

Willoughby Seago

Theoretical Physics

Lagrangian Dynamics

September 20, 2021

COURSE NOTES

Lagrangian Dynamics

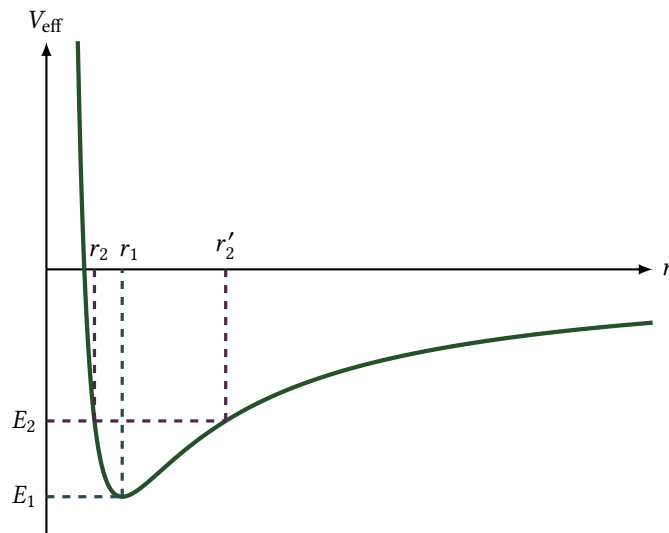
Willoughby Seago

September 20, 2021

Abstract

These are my notes from the course lagrangian dynamics. I took this course as a part of the theoretical physics degree at the University of Edinburgh.

These notes were last updated at 17:45 on 31st December 2021. For notes on other topics see <https://github.com/WilloughbySeago/Uni-Notes>.



Chapters

	Page
Chapters	ii
Contents	iv
1 Introduction	1
I Preparing for the Lagrangian Formalism	2
2 Newton's Laws and Conservation Laws	3
3 Orbits	8
4 Dynamics of a Particle System	13
5 Transforming Between Frames	17
6 Degrees of Freedom and Constraints	21
7 Generalised Quantities	25
II Lagrangian Mechanics	29
8 Lagrange's Equations	30
9 Lagrangian Methods in Action	35
10 Calculus of Variations	43
11 Symmetries and Conservation Laws	52
12 Moving Away From Newtonian Physics	62
13 Hamiltonian Dynamics	68

<i>CHAPTERS</i>	iii
III Rigid Body Motion	73
14 Euler's Method	74
15 Lagrange's Method	83
IV Oscillations and Lagrange Multipliers	91
16 Oscillations	92
17 Lagrange Multipliers	98
Appendices	101
A Differential Equations	102
B Electromagnetism	105
Bibliography	107
Acronyms	108
Index	109

Contents

	Page
Chapters	ii
Contents	iv
1 Introduction	1
I Preparing for the Lagrangian Formalism	2
2 Newton's Laws and Conservation Laws	3
2.1 Newton's Laws	3
2.1.1 Caveats	4
2.2 Conservation Laws	4
2.2.1 Conservation of Linear Momentum	4
2.2.2 Conservation of Angular Momentum	5
2.2.3 Conservation of Energy	5
3 Orbits	8
3.1 Central Force	8
3.1.1 Momenta	8
3.1.2 Coordinates	8
3.1.3 Velocity	9
3.1.4 Using Conservation Laws	9
3.1.5 Radial Velocity	10
3.2 Inverse Square Law Force	10
3.2.1 Circular Orbit	11
3.2.2 Perturbed Circular Orbit	12
3.2.3 Zones of Motion	12
4 Dynamics of a Particle System	13
4.1 Centre of Mass Motion	13
4.2 Angular Motion	14
4.3 Energy	15
5 Transforming Between Frames	17

5.1	Galilean Transformations	17
5.1.1	Centre of Momentum Frame	18
5.1.2	Intrinsic Angular Momentum	18
5.1.3	Kinetic Energy Transformation	19
5.2	Rotating Frame	19
6	Degrees of Freedom and Constraints	21
6.1	Degrees of Freedom	21
6.2	Constraint Forces	22
6.2.1	Holonomic Constraints	22
6.2.2	Nonholonomic Constraints	23
7	Generalised Quantities	25
7.1	Generalised Coordinates	25
7.1.1	Generalised Velocities	26
7.2	Generalised Force	26
7.3	Functions of $\{q\}$, $\{\dot{q}\}$, and t	27
II	Lagrangian Mechanics	29
8	Lagrange's Equations	30
8.1	Derivation	30
8.2	The Lagrangian	32
8.2.1	Notes	33
8.2.2	Generalisations	33
9	Lagrangian Methods in Action	35
9.1	Single Particle	35
9.1.1	Cartesian Coordinates	35
9.1.2	Plane Polar Coordinates	35
9.2	Atwood's Machine	36
9.2.1	Atwood's Monkey	37
9.3	Particle and Wedge	38
9.4	Pendula	39
9.4.1	Simple Pendulum	39
9.4.2	Spherical Pendulum	40
10	Calculus of Variations	43
10.1	Functionals	43
10.2	Euler's Equation	43
10.3	First Integrals	45
10.4	Examples	46
10.4.1	Shortest Path	46
10.4.2	Brachistochrone	47
10.4.3	Soap Film	48
10.5	Calculus of Variations for Many Variables	49
10.6	Hamilton's Principle	50
10.6.1	Hamilton's Principle from Newtonian Mechanics	50
11	Symmetries and Conservation Laws	52

11.1	Invariance Properties of the Lagrangian	52
11.1.1	Invariance Under Change of Generalised Coordinates	52
11.1.2	Invariance Under Addition of a Total Derivative	53
11.2	Generalised Momenta	54
11.3	The Energy Function	55
11.3.1	Energy Function and the Total Energy	56
11.4	Global Symmetries and Conservation Laws	58
11.4.1	Homogeneity	58
11.4.2	Isotropy	59
11.4.3	Notes	61
12	Moving Away From Newtonian Physics	62
12.1	Velocity-Dependent Forces	62
12.2	The Lorentz Force	63
12.2.1	Electromagnetic Lagrangian	63
12.2.2	Properties Derived from the Lagrangian	64
12.3	Special Relativity	65
12.4	Charged Relativistic Particle in an EM Field	66
12.4.1	Covariant Form	66
13	Hamiltonian Dynamics	68
13.1	Classical Mechanics Hamiltonian	68
13.1.1	Poisson Brackets	70
13.2	Quantum Mechanics	71
13.2.1	The Hamiltonian and Quantum Mechanics	71
13.2.2	The Lagrangian and Quantum Mechanics	72
III	Rigid Body Motion	73
14	Euler's Method	74
14.1	Recap	74
14.2	Euler's Theorem	74
14.3	The Inertia Tensor	76
14.4	Other Origins	77
14.5	The Eulerian Approach	78
14.6	Symmetric Top	79
14.7	Asymmetric Top	81
15	Lagrange's Method	83
15.1	The Lagrangian	83
15.2	Conservation Laws and Equations of Motion	86
15.3	The Zone Function	87
15.4	Steady Precession	88
15.5	Fast Top	88
15.6	Nutation	89
15.7	Sleeping Top	89

IV Oscillations and Lagrange Multipliers	91
16 Oscillations	92
16.1 Static Equilibrium	92
16.2 The Lagrangian	93
16.3 Triatomic Molecule	94
16.4 Simultaneous Diagonalisation	95
17 Lagrange Multipliers	98
17.1 Holonomic Example	99
17.2 Non-Holonomic Example	99
Appendices	101
A Differential Equations	102
B Electromagnetism	105
B.1 Maxwell's Equations	105
B.2 Potentials	105
Bibliography	107
Acronyms	108
Index	109

One

Introduction

The Lagrangian formalism of classical mechanics was pioneered by Joseph-Louis Lagrange. It is an alternative to Newton's formulation of classical mechanics. In many ways it is more powerful than Newton's formulation. For example, the Lagrangian formalism can easily include constraints on the system and is readily adapted to apply to electromagnetism, quantum mechanics, and relativistic systems. A task which is either impossible or incredibly difficult for the Newtonian formalism. Another benefit is that the Lagrangian formalism is built upon scalar quantities, which are often easier to work with than vectors.

Part I

Preparing for the Lagrangian Formalism

Two

Newton's Laws and Conservation Laws

2.1 Newton's Laws

Newtonian mechanics is based on Newton's laws, a set of three laws that combined allow us to describe mechanics at everyday scales, that is when things are: not too fast (so no need for special relativity), not too massive (so no need for general relativity), and not too small (so no need for quantum mechanics).

Before we can state Newton's laws we will need a few bits of notation and basic definitions. For now we consider point like particles and measure their position relative to some origin, O .

Notation 2.1.1 — Basic Quantities

- \mathbf{r} denotes the position vector.
- \mathbf{v} denotes the velocity vector.
- \mathbf{a} denotes the acceleration vector.
- \mathbf{p} denotes the linear momentum vector.
- Given some differentiable quantity, f , we denote by \dot{f} the total time derivative, df/dt , and by \ddot{f} the second total time derivative, d^2f/dt^2 . We should not have use of higher time derivatives but if we do then the it should be clear how this dot notation generalises.

Definition 2.1.2 — Basic Quantities

- The **velocity** is defined by $\mathbf{v} := \dot{\mathbf{r}}$.
- The **acceleration** is defined by $\mathbf{a} = \dot{\mathbf{v}} = \ddot{\mathbf{r}}$.
- The **momentum** is defined by $\mathbf{p} = m\mathbf{v}$.

We can now state Newton's laws.

N1. Newton's first law (N1) states that in the absence of forces \mathbf{v} is constant.

N2. **Newton's second law** (N2) states that in the presence of a force, F , the rate of change of linear momentum is $\dot{\mathbf{p}} = F$. A special case of this is when the mass is constant, which we will assume unless stated otherwise, in which case we recover the famous

$$\dot{\mathbf{p}} = m\dot{\mathbf{v}} = m\mathbf{a} = F. \quad (2.1.3)$$

In Newtonian mechanics we start most problems by assuming that the force, F is either known or can be easily found. We then solve N2 for the equations of motion.

N3. **Newton's third law** (N3) comes in two forms, a weak form and a strong form. In both we consider the force F_{ab} , which is the force on particle a due to particle b . The weak form states that $F_{ab} = -F_{ba}$. The strong form states that as well as this F_{ab} is directed along $\mathbf{r}_{ab} = \mathbf{r}_a - \mathbf{r}_b$ where \mathbf{r}_a is the position of a and \mathbf{r}_b the position of b .

2.1.1 Caveats

Newton's laws all come with caveats. Obviously, due to the need for quantum mechanics and relativity, Newton's laws aren't always valid. So, we need to be careful about when we apply them.

N1 and N2 are only valid in non-accelerating frames, that is when there are no external forces. We call these **inertial frames**. We typically assume frames are inertial unless stated otherwise. Apart from linearly accelerating frames it should be noted that rotating frames aren't inertial, even for a constant rotation speed. This, as well as the movement around the sun, and movement of the solar system, and movement of the solar system, and the movement of the galaxy and, ..., means that the Earth is *not* an inertial frame. However, it is a sufficiently good approximation of an inertial frame for most of our everyday requirements. There are a few cases, such as Foucault's pendulum, where the rotation of the Earth has a non-negligible effect.

Both forms of N3 only hold for types of force. For example, since the force from a magnetic field is orthogonal to the velocity, means that in general the strong form of N3 doesn't hold for charged particles.

2.2 Conservation Laws

Conservation laws are *incredibly* important in physics. One advantage of the Lagrangian formalism is how easy it is to find conservation laws with it. However, for now we restrict our selves to deriving the most important conservation laws from the Newtonian formalism.

2.2.1 Conservation of Linear Momentum

Suppose there is no force, so $F = 0$. It immediately follows from N2 that $\dot{\mathbf{p}} = F = 0$ and hence \mathbf{p} is constant. We say that in the absence of external forces linear momentum is conserved. A consequence of this is that if m is a constant then $\mathbf{a} = 0$ and so the particle travels at a constant velocity.

2.2.2 Conservation of Angular Momentum

Definition 2.2.1 — Angular Momentum and Torque For a particle at position \mathbf{r} and linear momentum \mathbf{p} we define its **angular momentum** with respect to the origin to be

$$\mathbf{L} := \mathbf{r} \times \mathbf{p}. \quad (2.2.2)$$

If a force, \mathbf{F} , is applied to the particle then we define the **torque**, also known as the **moment of force** about the origin to be

$$\mathbf{G} := \mathbf{r} \times \mathbf{F}. \quad (2.2.3)$$

Note that since they are defined via cross products both \mathbf{L} and \mathbf{G} are axial vectors (aka pseudovectors), meaning that if we change between a left and right handed coordinate system then these vectors change sign.

Applying N2 to the definition of torque we have

$$\mathbf{G} = \mathbf{r} \times \mathbf{F} = \mathbf{r} \times \dot{\mathbf{p}} = \frac{d}{dt}(\mathbf{r} \times \mathbf{p}) - \dot{\mathbf{r}} \times \mathbf{p} \quad (2.2.4)$$

where in the last step we have used the chain rule:

$$\frac{d}{dt}(\mathbf{r} \times \mathbf{p}) = \mathbf{r} \times \dot{\mathbf{p}} + \dot{\mathbf{r}} \times \mathbf{p} \implies \mathbf{r} \times \dot{\mathbf{p}} = \frac{d}{dt}(\mathbf{r} \times \mathbf{p}) - \dot{\mathbf{r}} \times \mathbf{p}. \quad (2.2.5)$$

We now look at this last term, $\dot{\mathbf{r}} \times \mathbf{p}$, in more detail. Noticing that $\dot{\mathbf{r}} = \mathbf{v}$ and $\mathbf{p} = m\mathbf{v}$ we see that this term is zero, since the cross product of any two collinear vectors is zero. Hence,

$$\mathbf{G} = \frac{d}{dt}(\mathbf{r} \times \mathbf{p}) = \dot{\mathbf{L}} \quad (2.2.6)$$

where we have used the definition of angular momentum, $\mathbf{L} = \mathbf{r} \times \mathbf{p}$. From this we see that if $\mathbf{G} = \mathbf{0}$ then $\dot{\mathbf{L}} = \mathbf{0}$ and hence \mathbf{L} is constant. We say that in the absence of external torques angular momentum is conserved.

2.2.3 Conservation of Energy

Definition 2.2.7 — Work Done The **work done** to move a particle along a path, Γ , requiring force \mathbf{F} is defined as

$$W := \int_{\Gamma} \mathbf{F} \cdot d\mathbf{r}. \quad (2.2.8)$$

Consider the work done to move a particle from A to B along some path, which we parametrise as a function of t (which may or may not be time) such that when $t = t_A$ the particle is at A and when $t = t_B$ the particle is at B . We will assume a constant mass but the following result holds for a variable mass, the derivation is

just more involved.

$$W_{AB} = \int_A^B \mathbf{F} \cdot d\mathbf{r} \quad (2.2.9)$$

$$= \int_{t_A}^{t_B} \mathbf{F} \cdot \frac{d\mathbf{r}}{dt} dt \quad (2.2.10)$$

$$= \int_{t_A}^{t_B} \dot{\mathbf{p}} \cdot \mathbf{v} dt \quad (2.2.11)$$

$$= \int_{t_A}^{t_B} m\dot{\mathbf{v}} \cdot \mathbf{v} dt \quad (2.2.12)$$

$$= \frac{m}{2} \int_{t_A}^{t_B} v^2 dt \quad (2.2.13)$$

$$= \frac{m}{2} [v^2]_{t_A}^{t_B} \quad (2.2.14)$$

$$= \frac{1}{2} m [v(t_B)]^2 - \frac{1}{2} [v(t_A)]^2 \quad (2.2.15)$$

$$= T_B - T_A \quad (2.2.16)$$

where $T := \frac{1}{2}mv^2$ defines the **kinetic energy** in Newtonian mechanics. Note that about halfway through the calculation we used

$$\frac{d}{dt}(\mathbf{v} \cdot \mathbf{v}) = \mathbf{v} \cdot \dot{\mathbf{v}} + \dot{\mathbf{v}} \cdot \mathbf{v} = 2\dot{\mathbf{v}} \cdot \mathbf{v} \implies \dot{\mathbf{v}} \cdot \mathbf{v} = \frac{1}{2} \frac{d}{dt}(\mathbf{v} \cdot \mathbf{v}). \quad (2.2.17)$$

We can see from this that the work done on a particle is equal to the change in kinetic energy of the particle.

Definition 2.2.18 — Conservative Force The following are equivalent:

- The force \mathbf{F} is **conservative**.
- The work done to move a particle from A to B is independent of the path taken, that is

$$\int_{\Gamma} \mathbf{F} \cdot d\mathbf{r} = \int_{\gamma} \mathbf{F} \cdot d\mathbf{r} \quad (2.2.19)$$

where Γ and γ are two paths from A to B .

- The following holds

$$\oint \mathbf{F} \cdot d\mathbf{r} = 0 \quad (2.2.20)$$

where \oint denotes an integral over a closed path, that is a path that starts and ends at the same point.

- There exists a scalar function, V , such that

$$\mathbf{F} = -\nabla V. \quad (2.2.21)$$

- The following holds

$$\nabla \times \mathbf{F} = \mathbf{0}. \quad (2.2.22)$$

Suppose F is a conservative force. Then we can write $F = -\nabla V$ for some scalar function V . The negative sign here is just a convention so that we can minimise the potential, instead of maximising, which better aligns with our intuition and the original definition of potentials before this more general notion. We can use this to redo the calculation above:

$$W_{AB} = \int_A^B \mathbf{F} \cdot d\mathbf{r} \quad (2.2.23)$$

$$= - \int_A^B \nabla V \cdot d\mathbf{r} \quad (2.2.24)$$

$$= - \int_A^B dV \quad (2.2.25)$$

$$= V_B - V_A \quad (2.2.26)$$

where V_A and V_B are the value of the potential at A and B respectively. Note here that we used

$$\int \nabla V \cdot d\mathbf{r} = \int dV \quad (2.2.27)$$

which is the three dimensional analogue of

$$\int \frac{df}{dx} dx = \int df \quad (2.2.28)$$

Combining [Equations \(2.2.16\)](#) and [\(2.2.26\)](#) we get

$$T_A + V_A = T_B + V_B \quad (2.2.29)$$

we define $E := T + V$ to be the **total energy** and since this equation holds for any two points A and B this shows that, for a conservative force, the total energy is conserved. Since most, but not all, forces we meet in classical mechanics are conservative this is a very useful property.

Three

Orbits

In this section we apply Newton's laws and conservation laws to one of the systems that motivated their development: orbiting planets. More generally we will consider some particle moving in a central force field described by a **central potential** $V(\mathbf{r}) = V(r)$. Our goal is to find an equation for the radial velocity of an orbiting particle.

3.1 Central Force

Recall that a **central force** is one which depends only on the magnitude, r , of the position vector, \mathbf{r} and acts along the line to the origin: $\mathbf{F}(\mathbf{r}) = F(r)\hat{\mathbf{r}} = -\nabla V(r)$. Here $\hat{\mathbf{r}} = \mathbf{r}/r$ is a unit vector in the direction of the position vector.

3.1.1 Momenta

Assuming that $\mathbf{F} \neq \mathbf{0}$ linear momentum will not be conserved. We can easily calculate the torque:

$$\mathbf{G} = \mathbf{r} \times \mathbf{F} = \mathbf{r} \times \hat{\mathbf{r}}F = \mathbf{0} \quad (3.1.1)$$

since the cross product of two collinear vectors is identically zero. Since there is no torque the angular momentum, \mathbf{L} , is conserved. In particular the magnitude of the angular momentum, L , is a constant. We have that

$$\mathbf{r} \cdot \mathbf{L} = \mathbf{r} \cdot (\mathbf{r} \times \mathbf{p}) = 0 \quad (3.1.2)$$

since $\mathbf{r} \times \mathbf{p}$ is perpendicular to \mathbf{r} . What we have here is of the form \mathbf{r} dotted with a constant vector equals a constant. This defines a plane perpendicular to the constant vector. The physical interpretation of this is that the particle is constrained to motion in the plane with normal \mathbf{L} containing the origin.

3.1.2 Coordinates

Since the particle is constrained to a plane we work in cylindrical polar coordinates, (r, φ, z) , with the $\hat{\mathbf{z}}$ -axis perpendicular to the plane of motion. In these coordinates

$$x = r \cos \varphi, \quad (3.1.3)$$

$$y = r \sin \varphi, \quad (3.1.4)$$

$$z = z, \quad (3.1.5)$$

where $r = |\mathbf{r}| = \sqrt{x^2 + y^2}$. In the plane of motion $z = 0$.

We define two orthonormal basis vectors, \mathbf{e}_r and \mathbf{e}_φ , which are parallel and perpendicular, respectively, to the position vector, \mathbf{r} . At the point (r, φ, z) these vectors are given in terms of the Cartesian unit vectors, \mathbf{e}_x , \mathbf{e}_y , and \mathbf{e}_z , by

$$\mathbf{e}_r := \cos \varphi \mathbf{e}_x + \sin \varphi \mathbf{e}_y \quad (3.1.6)$$

$$\mathbf{e}_\varphi := -\sin \varphi \mathbf{e}_x + \cos \varphi \mathbf{e}_y. \quad (3.1.7)$$

The position vector is then

$$\mathbf{r} = x\mathbf{e}_x + y\mathbf{e}_y = r\mathbf{e}_r. \quad (3.1.8)$$

3.1.3 Velocity

The velocity of the particle can be found by differentiating the position vector with respect to time:

$$\mathbf{v} = \dot{\mathbf{r}} \quad (3.1.9)$$

$$= \frac{d}{dt}(r\mathbf{e}_r) \quad (3.1.10)$$

$$= \dot{r}\mathbf{e}_r + r\dot{\mathbf{e}}_r \quad (3.1.11)$$

Notice that, unlike Cartesian unit vectors, the cylindrical unit vectors are time dependent¹ and so we have a second term involving the derivative of the unit vector. We can easily work out what this is by differentiating Equation (3.1.6):

¹rather, they're position dependent and the position is time dependent

$$\dot{\mathbf{e}}_r = \frac{d}{dt}(\cos \varphi \mathbf{e}_x + \sin \varphi \mathbf{e}_y) \quad (3.1.12)$$

$$= -\dot{\varphi} \sin \varphi \mathbf{e}_x + \dot{\varphi} \cos \varphi \mathbf{e}_y \quad (3.1.13)$$

$$= \dot{\varphi} \mathbf{e}_\varphi. \quad (3.1.14)$$

Substituting this into the velocity we have

$$\dot{\mathbf{r}} = \dot{r}\mathbf{e}_r + r\dot{\varphi}\mathbf{e}_\varphi. \quad (3.1.15)$$

We can see from this that the radial velocity is $v_r = \dot{r}$, and the tangential component of the velocity is $v_\varphi = r\dot{\varphi}$.

