p_i', p_j'\} = 0.$$

Definition 20.8. A **symmetry** in the Hamiltonian formulation is a canonical transformation

$$q_i \to q_i'(q_i, p_i) \quad \text{and} \quad p_i \to p'(q_i, p_i)$$

such that

$$H(q_i', p_i') = H(q_i, p_i).$$

Theorem 20.8

An infinitesimal canonical transformation is generated by any arbitrary function $f(q_i, p_i)$ (called a **generating function**) on a phase space via

$$q_i \to q_i' = q_i + \varepsilon\{q_i, f\} \equiv q_i + \delta q_i$$
$$p_i \to p_i' = p_i + \varepsilon\{p_i, f\} \equiv p_i + \delta p_i$$

to first order in ε for any function f.

Proof. We check that to first order the fundamental Poisson brackets are preserved. Firstly,

$$\{q_i', p_j'\} = \delta_{ij} + O(\varepsilon^2)$$

Theorem 20.9

TO FINIHS

$\square$

20.7 Noether's theorem - Hamiltonian formulation

Theorem 20.10 (Noether's theorem – Hamiltonian)

If an infinitesimal canonical transformation generated by a function f on the phase space is symmetry then, f is conserved.
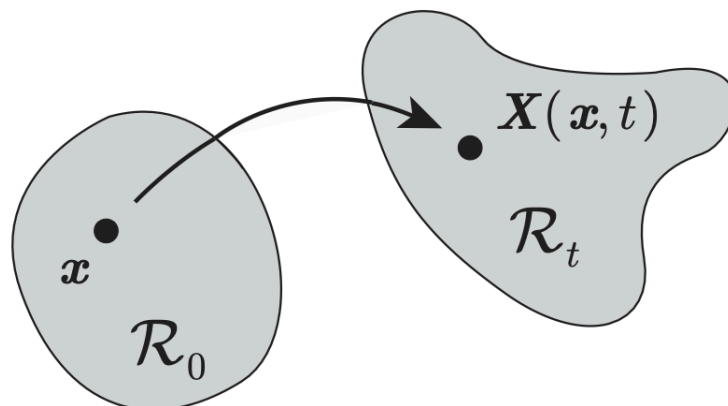

Proof. Let $\delta q_i = \varepsilon\{q_i, f\}$ and $\delta p_i = \varepsilon\{p_i, f\}$ then, we have that

$$\delta H = H(q_i + \delta q_i, p_i + \delta p_i) - H(q_i, p_i)$$

$$= \sum_{i=1}^{N} \left(\frac{\partial H}{\partial q_i} \delta q_i + \frac{\partial H}{\partial p_i} \delta p_i \right)$$

$$= \varepsilon \sum_{i=1}^{N} \left(\frac{\partial H}{\partial q_i} \{q_i, f\} + \frac{\partial H}{\partial p_i} \{p_i, f\} \right)$$

$$= \varepsilon \sum_{i=1}^{N} \left(\frac{\partial H}{\partial q_i} \frac{\partial f}{\partial p_i} - \frac{\partial H}{\partial p_i} \frac{\partial f}{\partial q_i} \right)$$

$$= \varepsilon\{H, f\}$$

$$= -\varepsilon \frac{\partial f}{\partial t}.$$

Therefore, $\delta H = 0 \iff \dfrac{\partial f}{\partial t} = 0$ i.e f is a conserved quantity. $\square$

20.8 Liouville's theorem and Poincare Recurrence

Consider phase point that at some instant, occupy the region $\mathcal{R}_0$ of the phase space, as shown by in the figure above. As t increases, tehse points move along their various phase paths in accordance with Hamilton's equations and, after time t, will occupy some new region $\mathcal{R}_t$ of the phase space. This new region will have a different shape to $\mathcal{R}_0$, but **Liouville's theorem** states that the volumes of the regions are equal.

> **Theorem 20.11** (Liouville's theorem)
>
> The volume of a region of phase spcae is constant along a Hamiltonian flow.