Be careful: $\dot{r} = v_r$ is not the same as $|\dot{\mathbf{r}}| = v$ here, for example, if motion is in a circle then $\dot{r} = 0$ but $\dot{\mathbf{r}} \neq \mathbf{0}$.

3.1.4 Using Conservation Laws

We can calculate the angular momentum now:

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} \quad (3.1.16)$$

$$= \mathbf{r} \times \dot{\mathbf{r}}m \quad (3.1.17)$$

$$= mr[\mathbf{e}_r \times (\dot{r}\mathbf{e}_r + r\dot{\varphi}\mathbf{e}_\varphi)] \quad (3.1.18)$$

$$= \mathbf{0} + mr^2\dot{\varphi}\mathbf{e}_r \times \mathbf{e}_\varphi \quad (3.1.19)$$

$$= mr^2\dot{\varphi}\mathbf{e}_z. \quad (3.1.20)$$

Since the angular momentum is conserved we know that $L = mr^2\dot{\varphi}$ is constant. Inverting this relationship we have $\dot{\varphi} = L/(mr^2)$, which will be useful in a moment.

The kinetic energy is easy to find once we have an expression for the velocity:

$$T = \frac{1}{2}mv^2 \quad (3.1.21)$$

$$= \frac{1}{2}m\dot{\mathbf{r}} \cdot \dot{\mathbf{r}} \quad (3.1.22)$$

$$= \frac{1}{2}m(\dot{r}^2 + r^2\dot{\phi}^2). \quad (3.1.23)$$

The total energy is then

$$E = T + V \quad (3.1.24)$$

$$= \frac{1}{2}m(\dot{r}^2 + r^2\dot{\phi}^2) + V(r) \quad (3.1.25)$$

$$= \frac{1}{2}m\dot{r}^2 + \frac{L^2}{2mr^2} + V(r) \quad (3.1.26)$$

where we have substituted for $\dot{\phi} = L/(mr^2)$.

Notice that we can now write the total energy as

$$E = \frac{1}{2}m\dot{r}^2 + V_{\text{eff}}(r) \quad (3.1.27)$$

where

$$V_{\text{eff}}(r) = \frac{L^2}{2mr^2} + V(r) \quad (3.1.28)$$

is an **effective potential**, that is the particle's motion in the radial direction (since we consider a 'kinetic energy' $m\dot{r}^2/2$, which considers motion only along the radial direction) behaves as if it is a particle constrained to one-dimensional motion in the potential V_{eff} .

3.1.5 Radial Velocity

Solving Equation (3.1.26) for \dot{r} we find

$$\dot{r} = \pm \left[\frac{2}{m} \left(E - V(r) - \frac{L^2}{2mr^2} \right) \right]^{1/2}. \quad (3.1.29)$$

Here E and L are conserved quantities fixed by the initial conditions. Since \dot{r} is real the quantity in square brackets on the right must be non-negative for all values of r . From this we see that $E \geq V_{\text{eff}}(r)$ for all r .

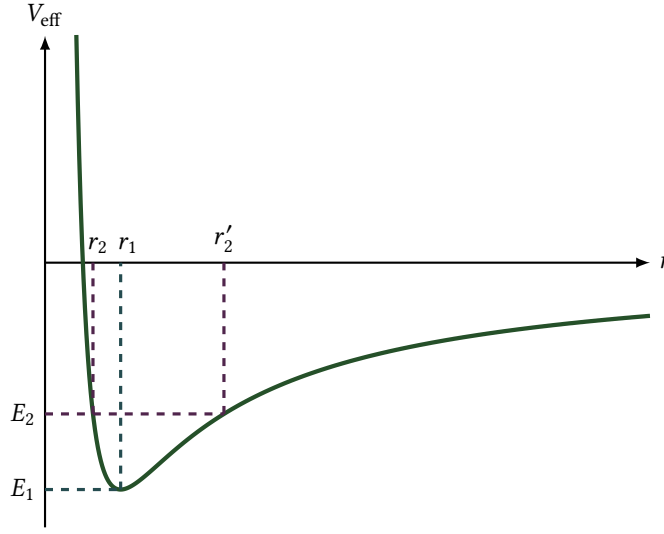
3.2 Inverse Square Law Force

We now restrict ourselves to the special case of

$$V(r) = -\frac{k}{r}. \quad (3.2.1)$$

The force from this is

$$\mathbf{F}(r) = -\frac{k}{r^2}\hat{\mathbf{r}}, \quad (3.2.2)$$

Figure 3.1: The effective potential, V_{eff} .

which is an inverse square law. Hence this potential can describe gravitational attraction or electrostatic interactions. The effective potential in this case is

$$V_{\text{eff}}(r) = \frac{L^2}{2mr^2} - \frac{k}{r}. \quad (3.2.3)$$

This is plotted in [Figure 3.1](#).

Now consider the particular case where $E = E_1 = \min V_{\text{eff}}(r)$. In this case the particle is constrained to have $r = r_1$ where r_1 is the value that minimises V_{eff} . This corresponds to a stable circular orbit.

If instead $E = E_2 > E_1$ then the particle will be able to move in the radial direction. If $E < 0$ then the particle will be constrained between two values r_2 and r'_2 which are such that $V_{\text{eff}}(r_2) = V_{\text{eff}}(r'_2) = E_2$. At r_2 and r'_2 we must have $\dot{r} = 0$ since these are turning points.

We can treat both of these two cases a bit more formally by considering the effective equation of motion for motion in the radial direction. We can find this using energy conservation since this implies $\dot{E} = 0$:

$$\frac{dE}{dt} = \frac{d}{dt} \left(\frac{1}{2} m \dot{r}^2 + V_{\text{eff}}(r) \right) \quad (3.2.4)$$

$$= m \dot{r} \ddot{r} + \frac{dV_{\text{eff}}}{dr} \dot{r}. \quad (3.2.5)$$

Since this holds for all \dot{r} we must have

$$m \ddot{r} = - \frac{dV_{\text{eff}}}{dr}. \quad (3.2.6)$$

3.2.1 Circular Orbit

Suppose the particle is in a circular orbit with some fixed radius, $r = r_1$. Then $\dot{r} = 0$. This requires that $E = V_{\text{eff}}(r_1) = E_1$. Therefore a stable circular orbit is possible where the effective potential is a minima or maxima.

Alternatively we know that r is constant so \dot{r} and \ddot{r} are zero meaning that [Equation \(3.2.6\)](#) gives

$$0 = \left. \frac{dV_{\text{eff}}}{dr} \right|_{r=r_1}. \quad (3.2.7)$$

3.2.2 Perturbed Circular Orbit

Suppose we have a circular orbit of radius r_1 , and we perturb it by moving the particle a small amount, ε , outwards. The radial velocity of the particle is simply

$$\dot{r} = \frac{d}{dt}(r_1 + \varepsilon) = \dot{\varepsilon} \quad (3.2.8)$$

For small ε we can Taylor expand the effective potential about $r = r_1$:

$$V_{\text{eff}}(r) = V_{\text{eff}}(r_1 + \varepsilon) \quad (3.2.9)$$

$$= V_{\text{eff}}(r_1) + \varepsilon \left. \frac{dV_{\text{eff}}}{dr} \right|_{r=r_1} + \frac{\varepsilon^2}{2} \left. \frac{d^2V_{\text{eff}}}{dr^2} \right|_{r=r_1} + O(\varepsilon^3) \quad (3.2.10)$$

$$= V_{\text{eff}}(r_1) + \frac{k}{2} \varepsilon^2 + O(\varepsilon^3). \quad (3.2.11)$$

Here we have used the fact that $V'_{\text{eff}}(r_1) = 0$ as this r_1 is the location of the minimum. $V_{\text{eff}}(r_1)$ is just a constant and we have defined $k = V''_{\text{eff}}(r_1)$. Ignoring terms $O(\varepsilon^3)$ or smaller and dropping the constant $V_{\text{eff}}(r_1)$ term we have

$$E = \frac{1}{2} m \dot{\varepsilon}^2 + \frac{k}{2} \varepsilon^2. \quad (3.2.12)$$

If $k > 0$, i.e. r_1 is a local minimum, then this is the energy of a harmonic oscillator of mass m and spring constant k . The small perturbation then leads to the radial position oscillating with angular frequency $\omega = \sqrt{k/m}$ about r_1 .

If $k < 0$, i.e. r_1 is a local maximum, then the circular orbit is unstable and the perturbation will cause it to turn into some other kind of orbit.

If $k = 0$ then we can't ignore the $O(\varepsilon^3)$ terms in the Taylor expansion.

3.2.3 Zones of Motion

We showed in the previous example that motion was restricted to the zone $r_{\min} \leq r \leq r_{\max}$ where the boundaries are the roots of $V_{\text{eff}}(r) = E$.

This can be generalised by noticing that [Equation \(3.1.29\)](#) is of the form

$$\dot{r} = \pm \sqrt{g(r)}, \quad (3.2.13)$$

in this case $g(r) = 2(E - V_{\text{eff}}(r))/m$. Since \dot{r} is strictly real this requires $g(r)$ be positive. Therefore whatever happens (barring things like quantum tunnelling) the motion is restricted to the zone $r_{\min} \leq r \leq r_{\max}$ where the boundaries are roots of $g(r) = 0$ at which g crosses the axis. There may be multiple regions in which motion can occur but the system will not be able to move between them unaided.

Analysing various zones of motion can be a powerful tool that allows us to find information about a system without needing to solve for the full trajectory, which is often a difficult task. To show that a coordinate is restricted to some zone first look for equations of the form $\dot{r} = \pm \sqrt{g(r)}$.

Four

Dynamics of a Particle System

In this section we will consider a system of N particles, each labelled by some index $a = 1, \dots, N$. We suppose that the a th particle has constant mass m_a and its position is \mathbf{r}_a .

We start with Newton's second law for an individual particle:

$$m_a \ddot{\mathbf{r}}_a = \mathbf{F}_a. \quad (4.0.1)$$

The force \mathbf{F}_a is due to two different sources. First there is the external force, $\mathbf{F}_a^{\text{ext}}$, which may be, for example, gravity. There is also then the force due to interactions with all of the other particles, perhaps due to an electromagnetic interaction. We use \mathbf{F}_{ab} to denote the force on particle a due to particle b . The total internal force on a is then given by summing over this interaction for all particles distinct from a . The net force on the particle is then

$$\mathbf{F}_a = \mathbf{F}_a^{\text{ext}} + \sum_{b \neq a} \mathbf{F}_{ab}. \quad (4.0.2)$$

4.1 Centre of Mass Motion

Summing Newton's law over all particles we get

$$\sum_a m_a \ddot{\mathbf{r}}_a = \sum_a \mathbf{F}_a = \sum_a \mathbf{F}_a^{\text{ext}} + \sum_a \sum_{b \neq a} \mathbf{F}_{ab}. \quad (4.1.1)$$

Focussing on this last term with a double sum we see that it is possible to write it as a single sum over pairs of particles,

$$\sum_a \sum_{b \neq a} \mathbf{F}_{ab} = \sum_{\text{pairs}} (\mathbf{F}_{ab} + \mathbf{F}_{ba}). \quad (4.1.2)$$

To see this consider the explicit example of three particles:

$$\sum_{a=1}^3 \sum_{\substack{b \in \{1,2,3\} \\ b \neq a}} \mathbf{F}_{ab} = \mathbf{F}_{12} + \mathbf{F}_{13} + \mathbf{F}_{21} + \mathbf{F}_{23} + \mathbf{F}_{31} + \mathbf{F}_{32} \quad (4.1.3)$$

$$= (\mathbf{F}_{12} + \mathbf{F}_{21}) + (\mathbf{F}_{13} + \mathbf{F}_{31}) + (\mathbf{F}_{23} + \mathbf{F}_{32}). \quad (4.1.4)$$

It follows, assuming the weak form of Newton's third law, that $F_{ba} = -F_{ab}$ and hence

$$\sum_a \sum_{b \neq a} F_{ab} = \sum_{\text{pairs}} (F_{ab} + F_{ba}) = \sum_{\text{pairs}} (F_{ab} - F_{ab}) = 0. \quad (4.1.5)$$

So, all internal forces cancel pairwise. Hence

$$\sum_a m_a \ddot{\mathbf{r}}_a = \sum_a \mathbf{F}_a^{\text{ext}} =: \mathbf{F}^{\text{ext}} \quad (4.1.6)$$

where we have defined \mathbf{F}^{ext} to be the net external force on the system. Defining the total mass

$$M := \sum_a m_a \quad (4.1.7)$$

we can define the **centre of mass** to be

$$\mathbf{R} := \frac{1}{M} \sum_a m_a \mathbf{r}_a. \quad (4.1.8)$$

We can think of it as the average position of the particles weighted by their masses. Differentiating this definition twice leads to

$$M \ddot{\mathbf{R}} = \mathbf{F}^{\text{ext}}. \quad (4.1.9)$$

We can also define the total linear momentum

$$\mathbf{P} := \sum_a \mathbf{p}_a = \sum_a m_a \dot{\mathbf{r}}_a = M \dot{\mathbf{R}}. \quad (4.1.10)$$

From this we see that

$$\dot{\mathbf{P}} = \mathbf{F}^{\text{ext}}. \quad (4.1.11)$$

This means that the centre of mass acts as a point particle obeying Newton's second law in the presence of the net external force. We use this implicitly whenever we do dynamics since all bodies are really made up of many particles but we almost exclusively consider rigid bodies to be point particles at their centre of mass.

4.2 Angular Motion

Starting with Newton's second law, this time in the form $\dot{\mathbf{p}}_a = \mathbf{F}_a$, we can take the cross product with \mathbf{r}_a and sum over all particles giving

$$\sum_a \mathbf{r}_a \times \dot{\mathbf{p}}_a = \sum_a \mathbf{r}_a \times \mathbf{F}_a. \quad (4.2.1)$$

Using the trick of recognising the left hand side as half of a product rule we have

$$\mathbf{r}_a \times \dot{\mathbf{p}}_a = \frac{d}{dt} (\mathbf{r}_a \times \mathbf{p}_a) - \dot{\mathbf{r}}_a \times \mathbf{p}_a = \frac{d\mathbf{L}_a}{dt} - \dot{\mathbf{r}}_a \times \mathbf{p}_a. \quad (4.2.2)$$

The second term is zero since $\mathbf{p}_a = m_a \dot{\mathbf{r}}_a$ and hence we have

$$\dot{\mathbf{L}} := \sum_a \mathbf{L}_a = \sum_a \mathbf{r}_a \times \dot{\mathbf{p}}_a. \quad (4.2.3)$$

Returning to the right hand side of Equation (4.2.1) we have

$$\sum_a \mathbf{r}_a \times \mathbf{F}_a = \sum_a \mathbf{r}_a \times \mathbf{F}_a^{\text{ext}} + \sum_a \sum_{b \neq a} \mathbf{r}_a \times \mathbf{F}_{ab}. \quad (4.2.4)$$

Using the weak form of Newton's third law we can again move to a sum over pairs of particles and the last term becomes

$$\sum_a \sum_{b \neq a} \mathbf{r}_a \times \mathbf{F}_{ab} = \sum_{\text{pairs}} (\mathbf{r}_a \times \mathbf{F}_{ab} + \mathbf{r}_b \times \mathbf{F}_{ba}) = \sum_{\text{pairs}} (\mathbf{r}_a - \mathbf{r}_b) \times \mathbf{F}_{ab}. \quad (4.2.5)$$

Assuming the strong form of Newton's third law this is $\mathbf{0}$ since $\mathbf{r}_a - \mathbf{r}_b$ and \mathbf{F}_{ab} both point along the line between the two particles. In this case the total angular momentum obeys

$$\dot{\mathbf{L}} = \sum_a \mathbf{r}_a \times \mathbf{F}_a^{\text{ext}} = \sum_a \mathbf{G}_a^{\text{ext}} =: \mathbf{G}^{\text{ext}} \quad (4.2.6)$$

where $\mathbf{G}_a^{\text{ext}}$ is the external torque on the a th particle and \mathbf{G}^{ext} is the net external torque. That is, these torques are due to external forces.

It should be noted that both \mathbf{L} and \mathbf{G} depend on our choice of origin.

4.3 Energy

For a single particle we have

$$\mathbf{F} \cdot d\mathbf{r} = d\left(\frac{1}{2}mv^2\right) = dT. \quad (4.3.1)$$

Here T is the kinetic energy. For N particles we then have

$$dT = \sum_a dT_a \quad (4.3.2)$$

$$= \sum_a \mathbf{F}_a \cdot d\mathbf{r}_a \quad (4.3.3)$$

$$= \sum_a \mathbf{F}_a^{\text{ext}} \cdot d\mathbf{r}_a + \sum_{\text{pairs}} \mathbf{F}_{ab} \cdot (d\mathbf{r}_a - d\mathbf{r}_b). \quad (4.3.4)$$

Again, we have converted a sum over a and $b \neq a$ into a sum over pairs of particles.

If the external force is conservative then there is a potential function of the form $V(\mathbf{r}_1, \dots, \mathbf{r}_N)$ such that

$$\mathbf{F}_a^{\text{ext}} = -\nabla_a V \quad (4.3.5)$$

where ∇_a is the gradient with respect to \mathbf{r}_a . It then follows that

$$\sum_a \mathbf{F}_a^{\text{ext}} \cdot d\mathbf{r}_a = -dV. \quad (4.3.6)$$

If the internal forces are conservative then there is a potential function of the form $U(\mathbf{r}_1, \dots, \mathbf{r}_N)$ such that

$$\mathbf{F}_{ab} = -\nabla_a U. \quad (4.3.7)$$

The weak form of Newton's third law then gives us

$$\mathbf{F}_{ba} = -\nabla_b U = \nabla_a U. \quad (4.3.8)$$

The only way that this can hold is if the potential depends only on the vector $\mathbf{r}_{ab} = \mathbf{r}_a - \mathbf{r}_b$, which connects the two particles.

Defining ∇_{ab} as the gradient with respect to \mathbf{r}_{ab} this becomes

$$\mathbf{F}_{ab} = -\nabla_{ab} U = \nabla_{ba} U. \quad (4.3.9)$$

We then have

$$\sum_{\text{pairs}} \mathbf{F}_{ab} \cdot (\mathbf{dr}_a - \mathbf{dr}_b) = - \sum_{\text{pairs}} \nabla_{ab} U \cdot \mathbf{dr}_{ab} = -dU. \quad (4.3.10)$$

Combining the results of this section we have

$$dT = -dV - dU \implies dT + dV + dU = 0. \quad (4.3.11)$$

What this means is that the total change in energy, which is the change in kinetic energy, and external and internal potential energies, is zero. So, assuming conservative forces and the weak form of Newton's third law energy is conserved.

Five

Transforming Between Frames

5.1 Galilean Transformations

A **Galilean transformation** is a non-relativistic transformation between inertial frames. Suppose that frame S' moves at constant velocity, \mathbf{V} , with respect to the frame S and that the origin of S' is related to the origin of S by $\mathbf{O}' = \mathbf{O} + \mathbf{b}$ at time $t = 0$. Note that it is common to assume $\mathbf{b} = \mathbf{0}$ and the origins coincide at $t = 0$, and also that the axes of the two frames are parallel and motion is in the x direction. This is called the standard configuration but we will deal with this slightly more general case.

Let \mathbf{r} be a position vector in frame S , and let \mathbf{r}' be the same position in frame S' . Notice that these two vectors refer to the same point in different frames. In the language of coordinate transforms this means that we transform between frames with passive transformations. It follows from the geometry of our set up that at time $t = 0$ $\mathbf{r} = \mathbf{r}' + \mathbf{b}$. At some later time t we then have

$$\mathbf{r} = \mathbf{r}' + \mathbf{b} + t\mathbf{V}. \quad (5.1.1)$$

This is the Galilean transformation for positions.

Differentiating this we get

$$\mathbf{v} = \mathbf{v}' + \mathbf{V} \quad (5.1.2)$$

where \mathbf{v} and \mathbf{v}' are the velocity of the point in frames S and S' , respectively. This is the Galilean transformation for velocities.

Suppose now that this point is the position of one of our particles from the previous chapter. Multiplying the velocity transform by the mass, m_a , we have

$$\mathbf{p}_a = \mathbf{p}'_a + m_a\mathbf{V} \quad (5.1.3)$$

where \mathbf{p}_a and \mathbf{p}'_a are the momentum of the particle as measured in frames S and S' respectively. Differentiating this we have

$$\dot{\mathbf{p}}_a = \dot{\mathbf{p}}'_a = \mathbf{F}_a, \quad (5.1.4)$$

which shows us that force is Galilean invariant.

From this the transformation for the properties of the entire system follow. In particular

$$\mathbf{P} = \mathbf{P}' + M\mathbf{V}, \quad \text{and} \quad \dot{\mathbf{P}} = \dot{\mathbf{P}}' = \mathbf{F}^{\text{ext}}. \quad (5.1.5)$$

This means that Newton's laws are unaffected by a Galilean transformation and are the same in any inertial frame. That is they are Galilean invariant.

5.1.1 Centre of Momentum Frame

Definition 5.1.6 — Centre of Momentum Frame We define the **centre of momentum frame** (CoM frame) as the *inertial* frame chosen *instantaneously* such that

$$\mathbf{P}' = \sum_a \mathbf{p}'_a = \sum_a m_a \dot{\mathbf{r}}'_a = 0. \quad (5.1.7)$$

That is, the total momentum is instantaneously zero.

The centre of mass frame is only defined at a single time, this is so that it is an inertial frame.

We often choose the centre of mass to be the origin of the centre of momentum frame, that is $\mathbf{R}' = 0$. Taking the position of each particle and multiplying by its mass before summing over all particles gives us

$$\sum_a m_a \mathbf{r}_a = \sum_a m_a \mathbf{r}'_a + \sum_a m_a t\mathbf{V} + \sum_a m_a \mathbf{b} \quad (5.1.8)$$

or in terms of global properties

$$M\mathbf{R} = M\mathbf{R}' + Mt\mathbf{V} + M\mathbf{b}. \quad (5.1.9)$$

Taking the origin as the centre of mass we find

$$\mathbf{b} = \mathbf{R} - t\mathbf{V}. \quad (5.1.10)$$

A related concept is the **centre of mass frame** which is the frame with its origin at the centre of mass at all times. This frame will be non-inertial if there is a net external force.

5.1.2 Intrinsic Angular Momentum

The angular momentum of a single particle at time $t = 0$ is given by

$$\mathbf{L}_a = \mathbf{r}_a \times \mathbf{p}_a = (\mathbf{r}'_a + \mathbf{b}) \times \mathbf{p}'_a = \mathbf{r}'_a \times \mathbf{p}'_a + \mathbf{b} \times \mathbf{p}'_a = \mathbf{L}'_a + \mathbf{b} \times \mathbf{p}'_a. \quad (5.1.11)$$

The total angular momentum at some time t is then

$$\mathbf{L} = \sum_a \mathbf{r}_a \times \mathbf{p}_a \quad (5.1.12)$$

$$= \sum_a (\mathbf{r}'_a + t\mathbf{V} + \mathbf{b}) \times (\mathbf{p}'_a + m_a \mathbf{V}) \quad (5.1.13)$$

$$= \mathbf{L}' + M\mathbf{R}' \times \mathbf{V} + (t\mathbf{V} + \mathbf{b}) \times \mathbf{P}' + M(t\mathbf{V} + \mathbf{b}) \times \mathbf{V}. \quad (5.1.14)$$

Define $\mathbf{J} = \mathbf{L}'$ to be the **intrinsic angular momentum**, that is the angular momentum due to the *internal* motion of the particles in the centre of momentum frame. In this frame we also have $\mathbf{P}' = 0$, $\mathbf{R}' = 0$, $t\mathbf{V} + \mathbf{b} = \mathbf{R}$, and $\mathbf{P} = M\mathbf{V}$ and so we get

$$\mathbf{L} = \mathbf{J} + \mathbf{R} \times \mathbf{P}. \quad (5.1.15)$$

What this shows is that the angular momentum in some frame, S , is the intrinsic angular momentum of the particles, plus the angular momentum due to the motion

of the centre of mass. From this we also see that L depends on the choice of origin (unless $\mathbf{P} = \mathbf{0}$).

Performing a similar calculation for the external torque we find

$$\mathbf{G}^{\text{ext}} = (\mathbf{G}^{\text{ext}})' + \mathbf{b} \times \mathbf{F}^{\text{ext}}. \quad (5.1.16)$$

Notice that the extra term from moving the origin *doesn't* vanish in the torque expression and therefore we must always specify the origin of the centre of momentum frame when we discuss torque.

5.1.3 Kinetic Energy Transformation

In frame S the kinetic energy is

$$T = \sum_a \frac{1}{2} m_a v_a^2 \quad (5.1.17)$$

$$= \sum_a \frac{1}{2} m_a (\mathbf{v}'_a + \mathbf{V}) \cdot (\mathbf{v}'_a + \mathbf{V}) \quad (5.1.18)$$

$$= \sum_a \frac{1}{2} m_a (v'_a)^2 + \sum_a m_a \mathbf{v}'_a \cdot \mathbf{V} + \sum_a \frac{1}{2} m_a V^2 \quad (5.1.19)$$

$$= T' + \sum_a m_a \mathbf{v}'_a \cdot \mathbf{V} + \frac{1}{2} M V^2. \quad (5.1.20)$$

Taking S' to be the centre of momentum frame this central term vanishes is $\mathbf{P}' \cdot \mathbf{V} = 0$, and we have

$$T = T' + \frac{1}{2} M V^2 \quad (5.1.21)$$

So, the kinetic energy in some frame, S , is the kinetic energy in the centre of momentum frame plus a contribution due to motion of the centre of mass.

5.2 Rotating Frame

Occasionally we will have need to work in a rotating frame. Note that rotating frames are necessarily non-inertial so the work from earlier on in this chapter *does not apply*. We will consider only the simplest example of a frame, R , rotating with respect to some inertial frame, S , with constant angular velocity, $\boldsymbol{\omega}$. Further we set up the axes and time such that at time $t = 0$ the axes and origins of R and S are coincident. We will use the notation $[f]_X$ to denote the quantity f as measured in frame X .

Consider a position vector in R , $[\mathbf{r}]_R$. Quantities without subscripts are taken to have been measured at time $t = 0$ and hence are the same in both frames. If this is stationary in R then from S it is seen as rotating with velocity

$$[\mathbf{v}]_S = \boldsymbol{\omega} \times \mathbf{r} \quad (5.2.1)$$

where $\mathbf{r} = [\mathbf{r}(t = 0)]_R = [\mathbf{r}(t = 0)]_S$ is the position at time $t = 0$. If the position instead has some nonzero velocity in R then the velocity in S is

$$[\dot{\mathbf{r}}]_S = [\dot{\mathbf{r}}]_R + \boldsymbol{\omega} \times \mathbf{r}. \quad (5.2.2)$$

Consider a general vector, \mathbf{A} . In a short time, dt , in the rotating frame

$$[\mathbf{A}]_R \rightarrow [\mathbf{A}]_R + [d\mathbf{A}]_R, \quad (5.2.3)$$

and in the stationary frame

$$[\mathbf{A}]_S \rightarrow ([\mathbf{A}]_R + [\mathbf{dA}]_R) + \boldsymbol{\omega} \times \mathbf{A}dt. \quad (5.2.4)$$

Together these imply that

$$[\mathbf{dA}]_S = \mathbf{dA} + \boldsymbol{\omega} \times \mathbf{A}dt. \quad (5.2.5)$$

Or, equivalently

$$[\dot{\mathbf{A}}]_S = [\dot{\mathbf{A}}]_R + \boldsymbol{\omega} \times \mathbf{A}. \quad (5.2.6)$$

Setting $\mathbf{A} = \mathbf{r}$ we recover the equation for velocity transformations in a rotating frame.

Now consider the case of $\mathbf{A} = m[\dot{\mathbf{r}}]_S$. We have

$$m \left[\frac{d}{dt} [\dot{\mathbf{r}}]_S \right]_S = m \left[\frac{d}{dt} [\dot{\mathbf{r}}]_S \right]_S. \quad (5.2.7)$$

Hence

$$m[\ddot{\mathbf{r}}]_S = m \left[\frac{d}{dt} ([\dot{\mathbf{r}}]_R + \boldsymbol{\omega} \times \mathbf{r}) \right]_R + m\boldsymbol{\omega} \times [[\dot{\mathbf{r}}]_R + \boldsymbol{\omega} \times \mathbf{r}]_S. \quad (5.2.8)$$

The left hand side is simply the mass times the acceleration in the inertial frame, S , and hence we can apply Newton's second law and write this term as \mathbf{F} , the force on the particle. We can expand the right hand side and we get

$$\mathbf{F} = m[\ddot{\mathbf{r}}]_R + m\boldsymbol{\omega} \times [\dot{\mathbf{r}}]_R + m\boldsymbol{\omega} \times \boldsymbol{\omega} \times [\dot{\mathbf{r}}]_R + m\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}). \quad (5.2.9)$$

Rearranging this we have

$$m[\ddot{\mathbf{r}}]_R = \mathbf{F} - m\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}) - 2m\boldsymbol{\omega} \times [\dot{\mathbf{r}}]_R \quad (5.2.10)$$

which is equivalent to

$$m\mathbf{a} = \mathbf{F} - m\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}) - 2m\boldsymbol{\omega} \times \mathbf{v} \quad (5.2.11)$$

where \mathbf{v} and \mathbf{a} are the velocity and acceleration in the rotating frame.

This is similar in form to Newton's second law and so we identify the two extra terms as **fictitious forces**. The first is the **centrifugal force** and the second the **Coriolis force**. While these aren't real forces, they arise due to an attempt to treat the non-inertial rotating frame as inertial, they do have observable effects. The centrifugal force always acts to cause objects to fly outwards in the rotating frame. The Coriolis force acts to cause objects moving in a plane perpendicular to the axis of rotation to be deflected at right angles.

Six

Degrees of Freedom and Constraints

6.1 Degrees of Freedom

Definition 6.1.1 — Degrees of Freedom The **degrees of freedom** of a system are a set of variables the values of which are precisely enough to specify the state of the system.

■ **Example 6.1.2 — Single Particle** A single particle has three degrees of freedom. These correspond to its location. There are several equally valid choices of degrees of freedom. We could use the three Cartesian coordinates, cylindrical coordinates, spherical coordinates, or many more.

The degrees of freedom aren't always as simple as the coordinates of the particles.

■ **Example 6.1.3 — Pendulum** A simple pendulum has only one degree of freedom, which we most commonly take to be the angle to the vertical. We see from this that constraints, such as being on the end of a rod of fixed length that can only move in a plane, can reduce the number of degrees of freedom.

■ **Example 6.1.4 — Rigid Body** A **rigid body** is a collection of particles fixed such that the particles positions relative to each other are fixed. Despite being made of arbitrarily many particles a rigid body has only six degrees of freedom. We need to specify location and orientation, this then fixes the position of all particles. Most commonly we do this by defining the centre of mass, which takes three degrees of freedom, and the orientation, which can be elegantly done with three Euler angles.

The example of a rigid body shows how we can greatly reduce the number of variables to deal with by considering degrees of freedom instead of the coordinates of every particle.

6.2 Constraint Forces

We have already seen one constraint in our examples above, the pendulum was constrained to motion in a plane that kept it a fixed distance from the point at which it is connected. Another example of a constraint may be a particle on the floor, or constrained to move within a box.

Consider some particle, a , with position vector \mathbf{r}_a , moving subject to some constraint. The motion is constrained by a constraint force, $\mathbf{F}_{\text{constraint}}$. This force acts at right angles to the motion and hence do no work for an instantaneous displacement since $\mathbf{F}_{\text{constraint}} \cdot \delta \mathbf{r} = 0$ for some small displacement, $\delta \mathbf{r}$, we will make this more rigorous shortly. An example here might be a ball constrained to roll down a tube. The walls of the tube exert a normal force on the ball only when it is in contact. When the ball is in contact with the wall it can only move parallel to the wall and hence perpendicular to the force.

6.2.1 Holonomic Constraints

Definition 6.2.1 — Holonomic Constraint A holonomic constraint is one which can be written as

$$f(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N, t) = 0. \quad (6.2.2)$$

Here f is an algebraic function of the coordinates and time. Specifically this is *not* a differential equation or an inequality.

An example of a holonomic constraint might be a particle on the floor. This is equivalent to the constraint that $z = 0$, where we choose the origin to be on the floor and the z -axis to be vertical. Another example is two particles constrained to be a fixed distance, ρ , apart, since we can write this as $|\mathbf{r}_a - \mathbf{r}_b| = \rho$.

Holonomic constraints can include an explicit time dependence, for example a particle on the floor of a lift is constrained to $z = vt$ where v is the speed of the lift which started at $z = 0$ at time $t = 0$.

■ **Example 6.2.3 — Rolling Cylinder** Consider a cylinder of radius a rolling down an inclined plane. We can set up coordinates such that x' points down the plane and y' is normal to the plane. In general there are three degrees of freedom. The position of the centre of the cylinder, (x', y') , and the angle through which it is rotated, ϑ .

If we impose **rolling conditions**, that is that the sphere doesn't bounce, so $y' = a$, and doesn't slide, so $x' = a\vartheta + C$ for some constant offset C . This gives us two holonomic constraints. Notice that the number of degrees of freedom is reduced to 1. We can specify either x' or ϑ , and are free to choose which, as the other degrees of freedom will then be fixed.

In general each holonomic constraint decreases the number of degrees of freedom by one.

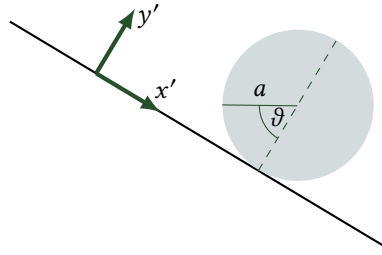


Figure 6.1: A cylinder of radius a rolls down an inclined plane. There are initially three degrees of freedom, x' , y' , and ϑ , which is reduced to one degree of freedom by imposing rolling conditions.

6.2.2 Nonholonomic Constraints

Unsurprisingly **nonholonomic constraints** are those which aren't holonomic, that is they can't be written as a simple algebraic equation of the coordinates and time. Most often they come in the form of inequalities or differential equations.

■ **Example 6.2.4 — Inequality Constraint** A particle is constrained to move in a rectangular box of width a and height b . This is a nonholonomic constraint since the constraint is equivalent to

$$0 \leq x \leq a, \quad \text{and} \quad 0 \leq y \leq b. \quad (6.2.5)$$

Notice that this is a two dimensional problem and as such there are initially two degrees of freedom. After imposing the constraint there are still two degrees of freedom.

■ **Example 6.2.6 — Differential Constraint** Consider a hoop of radius a rolling along a plane. This has five degrees of freedom, the location of the centre of the hoop, (x, y, z) , the angle through which the hoop is rotated, φ , and the direction in which the hoop rolls, for example the angle of this line to the x -axis, ϑ .

Unlike the one-dimensional case introducing rolling constraints only reduces the number of degrees of freedom by 1. We have now fixed $z = a$. Intuitively this is because the extra freedom of an second dimension allows us to return to the same spot with a different angle of rotation by rolling in a circle with a non-integer ratio of its radius to a . Mathematically this is because the constraints we now have are that a small change in the angle, $d\varphi$, corresponds to a change $dr = ad\varphi$, where dr is a small displacement along the direction of rolling. Decomposing into x and y changes we have that $dx = a \cos \vartheta d\varphi$ and $dy = a \sin \vartheta d\varphi$. This gives the nonholonomic constraints

$$\dot{x} = a\dot{\varphi} \cos \vartheta, \quad \text{and} \quad \dot{y} = a\dot{\varphi} \sin \vartheta. \quad (6.2.7)$$

These constraints *cannot* be integrated to get holonomic constraints as ϑ is a function of time.

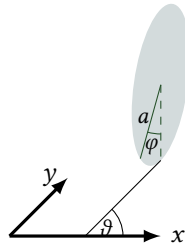


Figure 6.2: A hoop rolling in the plane has five initial degrees of freedom, x , y , z , ϕ , and ϑ . Introducing rolling constraints reduces this to four degrees of freedom by fixing $z = a$.

Note that we can write the rolling constraint for the one-dimensional example in a differential equation form as $\dot{x} = a\dot{\vartheta}$, but this can immediately be integrated to give $x = a\vartheta + C$ for some constant C and so this is still a holonomic constraint.

Seven

Generalised Quantities

Often our degrees of freedom are coordinates, in particular Cartesian coordinates. However, this isn't always the case, for example the angle through which a body has rotated is not a Cartesian coordinate. Our goal in this chapter is to come up with a generalisation for several classical quantities, specifically coordinates, velocities, and force, which keep the key properties but allow us to be more general. We will also see generalised momentum in later chapters.

Notation 7.0.1 Given some indexed quantity, x_i , with $i = 1, \dots, N$ we will use $\{x\}$ as shorthand for x_1, x_2, \dots, x_N .

7.1 Generalised Coordinates

Consider some system of particles in three dimensions. Particle a has three coordinates, $\mathbf{r}_a = (x_a, y_a, z_a)$. Given a list of all of the particles coordinates,

$$x_1, y_1, z_1, x_2, y_2, z_2, \dots, x_N, y_N, z_N, \quad (7.1.1)$$

we have $3N$ coordinates, for which we use the shorthand x_i with $i = 1, \dots, 3N$.

In the presence of holonomic constraints these x_i are not all independent. If we have M holonomic constraints then there are $3N - M$ independent variables, q_i , which describe the state of the system at a given time. We can view these as functions of the form

$$q_i = q_i(x_1, x_2, \dots, x_{3N}, t) = q_i(\{x\}, t), \quad \text{for } i = 1, \dots, 3N - M. \quad (7.1.2)$$

A general aim in this course will be to generate $3N - M$ second order differential equations in these q_i . These are the **equations of motion** (EoM). The variables $\{q\}$ are our **generalised coordinates**.

We assume that the relation in [Equation \(7.1.2\)](#) is invertible and so we have

$$x_i = x_i(q_1, q_2, \dots, q_{3N-M}, t) = x_i(\{q\}, t), \quad \text{for } i = 1, \dots, 3N. \quad (7.1.3)$$

This allows us to recover information about the original system by solving the equations of motion for $\{q\}$. The important point is that q_i can be individually varied, which we will use to derive the equations of motion, whereas we cannot vary all x_i without changing at least one other x_j .

7.1.1 Generalised Velocities

The **generalised velocities** are simply the time derivatives of the generalised coordinates. That is, \dot{q}_i , or collectively $\{\dot{q}\}$.

It should be noted that until we apply the equations of motion we can treat $\{q\}$ and $\{\dot{q}\}$ independently as we are considering arbitrary motion of the system. This means that we can just treat $\{\dot{q}\}$ as another set of variables until we start solving the equations of motion.

7.2 Generalised Force

Consider an infinitesimal change in the coordinates, $x_i \rightarrow x_i + dx_i$. This is a **generalised displacement**. A simple application of the chain rule gives us

$$dx_i = \sum_j \frac{\partial x_i}{\partial q_j} dq_j + \frac{\partial x_i}{\partial t} dt. \quad (7.2.1)$$

We define a **virtual displacement** to be an instantaneous, infinitesimal displacement with no dt component. To distinguish from a normal displacement we write this as δx_i . Again applying the chain rule, with the dt component vanishing, we have

$$\delta x_i = \sum_j \frac{\partial x_i}{\partial q_j} dq_j. \quad (7.2.2)$$

The work done in a virtual displacement due to forces F_i in the direction of δx_i is

$$\delta W = \sum_i F_i \delta x_i. \quad (7.2.3)$$

We can split the force into a constraint force, and other sources of force, F_i^{other} . Since constraint forces do no work the work done reduces to

$$\delta W = \sum_i F_i^{\text{other}} \delta x_i. \quad (7.2.4)$$

We can write this in terms of $\{\delta q\}$ using the chain rule:

$$\delta W = \sum_j \sum_i F_i^{\text{other}} \frac{\partial x_i}{\partial q_j} \delta q_j = \sum_j Q_j \delta q_j \quad (7.2.5)$$

where we have defined the **generalised force**

$$Q_j := \sum_i F_i^{\text{other}} \frac{\partial x_i}{\partial q_j}. \quad (7.2.6)$$

It should be noted that generalised quantities do not necessarily have the same dimensions as the quantity they generalise. For example an angle is dimensionless but may be a valid generalised coordinate, q , where we would expect coordinates to have dimensions of length. The work done has units of energy so in this case the generalised force, Q , must have units of energy so that the product qQ has units of energy.

■ **Example 7.2.7 — Particle on a Hoop** A particle of mass m slides under gravity around a smooth vertical hoop of radius a . Taking the centre of the hoop as the origin in Cartesian coordinates we have two coordinates, (x, z) , to describe the position of the particle. We also have one holonomic constraint, that $x^2 + z^2 = a^2$. Hence, there is one degree of freedom. The sensible choice for a generalised coordinate here is an angle. For no particular reason we choose the angle to the vertical measuring anticlockwise. This quantity, ϑ , is related to the Cartesian position by

$$x = -a \sin \vartheta, \quad \text{and} \quad z = a \cos \vartheta. \quad (7.2.8)$$

The force on the particle consists of a constraint force from the hoop, and the force due to gravity, $\mathbf{F}^{\text{other}} = (0, -mg)$. The work done on the particle in a virtual displacement, $\delta\vartheta$, is

$$\delta W = \left[F_x^{\text{other}} \frac{\partial x}{\partial \vartheta} + F_z^{\text{other}} \frac{\partial z}{\partial \vartheta} \right] \delta\vartheta \quad (7.2.9)$$

$$= 0 + (-mg) \frac{\partial}{\partial \vartheta} (a \cos \vartheta) \delta\vartheta \quad (7.2.10)$$

$$= (mga \sin \vartheta) \delta\vartheta. \quad (7.2.11)$$

From this we identify the generalised force

$$Q_\vartheta = mga \sin \vartheta. \quad (7.2.12)$$

This has dimensions of energy, so ϑQ_ϑ has units of energy also.

7.3 Functions of $\{q\}$, $\{\dot{q}\}$, and t

We will often deal with functions of $\{q\}$, $\{\dot{q}\}$, and t . For example the kinetic energy,

$$T = \sum_i \frac{1}{2} m_i \dot{x}_i^2. \quad (7.3.1)$$

Since $x_i = x_i(\{q\}, t)$ in generalised coordinates the kinetic energy can depend explicitly on $\{q\}$ and $\{\dot{q}\}$, for example in polar coordinates we saw that the kinetic energy was

$$T = \frac{1}{2} m (\dot{r}^2 + r^2 \dot{\phi}^2). \quad (7.3.2)$$

This depends on the generalised coordinate r and the generalised velocities \dot{r} and $\dot{\phi}$.

For such functions we can write a small increment, df , as

$$df = \sum_j \frac{\partial f}{\partial q_j} dq_j + \sum_j \frac{\partial f}{\partial \dot{q}_j} d\dot{q}_j + \frac{\partial f}{\partial t} dt. \quad (7.3.3)$$

Note that we are treating $\{q\}$ and $\{\dot{q}\}$ as independent variables here since we have not yet imposed the equations of motion.

Notation 7.3.4 When a function, f , nominally depends on the variables a, b, c, \dots but actually is independent of a given variable, say b , and we wish to emphasise this we will do so by writing $f(a, c, \dots) = f(a, \cancel{b}, c, \dots)$.

We now consider the special case where there is no dependence on the generalised velocities, that is $f = f(\{q\}, t) = f(\{q\}, \cancel{\dot{q}}, t)$. For this function $\partial f / \partial \dot{q}_j = 0$ and so

$$df = \sum_j \frac{\partial f}{\partial q_j} dq_j + \frac{\partial f}{\partial t} dt. \quad (7.3.5)$$

This means that

$$\dot{f} = \frac{df}{dt} = \sum_j \frac{\partial f}{\partial q_j} \frac{dq_j}{dt} + \frac{\partial f}{\partial t} = \sum_j \frac{\partial f}{\partial q_j} \dot{q}_j + \frac{\partial f}{\partial t}. \quad (7.3.6)$$

From this we see that $\dot{f} = \dot{f}(\{q\}, \{\dot{q}\}, t)$. Note that the partial derivatives of f in this expression are also independent of the velocities, the only contribution is from the dq_j terms in df :

$$\frac{\partial f}{\partial q_j} = f_{q_j}(\{q\}, \cancel{\dot{q}}, t), \quad \text{and} \quad \frac{\partial f}{\partial t} = f_t(\{q\}, \cancel{\dot{q}}, t). \quad (7.3.7)$$

Taking the partial derivative of \dot{f} with respect to \dot{q}_j we have

$$\frac{\partial \dot{f}}{\partial \dot{q}_j} = \frac{\partial}{\partial q_j} \left[\sum_j \frac{\partial f}{\partial q_j} \dot{q}_j + \frac{\partial f}{\partial t} \right] = \frac{\partial f}{\partial q_j}. \quad (7.3.8)$$

This gives us the so called **cancellation of dots** rule:

$$\frac{\partial \dot{f}}{\partial \dot{q}_j} = \frac{\partial f}{\partial q_j}. \quad (7.3.9)$$

This *only* applies if f is independent of $\{\dot{q}\}$.