Remark 20.1. By the volume of the region $\mathcal{R}$ we mean

$$\text{Volume}(\mathcal{R}) = \int_{\mathcal{R}} dq_1 \cdots dq_N \, dp_1 \cdots dp_N = \int_{\mathcal{R}} dV.$$

So, in general it is a high-dimensional integral in $2N$ dimensions.

Proof. Yah $\qquad\qquad\square$

> **Theorem 20.12** (Poincare Recurrence theorem)
>
> If the phase space of a system is bounded then given any open neighbourhood $\mathcal{B}$ of a point (q_i, p_i) then a system that at (q_i, p_i) will return to $\mathcal{B}$ in a finite time.

Remark 20.2. A system either repeats its motion in phase spcae or evolves to fill up a subregion of the phase space, or all of it. In the last case the system is said to be **ergodic**.

Proof. Let $\mathcal{R}_1$ be the region occupied by the points after time τ. Suppose that $\mathcal{R}_1$ does not overlap with $\mathcal{R}_0$ so that all the points that lay in $\mathcal{R}_0$ at time $t = 0$ have left $\mathcal{R}_0$ at time $t = \tau$. We must show that some of them eventually return to $\mathcal{R}_0$. Let $\mathcal{R}_2, \mathcal{R}_3, \ldots, \mathcal{R}_n$ be the regions occupied by the same points after times $2\tau, 3\tau, \ldots, n\tau$. By Liouville's theorem, all of these regions have the same volume. Therefore, if they never overlap, their total

volume will increase without limit. But, by the assumption, all these regions lie within some finite volume, so that eventually one of them must overlap a previous one. Now we show that it must overlap with original region $\mathcal{R}_0$. $\hfill\square$

Appendix

A   REMINDER

- The conic section in the BOOK is in the section called 7.3
  THE PATH EQUATION;

- FINISH THE TENSOR SECTION – VECTOR IDENTITIES

B   Multivariable Taylor theorem

> **Theorem B.1**
>
> Taylor Theorem for two variable about the point (a, b)
>
> $$f(x, y) = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b) + \cdots$$

C   Line integrals

C.1   Line integral of a vector field

Definition C.1. The **vector line integral** of vector field $\boldsymbol{F}$ along an oriented smooth curve C is

$$\int_C \boldsymbol{F} \cdot \hat{\boldsymbol{T}} \, ds,$$

where $\hat{\boldsymbol{T}}$ is the **unit tangent vector**.

> **Note C.1.** Imagine a particle moving along the path of curve C in which this curve is subjected to forces from a vector field, $\boldsymbol{F}$. As the particle moves along the path, **work** is done by the particle. In the simplest case (one-dimensional linear motion), the work done by a particle is given by the force acting on the particle multiplied by the distance that it travels; this idea of *work* is then extended to higher dimensions with vector line integrals. In a vector field the force acting on the curve are all in different directions thus, to evaluate the *work done* we need forces which act **along** the path of curve C, such forces are the ones which are **tangent** to the curve i.e. $\boldsymbol{F} \cdot \hat{\boldsymbol{T}}$. The distance travelled by the particle is the *arc length* of C i.e. ds. Hence, the work done by the particle over a C in a vector field $\boldsymbol{F}$, is the *sum* of the force multiplied by the distance travelled i.e.
>
> $$\int_C \boldsymbol{F} \cdot \hat{\boldsymbol{T}} \, ds.$$

> **Theorem C.1**
>
> The vector line integral is equivalently defined as
>
> $$\int_C \boldsymbol{F} \cdot d\boldsymbol{r} = \int_C \boldsymbol{F} \cdot \dot{\boldsymbol{r}} \, dt.$$

Proof. The vector line integral is defined as

$$\int_C \boldsymbol{F} \cdot \hat{\boldsymbol{T}}\, ds = \int_C \boldsymbol{F} \cdot \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}s}\frac{\mathrm{d}s}{\mathrm{d}t}\, dt$$
$$= \int_C \boldsymbol{F} \cdot \dot{\boldsymbol{r}}\, dt.$$