Another result we will need is

$$\frac{d}{dt} \left(\frac{\partial f}{\partial q_j} \right) = \frac{\partial}{\partial q_j} \left(\frac{df}{dt} \right) = \frac{\partial \dot{f}}{\partial q_j} \quad (7.3.10)$$

referred to as the **commuting derivatives rule**.

Part II

Lagrangian Mechanics

Eight

Lagrange's Equations

Finally, the time has come to derive the key object of our studies in this course and also the equations that make it so useful. We shall do so by a method that is rather displeasing since it starts with a random quantity that just so happens to be useful. Later in the course we will see a derivation using the calculus of variations that doesn't require a premonition of the answer.

8.1 Derivation

We start by noting that the Cartesian coordinates are functions of the generalised coordinates and *not* the generalised velocities: $x_i = x_i(\{q\}, \dot{q}, t)$. Because of this we can use the cancellation of dots rule (Equation (7.3.9)):

$$\frac{\partial \dot{x}_i}{\partial \dot{q}_j} = \frac{\partial x_i}{\partial q_j}. \quad (8.1.1)$$

This holds for all i and j .

Now consider the magic quantity

$$\mathcal{M} = \frac{d}{dt} \left(m_i \dot{x}_i \frac{\partial x_i}{\partial \dot{q}_j} \right). \quad (8.1.2)$$

Note that we are *not* employing the Einstein summation convention in this course.

Applying the cancellation of dots rule we get

$$\mathcal{M} = \frac{d}{dt} \left(m_i \dot{x}_i \frac{\partial x_i}{\partial q_j} \right) \quad (8.1.3)$$

$$= \frac{d}{dt} (m_i \dot{x}_i) \frac{\partial x_i}{\partial q_j} + m_i \dot{x}_i \frac{d}{dt} \left(\frac{\partial x_i}{\partial q_j} \right) \quad (8.1.4)$$

$$= \frac{d}{dt} (m_i \dot{x}_i) \frac{\partial x_i}{\partial q_j} + m_i \dot{x}_i \frac{\partial \dot{x}_i}{\partial q_j} \quad (8.1.5)$$

$$= \frac{d}{dt} (m_i \dot{x}_i) \frac{\partial x_i}{\partial q_j} + \frac{d}{dq_j} \left(\frac{1}{2} m_i \dot{x}_i^2 \right). \quad (8.1.6)$$

Here we have applied the commuting derivatives rule (Equation (7.3.10)) since x_i has no dependence on the velocities. The last equality can be shown by expanding out the last derivative with the product rule.

Instead of using the cancellation of dots rule we can instead write the magic quantity as

$$\mathcal{M} = \frac{d}{dt} \left[\frac{d}{d\dot{q}_j} \left(\frac{1}{2} m_i \dot{x}_i^2 \right) \right]. \quad (8.1.7)$$

Again, this can be shown by expanding the partial derivative.

Equating these two forms of the magic quantity we have

$$\mathcal{M} = \frac{d}{dt} \left[\frac{d}{d\dot{q}_j} \left(\frac{1}{2} m_i \dot{x}_i^2 \right) \right] = \frac{d}{dt} (m_i \dot{x}_i) \frac{\partial x_i}{\partial q_j} + \frac{d}{dq_j} \left(\frac{1}{2} m_i \dot{x}_i^2 \right). \quad (8.1.8)$$

Using Newton's second law we identify

$$\frac{d}{dt} (m_i \dot{x}_i) = \dot{p}_i = F_i. \quad (8.1.9)$$

Similarly we can identify

$$T_i = \frac{1}{2} m_i \dot{x}_i^2 \quad (8.1.10)$$

as the kinetic energy due to motion along the i th coordinate. Hence the magic quantity is

$$\mathcal{M} = \frac{d}{dt} \left(\frac{\partial T_i}{\partial \dot{q}_j} \right) = F_i \frac{\partial x_i}{\partial q_j} + \frac{\partial T_i}{\partial q_j}. \quad (8.1.11)$$

Next we sum over all i . In doing so since derivatives are linear operators we end up with

$$T(\{q\}, \{\dot{q}\}, t) = \sum_i T_i = \sum_i \frac{1}{2} m_i \dot{x}_i^2, \quad (8.1.12)$$

which is the total kinetic energy. We can also identify

$$Q_j(\{q\}, \{\dot{q}\}, t) = \sum_i F_i \frac{\partial x_i}{\partial q_j}, \quad (8.1.13)$$

the generalised force conjugate to the q_j coordinate. Note that the constraint forces don't contribute to this so we are free to replace F_i with F_i^{other} , the non-constraint forces. The constraint forces are accounted for when we reduce the number of variables by moving to generalised coordinates.

Putting this all together we get

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_j} \right) - \frac{\partial T}{\partial q_j} = Q_j. \quad (8.1.14)$$

These¹ are **Lagrange's equations of motion** in their most general form. They are derived from Newton's laws and are completely equivalent to them. However, it is often easier to use these equations as the constraint forces are automatically included.

¹yes these, there are $3N - M$ equations here, one for every $j = 1, \dots, 3N - M$.

■ **Example 8.1.15 — Particle on a Hoop** Consider the particle on a hoop from [Example 7.2.7](#). The particles speed is $a\dot{\vartheta}$ and so the kinetic energy of the particle is

$$T(\vartheta, \dot{\vartheta}, t) = \frac{1}{2}ma^2\dot{\vartheta}^2. \quad (8.1.16)$$

We have $\partial T/\partial \vartheta = 0$ and $\partial T/\partial \dot{\vartheta} = ma^2\dot{\vartheta}$. Putting this in Lagrange's equation of motion we get

$$\frac{d}{dt}(ma^2\dot{\vartheta}) = mga \sin \vartheta. \quad (8.1.17)$$

Other methods, such as conservation of energy, will give us the same result, although with more work.

8.2 The Lagrangian

Recall that in an instantaneous virtual displacement the work done is

$$\delta W = \sum_j Q_j \delta q_j. \quad (8.2.1)$$

Suppose we only have conservative forces. Then there is a potential energy function, V , such that

$$\delta W = -\delta V. \quad (8.2.2)$$

²we will treat velocity dependent potentials separately later, an example of such a potential is the potential energy of an electromagnetic field for a charged particle.

Suppose also that $V = V(\{q\}, \{\dot{q}\}, t)$, i.e. there is no velocity dependence². We then have

$$\delta W = \sum_j Q_j \delta q_j = -\delta V = -\sum_j \frac{\partial V}{\partial q_j} \delta q_j. \quad (8.2.3)$$

Since the generalised coordinates are by definition independent we can vary them individually keeping the others fixed. The only way to be able to do this and have the result above hold is if we have

$$Q_j = -\frac{\partial V}{\partial q_j} \quad (8.2.4)$$

for all j . Notice the similarity to $F = -\nabla V$, except that instead of having to consider the gradient, which becomes horrible if our coordinates are independent, such as for spherical coordinates, we can consider just single derivatives.

Lagrange's equations can then be written as

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_j} \right) = \frac{\partial T}{\partial q_j} - \frac{\partial V}{\partial q_j}. \quad (8.2.5)$$

Since $V = V(\{q\}, \{\dot{q}\}, t)$ we have $\partial V/\partial \dot{q}_j = 0$ so we are free to insert terms proportional to this at will. Doing so we get

$$\frac{d}{dt} \left(\frac{d}{dq_j} (T - V) \right) - \frac{d}{dq_j} (T - V) = 0. \quad (8.2.6)$$

Introducing the quantity

$$\mathcal{L} := T - V \quad (8.2.7)$$

we get the form of Lagrange's equations for conservative forces with a velocity independent potential, called the **Euler-Lagrange equations**:

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_j} \right) - \frac{\partial \mathcal{L}}{\partial q_j} = 0. \quad (8.2.8)$$

These hold for

$$\mathcal{L}(\{q\}, \{\dot{q}\}, t) = T(\{q\}, \{\dot{q}\}, t) - V(\{q\}, t). \quad (8.2.9)$$

These are the key equations of the entire course.

8.2.1 Notes

Notice the minus sign in $\mathcal{L} = T - V$. Its not clear *why* the difference between kinetic and potential energy is such a useful quantity, but it is as we shall see shortly.

The time derivative is a total derivative, and therefore must be applied using the chain rule for $f(\{q\}, \{\dot{q}\}, t)$:

$$\frac{df}{dt} = \sum_j \frac{\partial f}{\partial q_j} \dot{q}_j + \sum_j \frac{\partial f}{\partial \dot{q}_j} \ddot{q}_j + \frac{\partial f}{\partial t}. \quad (8.2.10)$$

The Lagrangian is always differentiated to get the equations of motion. Therefore \mathcal{L} and $\mathcal{L} + C$ will give the same equations for any constant C . In particular we are free to ignore any additive constant in the potential energy, as we are accustomed to doing.

Technically the potential we are considering, $V(\{q\}, \{\dot{q}\}, t)$, is not conservative in the standard meaning of the word as the explicit time dependence means that energy need not be conserved. However, the important properties of path independence apply instantaneously and this is what we used above.

8.2.2 Generalisations

There are several ways to generalise the Euler-Lagrange equations. We shall mention a few here.

- We have only considered holonomic constraints so far. We can also consider non-holonomic constraints. These do not decrease the degrees of freedom and instead must be imposed explicitly after solving the equations of motion.
- Not all forces are conservative, for example friction. Given a non-conservative force we can separate $\{Q\}$ into conservative and non-conservative parts. The conservative part acts the same as it does for an entirely conservative force and we are unable to get rid of the non-conservative force through a potential so it remains:

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_j} \right) - \frac{\partial \mathcal{L}}{\partial q_j} = Q_j^{\text{non-conservative}}. \quad (8.2.11)$$

Such non-conservative forces are macroscopic anyway meaning they don't appear in fundamental theories of matter and the Lagrangian formalism is well suited for these applications.

- If there are velocity dependent forces we have two choices. Both involve going back to the more general form of Lagrange's equations. We can always work directly from these but for common velocity dependent forces, such as the Lorentz force of electromagnetism, it can be advantageous to do some of the heavy lifting up front and derive a modified Lagrangian that behaves in the expected way, accounting for this velocity dependence. Indeed, setting up a Lagrangian is often the start point of a field theory.

Nine

Lagrangian Methods in Action

9.1 Single Particle

9.1.1 Cartesian Coordinates

Consider an unconstrained particle in a potential $V(\mathbf{r})$. There are three degrees of freedom. Taking these to be the three Cartesian coordinates we get the Lagrangian

$$\mathcal{L} = T - V = \sum_{i=1}^3 \frac{1}{2} m \dot{x}_i^2 - V(\{x\}). \quad (9.1.1)$$

We then have

$$\frac{\partial \mathcal{L}}{\partial \dot{x}_j} = m \dot{x}_j \quad (9.1.2)$$

$$\frac{\partial \mathcal{L}}{\partial x_j} = -\frac{\partial V}{\partial x_j} \quad (9.1.3)$$

$$\implies \frac{d}{dt}(m \dot{x}_j) = -\frac{dV}{dx_j}. \quad (9.1.4)$$

This is simply Newton's second law with force $\mathbf{F} = -\nabla V$. Notice that if the mass is time dependent we still get the correct result.

9.1.2 Plane Polar Coordinates

Consider a particle constrained to move in a plane with a conservative potential. There are two degrees of freedom, which we take to be r and ϑ , the coordinates in plane polar coordinates. The radial velocity is $v_r = \dot{r}$ and the tangential velocity is $v_\vartheta = r\dot{\vartheta}$. Since these are orthogonal coordinates the velocity is just $v = \sqrt{v_r^2 + v_\vartheta^2}$. Hence the Lagrangian is

$$\mathcal{L} = \frac{1}{2} m (\dot{r}^2 + r^2 \dot{\vartheta}^2) - V(r, \vartheta). \quad (9.1.5)$$

Considering first r we get

$$\frac{\partial \mathcal{L}}{\partial \dot{r}} = m \dot{r}, \quad (9.1.6)$$

$$\frac{\partial \mathcal{L}}{\partial r} = m r \dot{\vartheta}^2 - \frac{\partial V}{\partial r}. \quad (9.1.7)$$

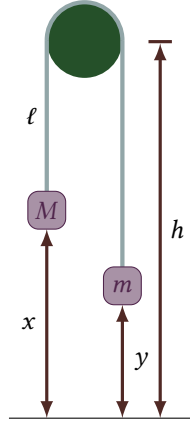


Figure 9.1: Atwood's machine.

Assuming a time independent mass this gives us the radial equation of motion

$$m(\ddot{r} - r\dot{\vartheta}^2) = -\frac{dV}{dr}. \quad (9.1.8)$$

We can identify the left hand side as ma_r where a_r is the radial acceleration.

Now for ϑ we have

$$\frac{\partial \mathcal{L}}{\partial \dot{\vartheta}} = mr^2 \dot{\vartheta} \implies \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{\vartheta}} \right) = 2mr\dot{r}\dot{\vartheta} + mr^2 \ddot{\vartheta}, \quad (9.1.9)$$

$$\frac{\partial \mathcal{L}}{\partial \vartheta} = -\frac{\partial V}{\partial \vartheta}. \quad (9.1.10)$$

This gives the tangential equation of motion

$$mr(2\dot{r}\dot{\vartheta} + mr\ddot{\vartheta}) = -\frac{\partial V}{\partial \vartheta}. \quad (9.1.11)$$

We can identify the left hand side as $mr a_\vartheta$, where a_ϑ is the tangential acceleration.

9.2 Atwood's Machine

Atwood's machine is a smooth pulley with a light inextensible string. Attached to each end of the string is are two masses, m and M . See [Figure 9.1](#).

There are two degrees of freedom, we take them to be x and y , the heights of M and m respectively, above some fixed floor a distance h below the pulley. There is one holonomic constraint:

$$x + y + \ell = 2h, \quad (9.2.1)$$

where ℓ is the length of the string¹. After applying this constraint we have one remaining degree of freedom. We take this to be x . This lets us write

$$y = 2h - x - \ell. \quad (9.2.2)$$

Differentiating this we have

$$\dot{y} = -\dot{x}. \quad (9.2.3)$$

¹we can either assume the radius of the pulley is negligible or we can slightly redefine our coordinates to account for the part of the string that goes over the pulley.

The total kinetic energy is just the kinetic energy of the two masses:

$$T = \frac{1}{2}M\dot{x}^2 + \frac{1}{2}m\dot{y}^2 = \frac{1}{2}M\dot{x}^2 + \frac{1}{2}m(-\dot{x})^2 = \frac{1}{2}(M+m)\dot{x}^2. \quad (9.2.4)$$

The potential energy is gravitational potential energy of the two blocks, we define the floor as where the potential is zero.

$$V = Mgx + mgy = Mgx + mg(2h - x - \ell) = (M - m)gx + \text{constant}. \quad (9.2.5)$$

Since we always differentiate the Lagrangian it is invariant under addition (or removal) of a constant term so we discard the constant. The Lagrangian for the system is then

$$\mathcal{L} = \frac{1}{2}(M - m)\dot{x}^2 - (M - m)gx. \quad (9.2.6)$$

Hence,

$$\frac{\partial \mathcal{L}}{\partial \dot{x}} = (M - m)\dot{x}, \quad (9.2.7)$$

$$\frac{\partial \mathcal{L}}{\partial x} = -(M - m)g. \quad (9.2.8)$$

Hence the equation of motion for the system is

$$(M - m)\ddot{x} = -(M - m)g \implies \ddot{x} = -\frac{M - m}{M + m}g. \quad (9.2.9)$$

Notice that we don't have to consider the tension in the string, which is a real pain with the Newtonian derivation.

9.2.1 Atwood's Monkey

Atwood's monkey is a modification of Atwood's machine where the mass M is replaced with a monkey of mass M which climbs up the rope at a known velocity, $\dot{\varphi}(t)$, which is such that at time t the monkey is a distance $\varphi(t)$ from the end of the rope.

Again we initially have two degrees of freedom, x and y , where now x is the distance of the monkey above the floor and y is still the height of m . We have one holonomic constraint,

$$x + y + \ell - \varphi(t) = 2h. \quad (9.2.10)$$

This allows us to eliminate y :

$$y = 2h - x - \ell + \varphi(t). \quad (9.2.11)$$

Differentiating this we now have

$$\dot{y} = -\dot{x} + \dot{\varphi}(t). \quad (9.2.12)$$

The kinetic energy is then

$$T = \frac{1}{2}M\dot{x}^2 + \frac{1}{2}m\dot{y}^2 \quad (9.2.13)$$

$$= \frac{1}{2}M\dot{x}^2 + \frac{1}{2}m(-\dot{x} + \dot{\varphi})^2 \quad (9.2.14)$$

$$= \frac{1}{2}(M + m)\dot{x}^2 + \frac{1}{2}m\dot{\varphi}^2 - m\dot{x}\dot{\varphi}. \quad (9.2.15)$$

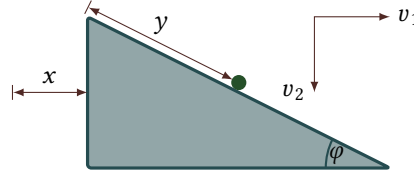


Figure 9.2: A particle on a wedge which is free to move.

The potential energy is

$$V = Mgx + mgy \quad (9.2.16)$$

$$= Mgx + mg(2h - x - \ell + \varphi) \quad (9.2.17)$$

$$= (M - m)gx + mg\varphi + \text{constant}. \quad (9.2.18)$$

The Lagrangian is then

$$\mathcal{L} = \frac{1}{2}(M + m)\dot{x}^2 + \frac{1}{2}m\dot{\varphi}^2 - m\dot{x}\dot{\varphi} - (M - m)gx - mg\varphi. \quad (9.2.19)$$

We then have

$$\frac{\partial \mathcal{L}}{\partial \dot{x}} = (M + m)\dot{x} - m\dot{\varphi} \implies \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}} \right) = (M + m)\ddot{x} - m\ddot{\varphi}, \quad (9.2.20)$$

$$\frac{\partial \mathcal{L}}{\partial x} = -(M - m)g. \quad (9.2.21)$$

The equation of motion for the system is then

$$(M + m)\ddot{x} - m\ddot{\varphi} = -(M - m)g \implies \ddot{x} = -\frac{M - m}{M + m}g + \frac{m}{M + m}\ddot{\varphi}. \quad (9.2.22)$$

We can integrate this to get

$$x(t) = -\frac{1}{2} \frac{M - m}{M + m}gt^2 + \frac{m}{M + m}\varphi(t) + At + B \quad (9.2.23)$$

for constants of integration A and B , which are fixed by the initial conditions.

Notice that the monkey does work on the system and so energy is not conserved.

9.3 Particle and Wedge

Consider a particle on a smooth wedge, which in turn is on a smooth table. What is the acceleration of the wedge? Suppose the particle has mass m , the wedge has mass M , and the wedge angle is φ .

There are two degrees of freedom. We choose them to be the position of the wedge and how far down the wedge the particle has slid. See Figure 9.2. Note that these are *not* orthogonal. We have been asked to find out the acceleration of the wedge and we have chosen its position as one of our generalised coordinates. It is generally a good idea to have desired quantities as generalised coordinates.

The kinetic energy has two contributions, first from the particle

$$\frac{1}{2}M\dot{x}^2, \quad (9.3.1)$$

and from the particle,

$$\frac{1}{2}m(v_1^2 + v_2^2) = \frac{1}{2}m[(\dot{x} + \dot{y} \cos \varphi)^2 + (\dot{y} \sin \varphi)^2] = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + 2\dot{x}\dot{y} \cos \varphi). \quad (9.3.2)$$

The potential energy is provided by the particles gravitational potential energy, so its given by

$$V = -mgy \sin \varphi, \quad (9.3.3)$$

taking the top of the ramp as zero potential energy.

The Lagrangian for the system is

$$\mathcal{L} = \frac{1}{2}M\dot{x}^2 + \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + 2\dot{x}\dot{y} \cos \varphi) + mgy \sin \varphi. \quad (9.3.4)$$

First consider x :

$$\frac{\partial \mathcal{L}}{\partial \dot{x}} = M\dot{x} + m\dot{x} + 2\dot{y} \cos \varphi \implies \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}} \right) = M\ddot{x} + m\ddot{x} + 2\ddot{y} \cos \varphi, \quad (9.3.5)$$

$$\frac{\partial \mathcal{L}}{\partial x} = 0. \quad (9.3.6)$$

Hence

$$(M + m)\ddot{x} + m\ddot{y} \cos \varphi = 0. \quad (9.3.7)$$

For y we get

$$\frac{\partial \mathcal{L}}{\partial \dot{y}} = m\dot{y} + 2\dot{x} \cos \varphi \implies \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{y}} \right) = m\ddot{y} + 2\ddot{x} \cos \varphi, \quad \frac{\partial \mathcal{L}}{\partial y} = mg \sin \varphi, \quad (9.3.8)$$

Hence

$$m\ddot{y} + 2\ddot{x} \cos \varphi = mg \sin \varphi. \quad (9.3.9)$$

We now have two simultaneous equations in \ddot{x} and \ddot{y} . Rearranging the first to give \ddot{y} and substituting into the second we can solve for \ddot{x} and we get

$$\ddot{x} = \frac{g \sin \varphi}{\cos \varphi - (m + M) \sec(\varphi)/m}. \quad (9.3.10)$$

9.4 Pendula

9.4.1 Simple Pendulum

Consider a **simple pendulum** consisting of a mass, m , on the end of a light, rigid rod of length a , fixed at the other end and constrained to move in a plane. Let ϑ be the angle the pendulum makes to the vertical.

The kinetic energy of the system is

$$T = \frac{1}{2}ma^2\dot{\vartheta}^2. \quad (9.4.1)$$

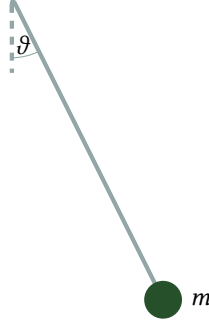


Figure 9.3: Simple pendulum.

The potential energy of the system is the gravitational potential energy of the bob. With a bit of trigonometry we can show that the distance of the pendulum bob below the attachment is $a \cos \vartheta$, and hence the potential energy is

$$V = -mga \cos \vartheta. \quad (9.4.2)$$

The Lagrangian of the system is

$$\mathcal{L} = \frac{1}{2}ma^2\dot{\vartheta}^2 + mga \cos \vartheta. \quad (9.4.3)$$

Hence

$$\frac{\partial \mathcal{L}}{\partial \dot{\vartheta}} = ma^2\dot{\vartheta} \implies \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{\vartheta}} \right) = ma^2\ddot{\vartheta}, \quad (9.4.4)$$

$$\frac{\partial \mathcal{L}}{\partial \vartheta} = -mga \sin \vartheta. \quad (9.4.5)$$

Hence the equation of motion for a pendulum is

$$ma^2\ddot{\vartheta} = -mga \sin \vartheta. \quad (9.4.6)$$

With the small angle approximation we recover the familiar simple harmonic motion equation

$$\ddot{\vartheta} = -\omega^2 \vartheta \quad (9.4.7)$$

with $\omega = \sqrt{g/a}$.

9.4.2 Spherical Pendulum

The **spherical pendulum** is like the simple pendulum but no longer restricted to swing in a plane. Instead we have an extra degree of freedom, which we take to be the angle, φ , about the vertical through which the pendulum swings.

The kinetic energy of the pendulum is

$$T = \frac{1}{2}m(v_{\vartheta}^2 + v_{\varphi}^2) \quad (9.4.8)$$

where $v_{\vartheta} = a\dot{\vartheta}$ and $v_{\varphi} = a\dot{\varphi} \sin \vartheta$, since $a \sin \vartheta$ is the radius of the circle defined by ϑ being constant and varying φ . That is

$$T = \frac{1}{2}ma^2\dot{\varphi}^2 \sin^2 \vartheta + \frac{1}{2}ma^2\dot{\vartheta}^2. \quad (9.4.9)$$

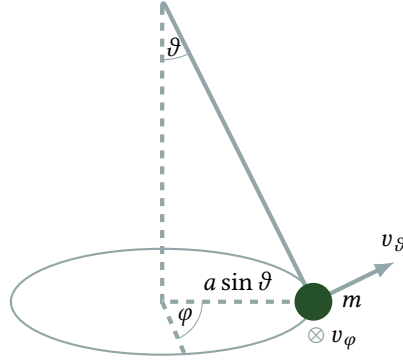


Figure 9.4: Spherical Pendulum.

The potential energy is the same as for the simple pendulum,

$$V = -mga \cos \vartheta. \quad (9.4.10)$$

The Lagrangian for the system is

$$\mathcal{L} = \frac{1}{2}ma^2\dot{\varphi}^2 \sin^2 \vartheta + \frac{1}{2}ma^2\dot{\vartheta}^2 + mga \cos \vartheta. \quad (9.4.11)$$

Considering ϑ we have

$$\frac{\partial \mathcal{L}}{\partial \dot{\vartheta}} = ma^2 \dot{\vartheta} \implies \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{\vartheta}} \right) = ma^2 \ddot{\vartheta}, \quad (9.4.12)$$

$$\frac{\partial \mathcal{L}}{\partial \vartheta} = ma^2 \dot{\varphi}^2 \sin \vartheta \cos \vartheta - mga \sin \vartheta \quad (9.4.13)$$

giving the equation of motion

$$ma^2 \ddot{\vartheta} = ma^2 \dot{\varphi}^2 \sin \vartheta \cos \vartheta - mga \sin \vartheta. \quad (9.4.14)$$

Now considering φ we have

$$\frac{\partial \mathcal{L}}{\partial \dot{\varphi}} = ma^2 \dot{\varphi} \sin^2 \vartheta \implies \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{\varphi}} \right) = ma^2 \ddot{\varphi} \sin^2 \vartheta + 2ma^2 \dot{\varphi} \dot{\vartheta} \sin \vartheta \cos \vartheta \quad (9.4.15)$$

$$\frac{\partial \mathcal{L}}{\partial \varphi} = 0. \quad (9.4.16)$$

Hence the equation of motion for φ is

$$ma^2 \ddot{\varphi} \sin^2 \vartheta + 2ma^2 \dot{\varphi} \dot{\vartheta} \sin \vartheta \cos \vartheta = 0. \quad (9.4.17)$$

We have two coupled differential equations. Integrating [Equation \(9.4.17\)](#) gives

$$ma^2 \dot{\varphi} \sin^2 \vartheta = \text{constant}. \quad (9.4.18)$$

We can identify this constant as L_z , the z component of the angular momentum. This equation implies conservation of angular momentum about the vertical axis.

We can then substitute in $\dot{\phi}$ into Equation (9.4.14) in terms of the angular momentum giving

$$ma^2 \ddot{\vartheta} = ma^2 \sin \vartheta \cos \vartheta \left(\frac{L_z}{a^2 \sin^2 \vartheta} \right)^2 - mga \sin \vartheta. \quad (9.4.19)$$

Multiplying through by $\dot{\vartheta}$ we can integrate both sides with respect to t which gives

$$\frac{1}{2} ma^2 \dot{\vartheta}^2 = mga \cos \vartheta - \frac{L_z^2}{2ma^2 \sin^2 \vartheta} + \text{constant}. \quad (9.4.20)$$

Rearranging we get

$$\frac{1}{2} ma^2 \dot{\vartheta}^2 + \frac{L_z^2}{2ma^2 \sin^2 \vartheta} - mga \cos \vartheta = \text{constant}. \quad (9.4.21)$$

We can identify this constant as the energy, E , in particular the first two terms are the ϑ and ϕ contributions to the kinetic energy and the last term is the potential energy.

9.4.2.1 Notes

Conservation of L_z arises because the Lagrangian is invariant under rotations about the z axis, we can see this because the Lagrangian is independent of ϕ . This gives us a general rule

$$\frac{\partial \mathcal{L}}{\partial q} = 0 \implies \frac{\partial \mathcal{L}}{\partial \dot{q}} = \text{constant}. \quad (9.4.22)$$

We will also show later that time independence of the Lagrangian (or equivalently invariance under time translations) implies conservation of energy. In fact every symmetry (i.e. something that leaves the Lagrangian invariant) implies a conservation law and vice versa. This is **Noether's theorem**, a concept we will return to later.

In this case it may have been easier to give the conservation laws by inspection:

$$L_z = ma^2 \dot{\phi} \sin^2 \vartheta = \text{constant}, \quad (9.4.23)$$

$$E = \frac{1}{2} ma^2 \dot{\vartheta}^2 + \frac{L_z^2}{2ma^2 \sin^2 \vartheta} - mga \cos \vartheta = \text{constant}. \quad (9.4.24)$$

This gives us two first order differential equations which are the two **first integrals** of Lagrange's equations. Each symmetry will give a first integral, which gets us a fair way to solving the system.

If there are as many symmetries as degrees of freedom, as is the case here, then we can often write down a set of first order differential equations like this without having to consider Lagrange's equations at all. We call systems where this is possible **integrable**.

Ten

Calculus of Variations

10.1 Functionals

Definition 10.1.1 — Functional A **functional** is a quantity that depends on the value of a function over some finite interval.

Notation 10.1.2 — Functionals We typically denote functionals with capital letters, for example I . If the functional depends on the value of the function y , which is itself a function of s , then we denote this either $I[y(s)]$ or $I[y]$, if s is obvious, with square brackets to distinguish it from a function.

Often, but not always, functionals are defined as integrals of functions. For example, the following are both functionals

$$I[y(s)] = \int_{s_1}^{s_2} |y(s)|^2 ds, \quad \text{and} \quad L[y(s)] = \int_{s_1}^{s_2} \sqrt{1 + \left(\frac{dy}{dx}\right)^2} ds. \quad (10.1.3)$$

The following are *not* functionals, they are merely functions

$$\sin(y(s)), \quad \text{and} \quad y'(s) = \frac{dy}{ds}. \quad (10.1.4)$$

Notice that the value of the derivative does depend on the value of y at both s and $s + ds$, but for infinitesimal ds and so the interval over which y' depends on y is not finite.

10.2 Euler's Equation

Consider a functional of the form

$$I[y(s)] = \int_a^b F(y(s), y'(s), s) ds \quad (10.2.1)$$

where F is a function of the variables y , y' , and s . We treat these as three independent variables.

The **calculus of variations** asks the question ‘what path, y , maximises or minimises $I[y(s)]$ subject to fixed endpoints $y(a)$ and $y(b)$?’

To answer this question we consider a small change in the path,

$$y(s) \rightarrow y(s) + \eta(s), \quad (10.2.2)$$

where $\eta(s)$ is small and vanishes at the endpoints, $\eta(a) = \eta(b) = 0$, so that the endpoints remain fixed. We then answer this question by finding out when the change in I is zero, that is when

$$\delta I = I[y(s) + \eta(s)] - I[y(s)] \stackrel{?}{=} 0, \quad (10.2.3)$$

since, as with functions, this corresponds to when the derivative is zero and hence the functional achieves some extreme value.

To find this in the most general case consider

$$I[y(s) + \eta(s)] = \int_a^b F(y(s) + \eta(s), y'(s) + \eta'(s), s) ds. \quad (10.2.4)$$

We can Taylor expand this for small η giving

$$I[y(s) + \eta(s)] \approx \int_a^b \left[F(y(s), y'(s), s) + \eta(s) \frac{\partial F}{\partial y} + \eta'(s) \frac{\partial F}{\partial y'} \right] ds. \quad (10.2.5)$$

Identifying the first term as $I[y(s)]$ we see that we have

$$\delta I = \int_a^b \left[\eta(s) \frac{\partial F}{\partial y} + \eta'(s) \frac{\partial F}{\partial y'} \right] ds. \quad (10.2.6)$$

Consider the second term,

$$\int_a^b \eta'(s) \frac{\partial F}{\partial y'} ds. \quad (10.2.7)$$

Integrating by parts we get

$$\left[\eta(s) \frac{\partial F}{\partial y'} \right]_a^b - \int_a^b \eta(s) \frac{d}{ds} \left(\frac{\partial F}{\partial y'} \right) ds = - \int_a^b \eta(s) \frac{d}{ds} \left(\frac{\partial F}{\partial y'} \right) ds. \quad (10.2.8)$$

The first term vanishes since $\eta(a) = \eta(b) = 0$. We then have

$$\delta I = \int_a^b \left[\frac{\partial F}{\partial y} - \frac{d}{ds} \left(\frac{\partial F}{\partial y'} \right) \right] \eta(s) ds. \quad (10.2.9)$$

We are looking for δI to vanish for arbitrary small changes, η , in the path, and therefore conclude we must have

$$\frac{\partial F}{\partial y} - \frac{d}{ds} \left(\frac{\partial F}{\partial y'} \right) = 0. \quad (10.2.10)$$

This is **Euler's equation**.

Notice that if we identify $F = \mathcal{L}$, $y = q$, and $s = t$ we get Lagrange's equation. For this reason these two equations are often referred to as the **Euler-Lagrange equation**. We will come back to what this means later.

10.3 First Integrals

Suppose that F doesn't depend on y . From Euler's equation we have

$$\frac{\partial F}{\partial y} = 0 \implies \frac{d}{ds} \left(\frac{\partial F}{\partial y'} \right) = 0 \implies \frac{\partial F}{\partial y'} = \text{constant}. \quad (10.3.1)$$

This can often be used to more easily solve Euler's equation. A similar, but less obvious, result is in the next lemma.

Lemma 10.3.2 If F doesn't depend explicitly on s , so that

$$\frac{\partial F}{\partial s} = 0, \quad (10.3.3)$$

then

$$y' \frac{\partial F}{\partial y'} = F = \text{constant}. \quad (10.3.4)$$

Proof. Consider the total derivative

$$\frac{dF}{ds} = \frac{\partial F}{\partial y} y' + \frac{\partial F}{\partial y'} y'' + \frac{\partial F}{\partial s} = \frac{\partial F}{\partial y} y' + \frac{\partial F}{\partial y'} y''. \quad (10.3.5)$$

Where the final derivative vanishes by the hypothesis. Using Euler's equation we can substitute for the first derivative giving

$$\frac{dF}{ds} = \left[\frac{d}{ds} \left(\frac{\partial F}{\partial y'} \right) \right] y' + \frac{\partial F}{\partial y'} y''. \quad (10.3.6)$$

Notice then that by the chain rule we have

$$\frac{d}{ds} \left(y' \frac{\partial F}{\partial y'} \right) = \left[\frac{d}{ds} \left(\frac{\partial F}{\partial y'} \right) \right] y' + \frac{\partial F}{\partial y'} y'' \quad (10.3.7)$$

and so

$$\frac{dF}{ds} = \frac{d}{ds} \left(y' \frac{\partial F}{\partial y'} \right). \quad (10.3.8)$$

It then follows that

$$\frac{d}{ds} \left(y' \frac{\partial F}{\partial y'} \right) - \frac{dF}{ds} = \frac{d}{ds} \left(y' \frac{\partial F}{\partial y'} - F \right) = 0. \quad (10.3.9)$$

From this we get the desired

$$y' \frac{\partial F}{\partial y'} - F = \text{constant}. \quad (10.3.10)$$

□

10.4 Examples

In this section we solve a few classic calculus of variation problems.

10.4.1 Shortest Path

Given two points (x_1, y_1) and (x_2, y_2) in the plane what is the shortest path between them?

Suppose we have a path described by $y(x)$, we are assuming here that the path doesn't double back on itself, which is reasonable for the shortest path, and therefore can be a function of x . If this were not the case we would need to parametrise y and x in terms of some third variable.

Consider a small step along the x -direction, dx . If we are to stay on the path we will need to make a small step in the y -direction, dy . The resulting step length along the path is

$$ds = \sqrt{dx^2 + dy^2} = dx \sqrt{1 + \left(\frac{dy}{dx}\right)^2}. \quad (10.4.1)$$

The length of a path is then described by the functional

$$L[y(s)] = \int_{x_1}^{x_2} ds = \int_{x_1}^{x_2} \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx = \int_{x_1}^{x_2} \sqrt{1 + (y'(x))^2} dx. \quad (10.4.2)$$

This is of the required form to apply Euler's equation with

$$F(y(x), y'(x), x) = \sqrt{1 + (y'(x))^2}. \quad (10.4.3)$$

This is independent of y and so we have

$$\frac{\partial F}{\partial y'} = \text{constant}. \quad (10.4.4)$$

That is

$$\frac{\partial F}{\partial y'} = \frac{y'}{\sqrt{1 + y'^2}} = p \quad (10.4.5)$$

for some constant, p . Inverting this we have

$$y' = \frac{p}{\sqrt{1 - p^2}} = m \quad (10.4.6)$$

for some constant m . Hence $y = mx + c$, that is the shortest line between two points is a straight line. With the initial condition $y(x_1) = y_1$ and $y(x_2) = y_2$ we get $y_1 = mx_1 + c$ and $y_2 = mx_2 + c$. Subtracting one from the other we have

$$y_1 - y_2 = m(x_1 - x_2) \implies m = \frac{y_1 - y_2}{x_1 - x_2}, \quad (10.4.7)$$

which is simply the gradient of the line as expected. Similarly by considering $y(0) = c$ we see that c is the y -intercept.

Here we used the Euclidean metric, $s^2 = \sum_{i,j} \delta_{ij} x_j x_i$, if instead we used a more general metric, $s^2 = \sum_{i,j} g_{ij} x_j x_i$, then we would get the general equations for a geodesic, which is just the fancy word for the shortest path in a given geometry.

10.4.2 Brachistochrone

A smooth wire connects two points, (x_1, y_1) and (x_2, y_2) , which are in a plane with gravity acting downwards in the negative y direction. Find the path such that a bead sliding down the wire from (x_1, y_1) , initially at rest, reaches (x_2, y_2) , in the least possible time.

R This problem was the original reason Euler invented variational calculus. Brachistochrone comes from the Greek $\beta\rho\acute{\alpha}\chi\iota\sigma\tau\omicron\varsigma$ (*brákhistos*) for *shortest* and $\chi\rho\acute{o}\nu\omicron\varsigma$ (*khronos*) for *time*.

Again, we assume the path doesn't double back on itself so we look for a function $y(x)$ which minimises the transit time, $\tau[y(x)]$. In order to find the correct functional, τ , we consider the velocity at some point, x , along the path. At (x_1, y_1) the velocity is 0, since the particle is at rest. At (x, y) the velocity is $v(x)$. We can then apply conservation of energy and we find that

$$0 + mgy_1 = \frac{1}{2}mv^2 + mgy(x) \implies v = \sqrt{2g(y_1 - y(x))}. \quad (10.4.8)$$

The time, $d\tau$, taken to fall along a section of the path of length ds is $d\tau = ds/v$. Again we have $ds = \sqrt{1 + y'^2}dx$. Putting this together we have

$$\tau[y] = \int_{x_1}^{x_2} d\tau = \int_{x_1}^{x_2} \frac{1}{v} ds = \int_{x_1}^{x_2} \sqrt{\frac{1 + y'^2}{2g(y_1 - y)}} dx. \quad (10.4.9)$$

This is of the form required to apply Euler's equation with

$$F(y(x), y'(x), x) = \sqrt{\frac{1 + y'^2}{2g(y_1 - y)}}. \quad (10.4.10)$$

This has no explicit x dependence and hence

$$y' \frac{\partial F}{\partial y'} - F = \text{constant}. \quad (10.4.11)$$

We have

$$\frac{\partial F}{\partial y'} = \frac{y'}{2g(y_1 - y)} \sqrt{\frac{1 + y'^2}{2g(y_1 - y)}} = \frac{Fy'}{1 + y'^2} \implies y' \frac{\partial F}{\partial y'} = \frac{Fy'^2}{1 + y'^2}. \quad (10.4.12)$$

From this we have

$$y' \frac{\partial F}{\partial y'} - F = \frac{Fy'^2}{1 + y'^2} - F = -\frac{F}{1 + y'^2} = \text{constant}. \quad (10.4.13)$$

Substituting in F and rearranging we get

$$\frac{1}{\sqrt{2g(y_1 - y)(1 + y'^2)}} = \text{constant} \implies (y_1 - y)(1 + y'^2) = \text{constant}. \quad (10.4.14)$$

This is the answer as far as variational calculus is concerned, a differential equation that can be solved for the path. Euler, being Euler, recognised it as the equation of a cycloid, which is given parametrically by

$$x = a(2\psi - \sin(2\psi)) + x_1, \quad \text{and} \quad y = a(\cos(2\psi) - 1) + y_1 \quad (10.4.15)$$

where ψ is the angle the curve makes to the the upward vertical. The cycloid starts vertical, which gives a large initial acceleration, and then levels out to give a relatively short path once the ring is going quickly. It should be noted that a cycloid is the path traced by a point on a circle rolling along the x -axis.

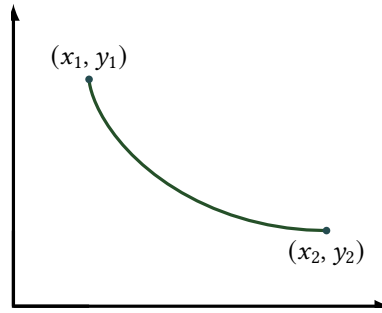


Figure 10.1: The Brachistochrone

10.4.3 Soap Film

... minimise its surface area,
something something, energy!

Matt Parker on why the bubble is
shaped to minimise its surface area
[1].

Sadly, that's super difficult

Matt Parker, on doing this
calculation [1].

Two wire rings of radius a are held at separation $2d$. What axially symmetric shape, $\rho(z)$, minimises the surface area, A , of a soap film connecting them?

R Minimising the surface area gives the shape of the soap film since the surface tension acts to minimise surface area. We assume that the shape is axially symmetric and so it is only the distance from the axis at some point along the axis that we need to know.

Consider taking a small step, dz , along the axis of symmetry. In order to stay on the soap film a small step, $d\rho$, must be taken along the soap film. The resulting step along the soap film is

$$ds = \sqrt{dz^2 + d\rho^2} = dz\sqrt{1 + \rho'(z)^2}. \quad (10.4.16)$$

Multiplying by 2π gives the area corresponding to this range of z values and so integrating we have

$$A[\rho(z)] = \int_{z=-d}^{z=d} 2\pi\rho(z) ds = 2\pi \int_{-d}^d \rho\sqrt{1 + \rho'^2} dz. \quad (10.4.17)$$

This is of form required to apply Euler's equation with

$$F(\rho(z), \rho'(z), z) = \rho\sqrt{1 + \rho'^2}. \quad (10.4.18)$$

This has no explicit z dependence and so

$$\rho' \frac{\partial F}{\partial \rho'} - F = \text{constant} = C. \quad (10.4.19)$$

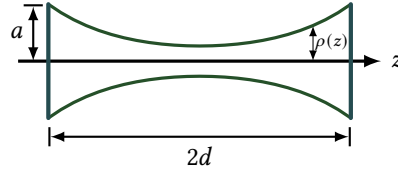


Figure 10.2: Soap film between two rings.

From this we get

$$\rho' \frac{\rho \rho'}{\sqrt{1 + \rho'^2}} - \rho \sqrt{1 + \rho'^2} = C. \quad (10.4.20)$$

This can be simplified to

$$\rho = -C \sqrt{1 + \rho'^2} \implies \frac{d\rho}{dz} = \sqrt{\left(\frac{\rho(z)}{C}\right)^2 - 1}. \quad (10.4.21)$$

This is the solution as far as calculus of variations is concerned. This can be solved by separation of variables to find

$$\rho(z) = C \cosh\left(\frac{z - z_0}{C}\right) \quad (10.4.22)$$

where z_0 is a constant of integration. Including the boundary condition that $\rho(\pm d) = a$ we get $z_0 = 0$ and hence

$$\frac{a}{C} = \cosh\left(\frac{d}{C}\right). \quad (10.4.23)$$

This cannot be solved analytically and numerically we find it has answers only if $d/a < 0.663 \dots$. We interpret this as meaning that if the rings are moved too far apart then the soap film pops, and instead fills the rings in flat discs, although the mathematics would suggest an infinitely thin strand connecting the two discs along the axis, but this doesn't happen in reality.

10.5 Calculus of Variations for Many Variables

Now consider a functional of many paths given by

$$I[\{y(s)\}] = \int_a^b F(\{y\}, \{y'\}, s) ds. \quad (10.5.1)$$

With $\{y\} = y_1, y_2, \dots, y_p$ being a set of p linearly independent functions of the variable s .