$\square$

D   Stuff
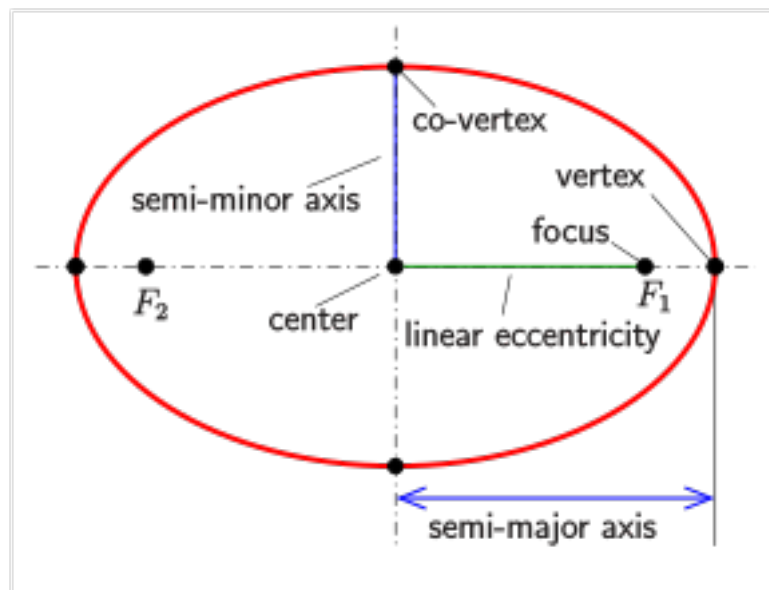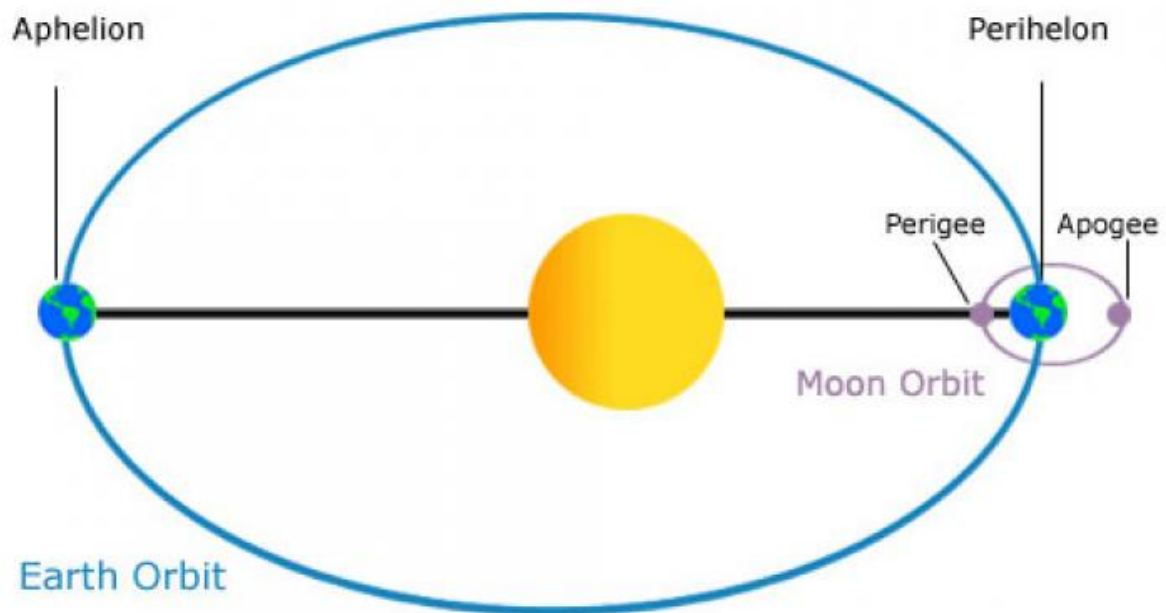
E Diagonalisation

F Links

- Moment of inertia tensor

- Parallel axes theorem

- MIT Classical Mechanics

- Matrices summation notation

- Kepler's laws