Consider a small variation $y_i \rightarrow y_i + \eta_i$, where η_i vanishes at the endpoints. We then have

$$F(\{y + \eta\}, \{y' + \eta'\}, s) \approx F(\{y\}, \{y'\}, s) + \sum_{i=1}^p \frac{\partial F}{\partial y_i} \eta_i(s) + \sum_{i=1}^p \frac{\partial F}{\partial y'_i} \eta'_i(s). \quad (10.5.2)$$

Requiring that $\delta I = 0$ and following the same derivation as the single variable case in [Section 10.2](#) we get a set of p equations, one for each y_i , which are

$$\frac{d}{ds} \left(\frac{\partial F}{\partial y'_j} \right) - \frac{\partial F}{\partial y_j} = 0. \quad (10.5.3)$$

We can similarly derive first integrals the obvious one being if for some specific i F doesn't depend on y_i then we have

$$\frac{\partial F}{\partial y'_i} = \text{constant}. \quad (10.5.4)$$

There is one of these equations for each variable, y_i , upon which F does not depend. If F doesn't depend on s then

$$\sum_i y'_i \frac{\partial F}{\partial y'_i} - F = \text{constant}. \quad (10.5.5)$$

There is only one equation here which includes all y_i .

10.6 Hamilton's Principle

Earlier we commented that Lagrange's equations are just the Euler–Lagrange equations with $F = \mathcal{L}$, $y_i = q_i$, and $s = t$. This implies that we can view Lagrangian mechanics as the result of demanding that the following functional is stationary:

$$S[\{q(t)\}] := \int_{t_1}^{t_2} \mathcal{L}(\{q\}, \{\dot{q}\}, t) dt. \quad (10.6.1)$$

We call this the **action**. **Hamilton's principle** states that the action is stationary, that is

$$\delta S = 0. \quad (10.6.2)$$

Note that Hamilton's principle as stated here involves choosing two times, t_1 and t_2 , and requiring that $\delta S = 0$ while keeping $\{q(t_1)\}$ and $\{q(t_2)\}$ fixed. Our preferred boundary conditions would involve one time, t_1 , and fixing the initial positions, $\{q(t_1)\}$, and velocities, $\{\dot{q}_i(t_1)\}$. Fortunately we can start from Hamilton's principle, derive the Lagrange equations, and then solve them for any boundary conditions we wish, so this isn't a problem.

The most general form of Hamilton's principle states that

$$\delta \int \mathcal{L} dt = 0, \quad (10.6.3)$$

which is coordinate independent, and hence means that it doesn't mean what choice we make for our generalised coordinates.

10.6.1 Hamilton's Principle from Newtonian Mechanics

Consider a system of N particles in three dimensions, so we have $3N$ degrees of freedom. The kinetic energy in Cartesian coordinates is

$$T = \sum_i \frac{1}{2} m_i \dot{x}_i^2. \quad (10.6.4)$$

Newton's second law can then be written as

$$\sum_i m_i \ddot{x}_i = \frac{d}{dt} \left(\frac{\partial T}{\partial \dot{x}_i} \right) = - \frac{\partial V}{\partial x_i} \quad (10.6.5)$$

assuming a conservative force, and hence a potential, $V = V(\{x\}, \{\dot{x}\}, t)$. We can therefore write

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{x}_i} \right) = - \frac{\partial V}{\partial x_i} + \frac{\partial T}{\partial x_i} + \frac{d}{dt} \left(\frac{\partial V}{\partial \dot{x}_i} \right) \quad (10.6.6)$$

where the two extra terms are both zero. Defining the Lagrangian to be $\mathcal{L} := T - V$ and rearranging this we have

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}_i} \right) - \frac{\partial \mathcal{L}}{\partial x_i} = 0. \quad (10.6.7)$$

We can then construct the action

$$S = \int \mathcal{L}(\{x\}, \{\dot{x}\}, t) dt, \quad (10.6.8)$$

and then Hamilton's principle gives us $\delta S = 0$ in Cartesian coordinates. However, Hamilton's principle is coordinate independent and so $\delta S = 0$ holds for any choice of generalised coordinates, $\{\tilde{q}\}$. We have not yet applied any constraints and so there must be $3N$ coordinates at this stage.

Suppose we have M holonomic constraints, which can be expressed in Cartesian coordinates as a set of M algebraic equalities

$$f_\alpha(\{x\}, t) = 0 \quad (10.6.9)$$

where $\alpha = 1, \dots, M$. We can then specify our choice of $3N$ coordinates, $\{\tilde{q}\} = (\tilde{q}_1, \tilde{q}_2, \dots, \tilde{q}_{3N})$, as we wish. We make the following choice:

$$\{\tilde{q}\} = (q_1, q_2, \dots, q_{3N-M}, f_1, f_2, \dots, f_M) = \{q, f\}. \quad (10.6.10)$$

In other words we choose M of our constraints to be precisely the values held constant by our constraints. We are therefore left with $3N - M$ generalised coordinates, $\{q\}$, which are linearly independent of each other and of the constraints.

Hamilton's principle states that for unconstrained motion we have $\delta S = 0$ and hence

$$\delta \int \mathcal{L}(\{q, f\}, \{\dot{q}, \dot{f}\}, t) dt = 0. \quad (10.6.11)$$

For constrained motion we demand that $f_\alpha = 0$, and hence $\dot{f}_\alpha = 0$, for all α . It then follows that if we define $\mathcal{L}(\{q\}, \{\dot{q}\}, t) := \mathcal{L}(\{q, 0\}, \{\dot{q}, 0\}, t)$ we have

$$\delta \int \mathcal{L}(\{q\}, \{\dot{q}\}, t) dt. \quad (10.6.12)$$

Hence Hamilton's principle applies to constrained motion in arbitrary generalised coordinates. From this we can derive the Euler-Lagrange equations

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) - \frac{\partial \mathcal{L}}{\partial q_i} = 0. \quad (10.6.13)$$

So we have derived the key equations for Lagrangian mechanics in a new way.

This will be useful when we come to consider modifications, such as allowing velocity dependent potentials, where we will use Hamilton's principle as a starting point.

Eleven

Symmetries and Conservation Laws

In this section we will study the Lagrangian, and Euler–Lagrange equations, and look for symmetries which leave the Lagrangian, and equations of motion, invariant.

11.1 Invariance Properties of the Lagrangian

11.1.1 Invariance Under Change of Generalised Coordinates

Consider two sets of generalised coordinates, $\{q\}$ and $\{q'\}$, for the same system. Let $\mathcal{L} = \mathcal{L}(\{q\}, \{\dot{q}\}, t)$ be the Lagrangian in the first set of coordinates and $\mathcal{L}' = \mathcal{L}'(\{q'\}, \{\dot{q}'\}, t)$ the Lagrangian in the second set of coordinates. Numerically for the same state we have $\mathcal{L} = \mathcal{L}'$, since both are equal to $T - V$, which is a physical property independent of the coordinates we choose to define it. However, the equations of motion depend on the *form* of the Lagrangian, not just its numerical value, and so we cannot just assume that being numerically equal means the same equations of motion will be produced. We need to prove that the two Lagrangians are **equivalent**, by which we mean they lead to the same equations of motion.

Theorem 11.1.1 — Invariance of the Lagrangian Under Change of Generalised Coordinates. Consider two sets of generalised coordinates, $\{q\}$ and $\{q'\}$, describing the same system. Then the Lagrangians $\mathcal{L}(\{q\}, \{\dot{q}\}, t)$ and $\mathcal{L}'(\{q'\}, \{\dot{q}'\}, t)$ are equivalent.

Proof. We start with \mathcal{L} , for which we know that the Euler–Lagrange equations hold:

$$0 = \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_j} \right) - \frac{\partial \mathcal{L}}{\partial q_j}. \quad (11.1.2)$$

If we can derive the Euler–Lagrange equations but for \mathcal{L}' , q' , and \dot{q}' , from this then the two Lagrangians are equivalent.

For the first step notice that since \mathcal{L} and \mathcal{L}' are numerically equal we can replace \mathcal{L} with \mathcal{L}' giving

$$0 = \frac{d}{dt} \left(\frac{\partial \mathcal{L}'}{\partial \dot{q}_j} \right) - \frac{\partial \mathcal{L}'}{\partial q_j}. \quad (11.1.3)$$

Applying the chain rule we get

$$0 = \frac{d}{dt} \left[\sum_i \frac{\partial \mathcal{L}'}{\partial \dot{q}_i'} \frac{\partial q_i'}{\partial \dot{q}_j} + \underbrace{\sum_i \frac{\partial \mathcal{L}'}{\partial \dot{q}_i'} \frac{\partial \dot{q}_i'}{\partial \dot{q}_j}}_{=0} \right] - \sum_i \frac{\partial \mathcal{L}'}{\partial q_i'} \frac{\partial q_i'}{\partial q_j} - \sum_i \frac{\partial \mathcal{L}'}{\partial \dot{q}_i'} \frac{\partial \dot{q}_i'}{\partial q_j}. \quad (11.1.4)$$

Noticing that we can write $q_i' = q_i'(\{q\}, \{\dot{q}\}, t)$ the first term vanishes. The time derivative then distributes over the second term and applying the chain rule the second term becomes

$$\left[\frac{d}{dt} \frac{\partial \mathcal{L}'}{\partial \dot{q}_i'} \right] \frac{\partial q_i'}{\partial \dot{q}_j} + \frac{\partial \mathcal{L}'}{\partial \dot{q}_i'} \left[\frac{d}{dt} \frac{\partial \dot{q}_i'}{\partial \dot{q}_j} \right]. \quad (11.1.5)$$

Applying the cancellation of dots rule we have $\partial \dot{q}_i' / \partial \dot{q}_j = \partial q_i' / \partial q_j$. We can also apply the commuting derivatives rule to the second term after the cancellation of dots rule giving $\frac{\partial \dot{q}_i'}{\partial q_j}$. Then this whole term becomes

$$\left[\frac{d}{dt} \frac{\partial \mathcal{L}'}{\partial \dot{q}_i'} \right] \frac{\partial q_i'}{\partial q_j} + \frac{\partial \mathcal{L}'}{\partial \dot{q}_i'} \frac{\partial \dot{q}_i'}{\partial q_j}. \quad (11.1.6)$$

Substituting this back into Equation (11.1.4) we have

$$0 = \sum_i \left(\left[\frac{d}{dt} \frac{\partial \mathcal{L}'}{\partial \dot{q}_i'} \right] \frac{\partial q_i'}{\partial q_j} + \frac{\partial \mathcal{L}'}{\partial \dot{q}_i'} \frac{\partial \dot{q}_i'}{\partial q_j} - \frac{\partial \mathcal{L}'}{\partial q_i'} \frac{\partial q_i'}{\partial q_j} - \frac{\partial \mathcal{L}'}{\partial \dot{q}_i'} \frac{\partial \dot{q}_i'}{\partial q_j} \right). \quad (11.1.7)$$

Noticing that the second and third term cancel and factoring out the common $\frac{\partial q_i'}{\partial q_j}$ we get

$$0 = \sum_i \left[\frac{d}{dt} \left(\frac{\partial \mathcal{L}'}{\partial \dot{q}_i'} \right) - \frac{\partial \mathcal{L}'}{\partial q_i'} \right] \frac{\partial q_i'}{\partial q_j}. \quad (11.1.8)$$

Since we can vary q_i' individually and this must always hold we must have that for every i

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}'}{\partial \dot{q}_i'} \right) - \frac{\partial \mathcal{L}'}{\partial q_i'} = 0. \quad (11.1.9)$$

These are exactly the Euler–Lagrange equations for \mathcal{L}' in terms of $\{q\}$ and $\{\dot{q}\}$. This proves the two Lagrangians are equivalent. \square

11.1.2 Invariance Under Addition of a Total Derivative

Theorem 11.1.10 — Invariance of the Lagrangian Under Addition of a Total Derivative. Let $F = F(\{q\}, \{\dot{q}\}, t)$ be a function of the generalised position and time, but not the generalised velocities. Then the Lagrangians \mathcal{L} and $\mathcal{L}' = \mathcal{L} + dF/dt$ are equivalent.

Proof. Consider the Euler–Lagrange equations for \mathcal{L}' :

$$0 = \frac{d}{dt} \left(\frac{\partial \mathcal{L}'}{\partial \dot{q}} \right) - \frac{\partial \mathcal{L}'}{\partial q}. \quad (11.1.11)$$

Substituting in for the definition of \mathcal{L}' and using the linearity of the derivative we have

$$0 = \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}} \right) + \frac{d}{dt} \left(\frac{\partial \dot{F}}{\partial \dot{q}} \right) - \frac{\partial \mathcal{L}}{\partial q} - \frac{\partial \dot{F}}{\partial q}. \quad (11.1.12)$$

Notice that since F is a function of the positions \dot{F} may be a function of the velocities. This means we may have $\frac{\partial \dot{F}}{\partial \dot{q}} \neq 0$. We can simplify this by taking the second term, applying the cancellation of dots rule and then the commuting derivatives rule which gives

$$0 = \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}} \right) + \frac{\partial \dot{F}}{\partial q} - \frac{\partial \mathcal{L}}{\partial q} - \frac{\partial \dot{F}}{\partial q}. \quad (11.1.13)$$

Noticing that the second and third terms cancel we are left with

$$0 = \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}} \right) - \frac{\partial \mathcal{L}}{\partial q}. \quad (11.1.14)$$

This is the Euler–Lagrange equation for \mathcal{L} , q , and \dot{q} , and so \mathcal{L} and $\mathcal{L}' = \mathcal{L} + \dot{F}$ are equivalent. \square

11.2 Generalised Momenta

Definition 11.2.1 — Generalised Momenta The **generalised momentum**, or **canonical momentum** conjugate to the generalised coordinate q is

$$p := \frac{\partial \mathcal{L}}{\partial \dot{q}}. \quad (11.2.2)$$

Using this definition we can write Lagrange's equation as

$$\frac{dp}{dt} = \frac{\partial \mathcal{L}}{\partial q}. \quad (11.2.3)$$

■ **Example 11.2.4** Consider a single particle in Cartesian coordinates in a velocity independent potential. We have

$$\mathcal{L} = \frac{1}{2} m \sum_i \dot{x}_i^2 - V(\{x\}, \{x\}, t) \quad (11.2.5)$$

and so

$$p_i = m\dot{x}_i \quad (11.2.6)$$

which means that the conjugate momentum of x_i is just the component of the linear momentum in the x_i direction.

■ **Example 11.2.7** Consider a rotating object. The kinetic energy will be of the form

$$T = T_{\text{linear}} + \frac{1}{2}I\dot{\vartheta}^2 \quad (11.2.8)$$

where the linear component of the kinetic energy, T_{linear} , is independent of $\dot{\vartheta}$, and I is the moment of inertia. For a velocity independent potential we then have

$$p_{\vartheta} = I\dot{\vartheta} \quad (11.2.9)$$

which is the angular momentum.

These two examples justify calling $\partial\mathcal{L}/\partial q$ the generalised *momentum*, however, it should be noted that $\partial\mathcal{L}/\partial q$ cannot always be identified as a momentum.

Notice that if the Lagrangian doesn't depend on a specific coordinate, q_j , then $\partial\mathcal{L}/\partial q_j = 0$ and so

$$\frac{dp_j}{dt} = 0 \implies p_j = \text{constant}. \quad (11.2.10)$$

So p_j is conserved. In this case we call q_j a **cyclic coordinate** or **ignorable coordinate**. We call the conserved momentum a **constant of motion**.

We can view $\partial\mathcal{L}/\partial q_j$ as a symmetry of the Lagrangian meaning that we are free to change q_j by some constant amount and without changing the resulting physics. For example conservation of linear momentum in the x_i direction corresponds to the ability to make the translation $x_i \rightarrow x_i + \Delta x$ without changing anything, and conservation of angular momentum when rotating about \mathbf{e}_{ϑ} corresponds to the ability to make the rotation $\vartheta \rightarrow \vartheta + \Delta\vartheta$ without changing anything. This leads to Noether's theorem, which informally states that every symmetry of the Lagrangian leads to a conserved quantity and vice versa. A more formal statement relates invariance of the Lagrangian under a given Lie group to a conserved current.

11.3 The Energy Function

Definition 11.3.1 — Energy Function We define the **energy function** as

$$h(\{q\}, \{\dot{q}\}, t) = \sum_j \frac{\partial\mathcal{L}}{\partial\dot{q}_j} \dot{q}_j - \mathcal{L}(\{q\}, \{\dot{q}\}, t) = \sum_j p_j \dot{q}_j - \mathcal{L}. \quad (11.3.2)$$

If the Lagrangian has no explicit time dependence, so $\mathcal{L} = \mathcal{L}(\{q\}, \{\dot{q}\}, t)$, then we have a first integral similar to the one in Equation (10.5.5), giving

$$h = \sum_j \dot{q}_j \frac{\partial \mathcal{L}}{\partial \dot{q}_j} - \mathcal{L} = \text{constant}. \quad (11.3.3)$$

This means that h is a conserved quantity. In terms of Noether's theorem this corresponds to invariance of the Lagrangian under time translation.

Numerically h has the same value as the Hamiltonian, \mathcal{H} , but h is a function of the position, velocity, and time, whereas the Hamiltonian is a function of position, momentum, and time.

11.3.1 Energy Function and the Total Energy

As the name suggests the energy function is a generalisation of the total energy, $E = T + V$. However, there are times when $h \neq E$.

It turns out that a sufficient condition for $h = E$ is that the kinetic energy be quadratic in velocity and the potential velocity independent. By this we mean we can write the kinetic energy as

$$T(\{q\}, \{\dot{q}\}, t) = \frac{1}{2} \sum_i \sum_j T_{ij}(\{q\}, t) \dot{q}_i \dot{q}_j. \quad (11.3.4)$$

Here T_{ij} is a symmetric matrix whose entries are a function of position and time, but not velocity.

An obvious example of a quadratic energy is the a set of particles in Cartesian coordinates in which

$$T = \sum_i \frac{1}{2} m_i \dot{x}_i^2 \implies T_{ij} = m_i \delta_{ij}. \quad (11.3.5)$$

If we have a kinetic energy that is quadratic in the velocities and a potential that is independent of the velocities then

$$p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i} = \frac{\partial T}{\partial \dot{q}_i} = \sum_j T_{ij} \dot{q}_j \quad (11.3.6)$$

and so it follows that

$$h = \sum_i p_i \dot{q}_i - \mathcal{L} \quad (11.3.7)$$

$$= \sum_i \sum_j T_{ij} \dot{q}_j \dot{q}_i - T + V \quad (11.3.8)$$

$$= 2T - T + V \quad (11.3.9)$$

$$= T + V \quad (11.3.10)$$

$$= E. \quad (11.3.11)$$

So we see that in systems where the kinetic energy is quadratic in the velocities (and the potential velocity independent) the energy is conserved. We will now study two similar examples, one where the energy is quadratic in the velocity and one where it isn't.

■ **Example 11.3.12 — Particle on a Rod** A particle slides on a massless rod which is fixed to a pivot at one end. A potential, $V(r)$, is applied to the particle and has rotational symmetry about the pivot. Find the two constants of motion. Since the rod can rotate and the particle can slide along it the particle can reach any point in the plane. The sensible choice of coordinates is polar coordinates. The Lagrangian is then

$$\mathcal{L} = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\vartheta}^2) - V(r). \quad (11.3.13)$$

The generalised momenta are then

$$p_r = \frac{\partial \mathcal{L}}{\partial \dot{r}} = m\dot{r}, \quad \text{and} \quad p_{\vartheta} = \frac{\partial \mathcal{L}}{\partial \dot{\vartheta}} = mr^2\dot{\vartheta}. \quad (11.3.14)$$

These are exactly the radial component of linear momentum and the angular momentum about the rotation axis.

The Lagrangian is independent of ϑ and therefore p_{ϑ} is conserved.

The Lagrangian is independent of time and therefore

$$h = p_r\dot{r} + p_{\vartheta}\dot{\vartheta} - \mathcal{L} \quad (11.3.15)$$

is conserved. Since T is quadratic in the velocities this quantity is the total energy.

■ **Example 11.3.16 — Particle on a Rod in Forced Motion** Consider the same set up as the previous example, but now the rod is forced to rotate at a fixed angular velocity $\dot{\vartheta} = \omega$. Find a constant of the motion.

There is now one degree of freedom, r , and the Lagrangian is

$$\mathcal{L} = \frac{1}{2}m(\dot{r}^2 + r^2\omega^2) - V(r). \quad (11.3.17)$$

Since ω is a constant the kinetic energy is *not* quadratic in the velocity.

The canonical momentum is

$$p_r = \frac{\partial \mathcal{L}}{\partial \dot{r}} = m\dot{r}. \quad (11.3.18)$$

However, now

$$\frac{\partial \mathcal{L}}{\partial r} = m\omega^2 r - \frac{\partial V}{\partial r} \neq 0 \quad (11.3.19)$$

and so the canonical momentum is not conserved^a, even if the potential is zero. Notice that the first term here is the centrifugal force due to the forced rotation.

The Lagrangian has no explicit time dependence and so

$$h = p_r\dot{r} - \mathcal{L} = \frac{1}{2}m(\dot{r}^2 - r^2\omega^2) + V(r) \quad (11.3.20)$$

is conserved. This is *not* the total energy, which is not conserved in this case. The total energy is

$$E = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\omega}^2) + V(r). \quad (11.3.21)$$

The tangential acceleration is

$$a_{\vartheta} = r\ddot{\vartheta} + 2\dot{r}\dot{\vartheta} \quad (11.3.22)$$

which corresponds to the constraint force

$$F = ma_{\vartheta} = 2m\dot{r}\dot{\omega} \quad (11.3.23)$$

where we have used $\ddot{\theta} = \dot{\omega} = 0$. The work done on the bead is then

$$Fv_{\vartheta} = 2m\omega^2 r\dot{r} = \frac{d}{dt}(mr^2\omega^2). \quad (11.3.24)$$

We can identify this quantity as

$$\frac{d}{dt}(E - h) = \frac{dE}{dt} \quad (11.3.25)$$

since h is conserved. This accounts for the change in energy.

Previously we stated that constraint forces do no work, this is only the case for virtual displacements consistent with the constraints. The only virtual displacement here is along the rod, which is perpendicular to F . Constraint forces can still do work in actual motion.

^afine, technically if $V(r) = mr^2\omega^2/2$ then p_r will be conserved, but this is just being pedantic.

11.4 Global Symmetries and Conservation Laws

In this section we develop the consequences of two global symmetries. By a **global symmetry** we mean a symmetry that applies to all space, rather than just a specific system. We will do so by considering a system of N particles, which for sufficiently large N covers all classical systems. We take \mathbf{r}_a to be the position of the a th particle.

11.4.1 Homogeneity

We assume that space is **homogeneous**, by this we mean that it is the same anywhere, in particular we are free to choose our origin to be any point in space. Movement of the origin corresponds to adding a constant to the particle positions, which we can also view as translating the entire system.

If space is homogenous then the Lagrangian must be invariant under a simultaneous translation of all particles, that is the coordinate transformation

$$\mathbf{r}_a \rightarrow \mathbf{r}_a + \Delta\mathbf{r}_a \quad (11.4.1)$$

for some small constant $\Delta\mathbf{r} = (\Delta x, \Delta y, \Delta z)$. Notice that the velocity is unchanged by this translation. The Lagrangian changes from $\mathcal{L} \rightarrow \mathcal{L} + \Delta\mathcal{L}$. The homogeneity of space implies that $\Delta\mathcal{L} = 0$.

In one dimension with a single particle then to first order in Δx we have

$$\Delta \mathcal{L} = \frac{\partial \mathcal{L}}{\partial x} \Delta x. \quad (11.4.2)$$

In three dimensions with a single particle then to first order in Δx_i we have

$$\Delta \mathcal{L} = \frac{\partial \mathcal{L}}{\partial x} \Delta x + \frac{\partial \mathcal{L}}{\partial y} \Delta y + \frac{\partial \mathcal{L}}{\partial z} \Delta z = (\nabla \mathcal{L}) \cdot \Delta \mathbf{r}. \quad (11.4.3)$$

For N particles in three dimensions we then have

$$\Delta \mathcal{L} = \sum_a (\nabla_{\mathbf{r}_a} \mathcal{L}) \cdot \Delta \mathbf{r}_a = 0 \quad (11.4.4)$$

where $\nabla_{\mathbf{r}_a} = (\partial_{x_a}, \partial_{y_a}, \partial_{z_a})$. Since $\Delta \mathbf{r}_a$ is arbitrary we have

$$\sum_a \nabla_{\mathbf{r}_a} \mathcal{L} = 0. \quad (11.4.5)$$

Recalling that we can write the Euler–Lagrange equations in terms of the canonical momenta as

$$\frac{dp_i}{dt} = \frac{\partial \mathcal{L}}{\partial q_i} \implies \frac{d\mathbf{p}}{dt} = \nabla \mathcal{L} \quad (11.4.6)$$

we find that

$$0 = \sum_a \frac{d\mathbf{p}_a}{dt} = \frac{d\mathbf{P}}{dt}. \quad (11.4.7)$$

This means that the total momentum, \mathbf{P} , is conserved.

Homogeneity of space implies conservation of total momentum.

11.4.2 Isotropy

We assume that space is **isotropic**, by this we mean that it is the same in all directions, in particular we are free to choose our axes to point in any direction. We can view rotations of the axes as rotations of the system by switching between active and passive transformations.

If space is isotropic then the Lagrangian must be invariant under a rotation of the entire system. A rotation takes $\mathbf{r}_a \rightarrow \mathbf{r}_a + \Delta \mathbf{r}_a$ where

$$\Delta \mathbf{r}_a = \Phi \times \mathbf{r}_a \quad (11.4.8)$$

with $\Phi = \hat{\Phi} \hat{\Phi}$ representing a rotation through a small angle, $\hat{\Phi}$, about an axis $\hat{\Phi}$.

Notice that the velocity is not conserved under this since the direction changes. The change in velocity is

$$\Delta \dot{\mathbf{r}}_a = \Phi \times \dot{\mathbf{r}}_a. \quad (11.4.9)$$

These rotations correspond to infinitesimal rotations but from infinitesimal rotations we can build finite rotations, and so the results of this derivation apply to any rotation.

Since the velocity is not constant under this transformation the resulting change in the Lagrangian also contains a term corresponding to the change in velocity, giving

$$\Delta \mathcal{L} = \sum_a (\nabla_{\mathbf{r}_a} \mathcal{L}) \cdot \Delta \mathbf{r}_a + \sum_a (\nabla_{\dot{\mathbf{r}}_a} \mathcal{L}) \cdot \Delta \dot{\mathbf{r}}_a. \quad (11.4.10)$$

We demand that this is zero if space is isotropic.

Substituting for $\Delta \mathbf{r}_a$ and $\Delta \dot{\mathbf{r}}_a$ we get

$$0 = \sum_a (\nabla_{\mathbf{r}_a} \mathcal{L}) \cdot (\boldsymbol{\Phi} \times \mathbf{r}_a) + \sum_a (\nabla_{\dot{\mathbf{r}}_a} \mathcal{L}) \cdot (\boldsymbol{\Phi} \times \dot{\mathbf{r}}_a). \quad (11.4.11)$$

Using the Euler–Lagrange equations we have

$$\frac{\partial \mathcal{L}}{\partial x_i} = \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}_i} \right) \implies \nabla_{\mathbf{r}} \mathcal{L} = \frac{d}{dt} \nabla_{\dot{\mathbf{r}}} \mathcal{L}. \quad (11.4.12)$$

It then follows that

$$0 = \sum_a \left(\frac{d}{dt} \nabla_{\dot{\mathbf{r}}_a} \mathcal{L} \right) \cdot (\boldsymbol{\Phi} \times \mathbf{r}_a) + \sum_a (\nabla_{\dot{\mathbf{r}}_a} \mathcal{L}) \cdot (\boldsymbol{\Phi} \times \dot{\mathbf{r}}_a). \quad (11.4.13)$$

The final term is the scalar triple product, which follows the relation $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{b} \cdot (\mathbf{c} \times \mathbf{a}) = (\mathbf{c} \times \mathbf{b}) \cdot \mathbf{a}$. Applying this we have

$$0 = \sum_a \left[\mathbf{r}_a \times \left(\frac{d}{dt} \nabla_{\dot{\mathbf{r}}_a} \mathcal{L} \right) \right] \cdot \boldsymbol{\Phi} + \sum_a [\dot{\mathbf{r}}_a \times (\nabla_{\dot{\mathbf{r}}_a} \mathcal{L})] \cdot (\boldsymbol{\Phi} \times \dot{\mathbf{r}}_a) \quad (11.4.14)$$

$$= \sum_a \left[\left(\frac{d}{dt} \nabla_{\dot{\mathbf{r}}_a} \mathcal{L} \right) \cdot (\times \mathbf{r}_a) + \sum_a [\dot{\mathbf{r}}_a \times (\nabla_{\dot{\mathbf{r}}_a} \mathcal{L})] \times \dot{\mathbf{r}}_a \right] \cdot \boldsymbol{\Phi} \quad (11.4.15)$$

Noticing that since $\boldsymbol{\Phi}$ is constant this is the result of applying the product rule we have

$$0 = \left[\frac{d}{dt} \left(\sum_a \mathbf{r}_a \times (\nabla_{\dot{\mathbf{r}}_a} \mathcal{L}) \right) \right] \cdot \boldsymbol{\Phi}. \quad (11.4.16)$$

Since this holds for arbitrary small $\boldsymbol{\Phi}$ we must have

$$\frac{d}{dt} \left(\sum_a \mathbf{r}_a \times (\nabla_{\dot{\mathbf{r}}_a} \mathcal{L}) \right) = 0 \quad (11.4.17)$$

which gives

$$\sum_a \mathbf{r}_a \times (\nabla_{\dot{\mathbf{r}}_a} \mathcal{L}) = \sum_a \mathbf{r}_a \times \mathbf{p}_a = \mathbf{L} = \text{constant}. \quad (11.4.18)$$

That is the total angular momentum is conserved.

Isotropy of space implies conservation of total angular momentum.

11.4.3 Notes

We have proven the above without needing Newton's third law, which shows that conservation of total linear/angular momentum is very fundamental (assuming of course that space is homogenous and isotropic).

What we have seen here is another example of Noether's theorem. In particular we have seen that infinitesimal symmetries give rise to global conservation laws. A more mathematically precise statement is that the 'infinitesimal symmetries' are generators for the Lie algebra of the relevant Lie group. The group describing translational symmetries is \mathbb{R}^3 , and the associated Lie algebra is \mathbb{R}^3 with the Lie bracket $[\mathbf{a}, \mathbf{b}] = 0$ for all $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$. The generators of \mathbb{R}^3 , as a Lie group, are ∂_i . The rotation group describing rotational symmetry is $\text{SO}(3)$, and the infinitesimal rotations we have considered are elements of the Lie algebra $\mathfrak{so}(3)$, the generators of $\mathfrak{so}(3)$ are T_i and satisfy the commutation relations $[T_i, T_k] = \varepsilon_{ijk} T_k$.

Notice that, up to factors of \hbar and i the generators of \mathbb{R}^3 (associated with conservation of momentum) are the momentum operators in quantum mechanics and the generators of $\mathfrak{so}(3)$ (associated with conservation of angular momentum) are the angular momentum operators. This hints at a deep connection between symmetries and fundamental concepts of quantum mechanics, and indeed the Lagrangian formalism can be applied to quantum mechanics to exploit this connection.

Twelve

Moving Away From Newtonian Physics

Physics is not broken.

Jenni Smillie

12.1 Velocity-Dependent Forces

In [Section 8.2](#) when we derived the Euler–Lagrange equations we explicitly ruled out velocity dependent potentials to derive the Euler-Lagrange equation as a special case of the more general

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}} \right) - \frac{\partial T}{\partial q} = Q. \quad (12.1.1)$$

For the velocity-independent case we can find $V = V(\{q\}, \{\dot{q}\}, t)$ such that

$$Q = -\frac{\partial V}{\partial q}. \quad (12.1.2)$$

We also have $\partial V / \partial \dot{q} = 0$, which we used to insert zero into the derivation.

If we have velocity dependent forces however then we cannot do this. Suppose instead that we can find a function $\tilde{V} = \tilde{V}(\{q\}, \{\dot{q}\}, t)$ such that

$$Q = \frac{d}{dt} \left(\frac{\partial \tilde{V}}{\partial \dot{q}} \right) - \frac{\partial \tilde{V}}{\partial q}. \quad (12.1.3)$$


This is such that if we substitute it in and follow the same process as deriving the original Lagrangian we get the familiar equations

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}} \right) - \frac{\partial \mathcal{L}}{\partial q}, \quad (12.1.4)$$

except now $\mathcal{L} = T - \tilde{V}$.

While \tilde{V} is not a potential as we normally think of them all of the important results carry over such as Hamilton’s principle and conservation laws since these were derived from the form of the Euler–Lagrange equations, which is unchanged.

12.2 The Lorentz Force


 See [Chapter B](#) for more details on electromagnetism.

Perhaps the most important velocity dependent force is the **Lorentz force**,

$$m\ddot{\mathbf{r}} = \mathbf{F} = e(\mathbf{E} + \dot{\mathbf{r}} \times \mathbf{B}). \quad (12.2.1)$$

Here m is the mass of the particle, e is its charge, \mathbf{E} and \mathbf{B} are some position and time dependent electric and magnetic fields, \mathbf{F} is the Lorentz force, and \mathbf{r} is the position of the particle.

12.2.1 Electromagnetic Lagrangian

 In this section we use index notation and the Einstein summation convention.

We can now derive a Lagrangian that works with velocity dependent Lorentz force. We start with the following result relating the product of Levi-Civita symbols, ε_{ijk} , to Kronecker deltas, δ_{ij} ¹:

$$\varepsilon_{ijk}\varepsilon_{klm} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}. \quad (12.2.2)$$

¹see methods of theoretical physics: vectors, tensors and continuum mechanics notes for more details

In index notation the Lorentz force has a component in the x_i direction given by

$$F_i = e [\mathbf{E} + \dot{\mathbf{r}} \times \mathbf{B}]_i \quad (12.2.3)$$

$$= e \left[-\nabla\varphi - \frac{\partial \mathbf{A}}{\partial t} + \dot{\mathbf{r}} \times (\nabla \times \mathbf{A}) \right]_i \quad (12.2.4)$$

where we have substituted the fields for the potentials. Considering the last term we have

$$[\dot{\mathbf{r}} \times (\nabla \times \mathbf{A})]_i = \varepsilon_{ijk}\dot{x}_j(\nabla \times \mathbf{A})_k \quad (12.2.5)$$

$$= \varepsilon_{ijk}\dot{x}_j\varepsilon_{klm}\partial_l A_m \quad (12.2.6)$$

$$= \varepsilon_{ijk}\varepsilon_{klm}\dot{x}_j\partial_l A_m \quad (12.2.7)$$

$$= (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})\dot{x}_j\partial_l A_m \quad (12.2.8)$$

$$= \dot{x}_j\partial_i A_j - \dot{x}_j\partial_j A_i. \quad (12.2.9)$$

Here $\partial_i = \partial/\partial x_i$, we will also use $\partial_t = \partial/\partial t$. With this we have $(\nabla\varphi)_i = \partial_i\varphi$ and so

$$F_i = e \left[-\partial_i\varphi - \partial_t A_i + \dot{x}_j(\partial_i A_j - \partial_j A_i) \right]. \quad (12.2.10)$$

We wish to identify a function, \tilde{V} , such that

$$F_i = \frac{d}{dt} \left(\frac{\partial \tilde{V}}{\partial \dot{x}_i} \right) - \frac{\partial \tilde{V}}{\partial x_i}. \quad (12.2.11)$$

There isn't really a trick to this, it just has to be done by inspection and then verified. We start by noting that if we choose

$$\tilde{V} = e(\varphi - \dot{x}_j A_j) \quad (12.2.12)$$

then we have

$$-\frac{\partial \tilde{V}}{\partial x_i} = -e\partial_i\varphi + \frac{\partial}{\partial x_i}(\dot{x}_j A_j) = -e\partial_i\varphi + \dot{x}_j\partial_i A_j. \quad (12.2.13)$$

Note that we treat \dot{x}_j and x_i as independent variables. This matches the first and third term of the Lorentz force.

Now considering the other terms, and remembering that both the scalar and vector potential are independent of the velocity, so the only velocity dependence is in the \dot{x}_j term, we have

$$\frac{d}{dt} \left(\frac{\partial \tilde{V}}{\partial \dot{x}_i} \right) = e \frac{d}{dt} \left[0 - \frac{\partial \dot{x}_j}{\partial \dot{x}_i} A_j \right] \quad (12.2.14)$$

$$= -e \frac{dA_i}{dt} \quad (12.2.15)$$

$$= -e \frac{\partial A_i}{\partial x_j} \frac{\partial x_j}{\partial t} - e \frac{\partial A_i}{\partial t} \quad (12.2.16)$$

$$= -e \left[\partial_t A_i + \dot{x}_j \partial_j A_i \right]. \quad (12.2.17)$$

This matches the second and fourth terms of the Lorentz force. We have therefore demonstrated that \tilde{V} satisfies the necessary equations and therefore we can use the Lagrangian

$$\mathcal{L} = \frac{1}{2} m \dot{x}_i \dot{x}_i - e (\varphi - \dot{x}_i A_i), \quad (12.2.18)$$

or in vector notation

$$\mathcal{L} = \frac{1}{2} m \sum_i \dot{x}_i^2 - e [\varphi - \dot{\mathbf{r}} \cdot \mathbf{A}]. \quad (12.2.19)$$

12.2.2 Properties Derived from the Lagrangian

The canonical momentum is given by

$$p_i = \frac{\partial \mathcal{L}}{\partial \dot{x}_i} = m \dot{x}_i + e A_i, \quad (12.2.20)$$

or in vector form

$$\mathbf{p} = m\mathbf{v} + e\mathbf{A}, \quad (12.2.21)$$

which is *not* equal to the mechanical momentum, $m\mathbf{v}$. It is important that when we apply rules, such as conservation of momentum if the Lagrangian is independent of position, that we remember it is the *canonical* momentum that we work with, not the mechanical momentum.

The energy function is

$$h = \mathbf{p} \cdot \dot{\mathbf{r}} - \mathcal{L} \quad (12.2.22)$$

$$= [m\mathbf{v} + e\mathbf{A}] \cdot \mathbf{v} - \frac{1}{2} m v^2 + e(\varphi - \mathbf{v} \cdot \mathbf{A}) \quad (12.2.23)$$

$$= \frac{1}{2} m v^2 + e\varphi. \quad (12.2.24)$$

This quantity is conserved only if \mathcal{L} has no explicit time dependence. This requires that φ and \mathbf{A} have no explicit time dependence, and hence \mathbf{E} and \mathbf{B} cannot have any explicit time dependence.

When this is the case h is conserved, and since the kinetic energy is quadratic in the velocities $h = E$. This result for the energy is what we expect, it is equal to the sum of the kinetic energy and electrostatic energy. Remember that \mathbf{B} does no work since $\mathbf{v} \times \mathbf{B}$ is perpendicular to the direction of the velocity. If \mathbf{B} is time dependent then the magnetic field still does no work but there is a term $-\partial\mathbf{A}/\partial t$ in the electric field, which contributes to the energy, but is not accounted for in the energy function.

■ **Example 12.2.25** Consider a single particle moving in a toroidal \mathbf{B} -field with $\varphi = 0$. Such a field has potential

$$\mathbf{A} = (0, 0, A_z) = (0, 0, B\rho) \quad (12.2.26)$$

where $\rho = \sqrt{x^2 + y^2}$. Find two constants of motion of the field.

We note here that the magnetic field arising from this potential is

$$\mathbf{B} = \nabla \times \mathbf{A} = \frac{B}{\rho}(y, -x, 0) \quad (12.2.27)$$

which gives circles in the (x, y) -plane. It should be noted that we don't actually need to compute \mathbf{B} , we do so here only to justify calling this a toroidal field.

The Lagrangian is

$$\mathcal{L} = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) + e\dot{z}A_z. \quad (12.2.28)$$

Since $A_z = B\sqrt{x^2 + y^2}$ has no explicit z dependence the canonical momentum in the z -direction,

$$p_z = \frac{\partial \mathcal{L}}{\partial \dot{z}} = m\dot{z} + eB\sqrt{x^2 + y^2}, \quad (12.2.29)$$

is conserved. This is one constant of the motion.

The Lagrangian also has no explicit time dependence, hence the energy function,

$$h = \frac{1}{2}m\dot{v}^2 + e\varphi = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2), \quad (12.2.30)$$

is conserved. That is the kinetic energy is conserved, which makes sense since the \mathbf{B} -field can do no work to change the kinetic energy.

12.3 Special Relativity

Lagrangian methods can also be applied to special relativity². The problem we have is that in special relativity we work with the relativistic three-momentum,

$$\mathbf{p} = \gamma(u)m\mathbf{u} \quad (12.3.1)$$

where m is the mass, \mathbf{u} the velocity, and $\gamma(u) = (1 - u^2/c^2)^{-1/2}$, with c being the speed of light. The relativistic kinetic energy is

$$T = (\gamma - 1)mc^2. \quad (12.3.2)$$

²For more details on special relativity see the relativity section of the relativity, nuclear, and particle physics course.

We want the canonical momentum calculated from the Lagrangian for a free particle to be the relativistic three-momentum, but this isn't what we get with $\mathcal{L} = (\gamma - 1)mc^2$. Instead it turns out that the correct Lagrangian is

$$\mathcal{L} = -mc^2 \sqrt{1 - \frac{u^2}{c^2}} - V(\mathbf{r}) = -\frac{1}{\gamma(u)} mc^2 - V(\mathbf{r}). \quad (12.3.3)$$

We then have that

$$p_x = \frac{\mathcal{L}}{\dot{x}} = \frac{m\dot{x}}{\sqrt{1 - u^2/c^2}} = \gamma(u)m\dot{x}. \quad (12.3.4)$$

We then have the required $\mathbf{p} = \gamma(u)m\mathbf{u}$. Using Lagrange's equations written in terms of the momentum,

$$\frac{dp_i}{dt} = \frac{\partial \mathcal{L}}{\partial x_i} \quad (12.3.5)$$

we have

$$\dot{\mathbf{p}} = \nabla \mathcal{L} = -\nabla V. \quad (12.3.6)$$

This is what we would expect, so the Lagrangian gives the correct equations of motion, which justifies the choice.

12.4 Charged Relativistic Particle in an EM Field


We want to find a Lagrangian that gives the equations of motion

$$\frac{d\mathbf{p}}{dt} = \mathbf{F} = e(\mathbf{E} + \mathbf{u} \times \mathbf{B}). \quad (12.4.1)$$

Where \mathbf{p} is the relativistic three-momentum. Fortunately we have already corrected for the velocity dependent potential and different momentum individually and both corrections hold when combined, so the correct Lagrangian is

$$\mathcal{L} = -\frac{1}{\gamma(u)} mc^2 - e[\varphi - \mathbf{u} \cdot \mathbf{A}]. \quad (12.4.2)$$

12.4.1 Covariant Form

 This section is non-examinable.

In special relativity we want to write everything in a covariant way using four vectors. Recall that the four-velocity is defined as

$$u^\mu := \frac{dx^\mu}{d\tau} = \gamma(u)(c, \mathbf{u}) \quad (12.4.3)$$

where $x^\mu = (ct, \mathbf{x})$ is the four-position and τ is the proper time. We also define the four-potential as $A^\mu := (\varphi/c, \mathbf{A})$. Now consider the scalar product of these two quantities:

$$u^\mu A_\mu = \gamma(u)c \frac{\varphi}{c} - \gamma(u)\mathbf{u} \cdot \mathbf{A} = \frac{dt}{d\tau}(\varphi - \mathbf{u} \cdot \mathbf{A}) \quad (12.4.4)$$

where we have used

$$\frac{dt}{d\tau} = \gamma \iff \gamma d\tau = dt, \quad (12.4.5)$$

which is time dilation. Hence the Lagrangian can be written in a covariant form as

$$\mathcal{L} = -mc^2 \frac{d\tau}{dt} - eu^\mu A_\mu \frac{d\tau}{dt}. \quad (12.4.6)$$

Using this we can write the action as

$$S := \int \mathcal{L} dt \quad (12.4.7)$$

$$= \int \left(-mc^2 \frac{d\tau}{dt} - eu^\mu A_\mu \frac{d\tau}{dt} \right) dt \quad (12.4.8)$$

$$= - \int (mc^2 + eu^\mu A_\mu) d\tau. \quad (12.4.9)$$

For the special case of a free particle we have

$$S = -mc^2 \int d\tau, \quad (12.4.10)$$

which is just $-mc^2$ times the length of the world line of the particle. We can then apply Hamilton's principle, $\delta S = 0$, to the action to derive the dynamics of the system.

Thirteen

Hamiltonian Dynamics

13.1 Classical Mechanics Hamiltonian

Recall that we defined the energy function for a system of $3N - M$ degrees of freedom as

$$h(\{q\}, \{\dot{q}\}, t) = \sum_i p_i \dot{q}_i - \mathcal{L}(\{q\}, \{\dot{q}\}, t) \quad (13.1.1)$$

where

$$p_i := \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \quad (13.1.2)$$

are the canonical momenta. The energy function is numerically equal to the function we call the **Hamiltonian**, \mathcal{H} , its just that this function is written in terms of the canonical momenta instead of the velocity.

Definition 13.1.3 — Hamiltonian

$$\mathcal{H}(\{q\}, \{p\}, t) := \sum_i p_i \dot{q}_i - \mathcal{L}(\{q\}, \{\dot{q}\}, t). \quad (13.1.4)$$

■ **Example 13.1.5** Consider a particle moving in one dimension in a conservative, time-independent potential $V(x)$. In this case we have

$$\mathcal{L} = \frac{1}{2} m \dot{x}^2 - V(x) \quad (13.1.6)$$

and so

$$p = \frac{\partial \mathcal{L}}{\partial \dot{x}} = m \dot{x} \implies \dot{x} = \frac{p}{m} \quad (13.1.7)$$

from which we get

$$\mathcal{H} = p \dot{x} - \mathcal{L} = p \frac{p}{m} - \left(\frac{1}{2} m \dot{x}^2 - V(x) \right) = \frac{p^2}{2m} + V(x). \quad (13.1.8)$$

In this case we can identify $p^2/(2m) = m\dot{x}^2/2 = T$ and so $\mathcal{H} = E = T + V$, although this will not always be the case.

Consider the variation of \mathcal{H} due to some arbitrary small variation in its arguments. The change in \mathcal{H} is given by

$$\delta\mathcal{H} = \sum_i \left[\frac{\partial\mathcal{H}}{\partial q_i} \delta q_i + \frac{\partial\mathcal{H}}{\partial p_i} \delta p_i \right] + \frac{\partial\mathcal{H}}{\partial t} \delta t. \quad (13.1.9)$$

We can consider the result of the same change in the definition of the Hamiltonian. For example, considering the first term $p_i \dot{q}_i$ a small change $p_i \rightarrow p_i + \delta p_i$ and $\dot{q}_i \rightarrow \dot{q}_i + \delta \dot{q}_i$ results in a change

$$(p_i + \delta p_i)(\dot{q}_i + \delta \dot{q}_i) - p_i \dot{q}_i = p_i \delta \dot{q}_i + \dot{q}_i \delta p_i \quad (13.1.10)$$

where we ignore second order terms. Hence

$$\delta\mathcal{H} = \sum_i [\dot{q}_i \delta p_i + p_i \delta \dot{q}_i] - \sum_i \left[\frac{\partial\mathcal{L}}{\partial q_i} \delta q_i + \frac{\partial\mathcal{L}}{\partial \dot{q}_i} \delta \dot{q}_i \right] - \frac{\partial\mathcal{L}}{\partial t} \delta t. \quad (13.1.11)$$

Identifying $\partial\mathcal{L}/\partial \dot{q}_i = p_i$ in the last term we see it cancels with the second term. We can also rewrite $\partial\mathcal{L}/\partial q_i = \dot{p}_i$ for the third term, which is simply a statement of the Euler–Lagrange equations. This then reduces to

$$\delta\mathcal{H} = \sum_i [\dot{q}_i \delta p_i - \dot{p}_i \delta q_i] - \frac{\partial\mathcal{L}}{\partial t} \delta t. \quad (13.1.12)$$

Since this must be true for any variation the coefficients of $\delta \dot{q}_i$, δp_i , and δt in [Equations \(13.1.9\) and \(13.1.12\)](#) must match, which gives us

$$\dot{q}_i = \frac{\partial\mathcal{H}}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial\mathcal{H}}{\partial q_i}, \quad \text{and} \quad \frac{\partial\mathcal{H}}{\partial t} = -\frac{\partial\mathcal{L}}{\partial t}. \quad (13.1.13)$$

These are **Hamilton's equations** of motion. We have replaced the $3N - M$ second order Euler–Lagrange equations with $6N - 2M + 1$ first order differential equations, the last being trivial if \mathcal{L} has no explicit time dependence.

■ **Example 13.1.14** Consider the one dimensional particle in a potential $V(x)$ again. We saw in [Example 13.1.5](#) that $\mathcal{H} = p^2/(2m) + V(x)$. From Hamilton's equations we then get

$$\dot{x} = \frac{\partial\mathcal{H}}{\partial p} = \frac{p}{m}, \quad \dot{p} = -\frac{\partial\mathcal{H}}{\partial x} = -V'(x), \quad \text{and} \quad 0 = 0. \quad (13.1.15)$$

The first of these is just the standard momentum definition, $p = m\dot{x}$, and the second is the standard force-potential relationship, $F = -\nabla V$, in one dimension.

It is also possible to Derive Hamilton's equations from Hamilton's principle, $\delta S = 0$, by writing the action as

$$S = \int \left[\sum_i p_i \dot{q}_i - \mathcal{H} \right] dt \quad (13.1.16)$$

where the integrand is simply the inversion of [Equation \(13.1.4\)](#) to give \mathcal{L} .

■ **Example 13.1.17** Consider a charged, non-relativistic particle. We found in Equation (12.2.24) that the energy function was

$$h = \frac{1}{2}mv^2 + e\varphi. \quad (13.1.18)$$

We also saw in Equation (12.2.21) that the canonical momentum is

$$\mathbf{p} = m\mathbf{v} + e\mathbf{A} \implies \mathbf{v} = \frac{1}{m}(\mathbf{p} - e\mathbf{A}). \quad (13.1.19)$$

We can find the Hamiltonian by replacing the velocity in the energy function with the momentum, giving

$$\mathcal{H} = \frac{1}{2}m \left[\frac{1}{m}(\mathbf{p} - e\mathbf{A}) \right]^2 + e\varphi = \frac{1}{2m}|\mathbf{p} - e\mathbf{A}|^2 + e\varphi. \quad (13.1.20)$$

Notice that the vector potential, \mathbf{A} , appears in the Hamiltonian, even though it didn't appear in the energy function. We can apply Hamilton's equations to this to get the Lorentz force law.

13.1.1 Poisson Brackets

Consider the arbitrary function \mathcal{A} , which is a function of positions, momenta, and time. That is $\mathcal{A} = \mathcal{A}(\{q\}, \{p\}, t)$. The total time derivative of \mathcal{A} is

$$\frac{d\mathcal{A}}{dt} = \sum_i \left[\frac{\partial \mathcal{A}}{\partial q_i} \dot{q}_i + \frac{\partial \mathcal{A}}{\partial p_i} \dot{p}_i \right] + \frac{\partial \mathcal{A}}{\partial t}. \quad (13.1.21)$$

Applying Hamilton's equations this becomes

$$\frac{d\mathcal{A}}{dt} = \sum_i \left[\frac{\partial \mathcal{A}}{\partial q_i} \frac{\partial \mathcal{H}}{\partial p_i} - \frac{\partial \mathcal{A}}{\partial p_i} \frac{\partial \mathcal{H}}{\partial q_i} \right] + \frac{\partial \mathcal{A}}{\partial t}. \quad (13.1.22)$$

Definition 13.1.23 — Poisson Bracket For two functions, \mathcal{A} and \mathcal{B} , which depend on position, momenta, and time, we define the **Poisson bracket** as

$$\{\mathcal{A}, \mathcal{B}\} := \sum_i \left[\frac{\partial \mathcal{A}}{\partial q_i} \frac{\partial \mathcal{B}}{\partial p_i} - \frac{\partial \mathcal{A}}{\partial p_i} \frac{\partial \mathcal{B}}{\partial q_i} \right]. \quad (13.1.24)$$

We see that

$$\frac{d\mathcal{A}}{dt} = \{\mathcal{A}, \mathcal{H}\} + \frac{\partial \mathcal{A}}{\partial t}. \quad (13.1.25)$$

We can use the Poisson brackets to rewrite Hamilton's equations as

$$\dot{q}_i = \{q_i, \mathcal{H}\}, \quad \dot{p}_i = \{p_i, \mathcal{H}\}, \quad \text{and} \quad \frac{\partial \mathcal{H}}{\partial t} = -\frac{\partial \mathcal{L}}{\partial t}. \quad (13.1.26)$$

This is a nice way to write them as the first two equations are symmetric in their signs, although we still have to remember which order the terms in the Poisson bracket are.

13.2 Quantum Mechanics

13.2.1 The Hamiltonian and Quantum Mechanics

R For more details on quantum mechanics see the notes for principles of quantum mechanics.

Given a classical system with a corresponding Hamiltonian, \mathcal{H} , we can create an operator in quantum mechanics by making the replacement

$$\mathcal{H}(q, p, t) \rightarrow \hat{\mathcal{H}}\left(q, -i\hbar \frac{\partial}{\partial q}, t\right) \quad (13.2.1)$$

which then acts on the wave function in the position basis. This process is called **canonical quantisation**.

■ **Example 13.2.2** For a single particle in one dimension we have

$$\mathcal{H} = \frac{p^2}{2m} + V(x) \rightarrow \hat{\mathcal{H}} = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x). \quad (13.2.3)$$

For a charged particle in an electromagnetic field we have

$$\mathcal{H} = \frac{1}{2m} |\mathbf{p} - e\mathbf{A}|^2 + e\varphi \quad (13.2.4)$$

and hence

$$\hat{\mathcal{H}} = -\frac{\hbar^2}{2m} \nabla^2 - \frac{e\hbar}{2im} [2\mathbf{A} \cdot \nabla + (\nabla \cdot \mathbf{A})] + \frac{e^2}{2m} |\mathbf{A}|^2 + e\varphi. \quad (13.2.5)$$

R There is some ambiguity in this prescription. For example, classically $\mathbf{p} \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{p}$, but $\mathbf{A} \cdot \nabla \neq \nabla \cdot \mathbf{A}$. This is a result of multiple quantum systems having the same classical limit.

To continue on with our quantisation we note that for two classical quantities, \mathcal{A} and \mathcal{B} , we can make the same change, $q \rightarrow q, p \rightarrow -i\hbar \partial_q, t \rightarrow t$, to get operators, $\hat{\mathcal{A}}$ and $\hat{\mathcal{B}}$ for these two quantities. The commutation relations for these operators follow from

$$i\hbar\{\mathcal{A}, \mathcal{B}\} = [\hat{\mathcal{A}}, \hat{\mathcal{B}}], \quad (13.2.6)$$

where

$$[\hat{\mathcal{A}}, \hat{\mathcal{B}}] := \hat{\mathcal{A}}\hat{\mathcal{B}} - \hat{\mathcal{B}}\hat{\mathcal{A}} \quad (13.2.7)$$

is the **commutator**.

Multiplying Equation (13.1.25) by $i\hbar$ and then making the replacement of Poisson brackets for commutators and functions for operators we see that

$$\frac{d\hat{\mathcal{A}}}{dt} = [\hat{\mathcal{A}}, \hat{\mathcal{H}}] + i\hbar \frac{\partial \hat{\mathcal{A}}}{\partial t}. \quad (13.2.8)$$

This is Heisenberg's equation of motion for a quantum operator in the Heisenberg picture.

13.2.2 The Lagrangian and Quantum Mechanics

To finish of the section on quantum mechanics we note that the Lagrangian formalism can also be used in quantum mechanics. In this case we define the transition amplitude for a particle to travel from x_a at time t_a to x_b at time t_b as

$$\langle x_b, t_b | x_a, t_a \rangle = A \sum_{\text{paths}} \exp \left[\frac{i}{\hbar} S[x(t)] \right] = \int \exp \left[\frac{i}{\hbar} \int_{t_a}^{t_b} \mathcal{L} dt \right] \mathcal{D}x \quad (13.2.9)$$

where A is a normalisation factor, the sum is over all paths from x_a to x_b , and $S[x(t)]$ is the classical action,

$$S[x(t)] = \int_{t_a}^{t_b} \mathcal{L}(x, \dot{x}) dt. \quad (13.2.10)$$

We define the integral in Equation (13.2.9) as the **path integral**, from this we can derive all of the same results as we could with the more familiar operator formalism.¹

In the limit of $\hbar \rightarrow 0$, i.e. the classical limit, we see that the exponential in the path integral oscillates rapidly, unless we are near an area where the action is stationary. The oscillations mostly cancel out and the integral is dominated by the regions near the stationary action², i.e. where $\delta S = 0$, so we recover Hamilton's principle.

¹see the notes for the quantum theory course for more details on path integrals

²this is the same logic as the method of stationary phase, discussed in the notes for methods of mathematical physics.

Part III

Rigid Body Motion

Fourteen

Euler's Method

14.1 Recap

Recall that a rigid body is a system of N particles, labelled $a = 1, \dots, N$, which have positions $\mathbf{r}_a = (x_a, y_a, z_a) = (x_{a1}, x_{a2}, x_{a3})$ and masses m_a . There is also the constraint that the distance between particles is fixed. That is

$$|\mathbf{r}_a - \mathbf{r}_b| = \rho_{ab} = \text{constant}. \quad (14.1.1)$$

These are not linearly independent constraints, in fact we have already shown that they combine to imply

$$M\ddot{\mathbf{R}} = \mathbf{F}^{\text{ext}}, \quad \text{and} \quad \dot{\mathbf{L}} = \mathbf{G}^{\text{ext}} \quad (14.1.2)$$

where $M = \sum_a m_a$ is the total mass, $\mathbf{R} = \sum_a \mathbf{r}_a / M$ is the centre of mass, \mathbf{F}^{ext} is the net external force on the body, \mathbf{L} is the total angular momentum, and \mathbf{G}^{ext} is the net external torque.

We also showed that

$$\mathbf{L} = \mathbf{J} + \mathbf{R} \times \mathbf{P}, \quad \text{and} \quad T = T_{\text{com}} + \frac{1}{2}MV^2 \quad (14.1.3)$$

where \mathbf{J} is the intrinsic angular momentum, which is independent of the choice of origin, \mathbf{P} is the total linear momentum, T is the total kinetic energy, T_{com} is the kinetic energy in the centre-of-mass frame, and $V = |\dot{\mathbf{R}}|$ is the velocity of the centre of mass.

Recall that we defined the centre-of-mass frame as the inertial (i.e. non-rotating) frame in which the centre of mass is stationary, but not necessarily at the origin. In this frame the instantaneous motion of the body is purely rotational.

14.2 Euler's Theorem

Theorem 14.2.1 — Euler's Theorem. Any displacement of a rigid body with one point fixed in space^a can be described as a rotation about some single axis.

^aThe requirement that a point stays fixed essentially rules out translations.

Proof. Rotations are represented by 3×3 orthogonal matrices, $R \in \text{O}(3)$. An equivalent statement of the theorem is that there is a non-zero vector, \mathbf{n} for

which $R\mathbf{n} = \mathbf{n}$, i.e. the vector is invariant under this particular rotation. This is the aforementioned 'single axis' of rotation. This statement is equivalent to saying that \mathbf{n} is an eigenvector of R with eigenvalue 1. Therefore to prove the statement it suffices to prove that all $R \in O(3)$ have 1 as an eigenvalue. We know that $R \in O(3)$, so by definition $RR^T = R^T R = I$, i.e. $R^{-1} = R^T$. By properties of the determinant we then have

$$1 = \det(I) = \det(RR^T) = \det(R) \det(R^T) = [\det(R)]^2 \quad (14.2.2)$$

since $\det(AB) = \det(A) \det(B)$ and $\det(A^T) = \det A$. and so $\det(R) = \pm 1$, indeed we can define $O(3)$ as the set of 3×3 real matrices with determinant ± 1 . We suppose that R is a proper rotation, so $\det(R) = 1$, that is $R \in SO(3)$. Another property of determinants is that $\det(\lambda A) = \lambda^3 \det(A)$ for a 3×3 matrix A . Using this we have

$$\det(-A) = (-1)^3 \det A = -\det A. \quad (14.2.3)$$

Hence, since we assume $\det R = 1$, we have $\det R^{-1} = \det R^T = \det R = 1$. Thus,

$$\det(R - I) = \det((R - I)^T) \quad (14.2.4)$$

$$= \det(R^T - I) \quad (14.2.5)$$

$$= \det(R^{-1} - R^{-1}R) \quad (14.2.6)$$

$$= \det(R^{-1}(I - R)) \quad (14.2.7)$$

$$= \det(R^{-1}) \det(I - R) \quad (14.2.8)$$

$$= \det(I - R) \quad (14.2.9)$$

$$= \det(-(R - I)) \quad (14.2.10)$$

$$= -\det(R - I). \quad (14.2.11)$$

Hence $\det(R - I) = 0$. This is equivalent to stating that 1 is an eigenvalue of R , and so we are finished. \square

Note that this theorem only works in odd dimensions, since we used the dimension of the space in the $\det(\lambda A) = \lambda^3 \det(A)$ step to get a negative.

To specify a rotation we need 6 degrees of freedom:

- 2 degrees of freedom to specify the axis of rotation (we can either give its angle to the x and z axis (as in spherical polars) or a unit vector, which is two degrees of freedom since the three degrees of freedom of a vector are reduced to two by the requirement of unit modulus).
- 1 degree of freedom to specify the size of rotation, i.e. the angle through which the object is rotated.
- 3 degrees of freedom to specify the fixed position.

There are other ways to specify rotations, which we will come to later, but the number of degrees of freedom are the same.

The velocity of the system also has six degrees of freedom, typically given as the velocity of the centre of mass, $\dot{\mathbf{R}}$, and angular velocity, $\boldsymbol{\omega}$. Note that we *cannot*

use the quantity $\Omega(t) = \int_0^t \omega(t') dt'$ as a generalised coordinate, in the way we use $R(t) = \int_0^t \dot{R}(t') dt'$, since rotations don't commute and it is possible that multiple rotations will give the same value for $\Omega(t)$. We will solve this problem later.

14.3 The Inertia Tensor

In the centre-of-momentum frame, taking the origin as the centre-of-mass the instantaneous velocity of the a th particle is

$$\dot{\mathbf{r}}_a = \boldsymbol{\omega}(t) \times \mathbf{r}_a. \quad (14.3.1)$$

The intrinsic angular momentum is then

$$\mathbf{J} = \sum_a \mathbf{r}_a \times \mathbf{p}_a \quad (14.3.2)$$

$$= \sum_a m_a (\mathbf{r}_a \times \dot{\mathbf{r}}_a) \quad (14.3.3)$$

$$= \sum_a m_a (\mathbf{r}_a \times (\boldsymbol{\omega} \times \mathbf{r}_a)) \quad (14.3.4)$$

$$= \sum_a m_a r_a^2 \boldsymbol{\omega} - \sum_a m_a \mathbf{r}_a (\mathbf{r}_a \cdot \boldsymbol{\omega}). \quad (14.3.5)$$

We have used $\dot{\mathbf{r}}_a = \boldsymbol{\omega} \times \mathbf{r}_a$ and the triple vector identity

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}. \quad (14.3.6)$$

Here we have derived an relation between \mathbf{J} and $\boldsymbol{\omega}$ that is linear in $\boldsymbol{\omega}$, but where \mathbf{J} doesn't necessarily point in the direction of $\boldsymbol{\omega}$ since the second sum is in the direction of \mathbf{r}_a and is in general non-zero and not in the direction of $\boldsymbol{\omega}$. A linear relation that changes the direction of the vector can be encoded in a rank 2 tensor as

$$\mathbf{J} = I\boldsymbol{\omega}, \quad \text{or} \quad J_i = \sum_{j=1}^3 I_{ij} \omega_j. \quad (14.3.7)$$

We call I the **inertia tensor**. In Cartesian coordinates by writing out the components of Equation (14.3.5) as

$$J_i = \sum_a m_a r_a^2 \omega_i - \sum_a m_a x_{ai} \sum_j x_{aj} \omega_j. \quad (14.3.8)$$

From which we can identify the Cartesian components are

$$I_{ij} = \sum_a m_a [r_a^2 \delta_{ij} - x_{ai} x_{aj}]. \quad (14.3.9)$$

Writing this out in a matrix we have

$$I = \sum_a m_a \begin{pmatrix} y_a^2 + z_a^2 & -x_a y_a & -x_a z_a \\ -x_a y_a & x_a^2 + z_a^2 & -y_a z_a \\ -x_a y_a & -y_a z_a & x_a^2 + y_a^2 \end{pmatrix}. \quad (14.3.10)$$

This explicitly involves the particle coordinates which are measured relative to the centre of mass. For a rigid body we can take these coordinates to be constants in

the frame of reference in which the body is completely stationary, called the **body frame**. This may not be an inertial frame if the body is accelerating, however it is still useful to evaluate the inertia tensor in this frame.

In the centre of momentum frame the kinetic energy is

$$T = \sum_a \frac{1}{2} m_a (\boldsymbol{\omega} \times \mathbf{r}_a) \cdot (\boldsymbol{\omega} \times \mathbf{r}_a). \quad (14.3.11)$$

Applying the scalar triple product identity, $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$, we get

$$T = \frac{1}{2} \boldsymbol{\omega} \cdot \sum_a m_a \mathbf{r}_a \times (\boldsymbol{\omega} \times \mathbf{r}_a) = \frac{1}{2} \boldsymbol{\omega} \cdot \mathbf{J} = \frac{1}{2} \boldsymbol{\omega} \cdot I \boldsymbol{\omega}. \quad (14.3.12)$$

The inertia tensor is symmetric, and hence there exist real eigenvalues, $\{I_i\}$, which we call the **principle moments of inertia**, and mutually orthogonal eigenvectors, $\{\mathbf{e}_i\}$, which we take to be normalised. We define the \mathbf{e}_i -axes as the **principle axes** (PA). In the $\{\mathbf{e}_i\}$ basis the inertia tensor is diagonal:

$$I = \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix}. \quad (14.3.13)$$

In this basis the intrinsic angular momentum is given by

$$\mathbf{J} = (I_1 \omega_1, I_2 \omega_2, I_3 \omega_3), \quad (14.3.14)$$

and the kinetic energy by

$$T = \frac{1}{2} (I_1 \omega_1^2 + I_2 \omega_2^2 + I_3 \omega_3^2) \quad (14.3.15)$$

which is the centre-of-mass kinetic energy since the principle-axes frame is a centre of momentum frame.

14.4 Other Origins

So far we have considered the origin, O , to be the centre of mass. This is a good choice for a body undergoing arbitrary rotations, since there is no other special point to pick. A common case however is of a body in which some other point, O' , is fixed. The most common example being a spinning top which has its point of contact with the table fixed (of course, real spinning tops tend to move about as well but we aren't considering translations here).

The total angular momentum about O' is

$$\mathbf{L} = \mathbf{J} + \mathbf{R} \times \mathbf{P} \quad (14.4.1)$$

where \mathbf{J} is the angular momentum about O , \mathbf{R} is the centre of mass, and \mathbf{P} is the total linear momentum. Notice that in the case of $O = O'$ this reduces to $\mathbf{L} = \mathbf{J}$ since $\mathbf{R} = \mathbf{0}$. Euler's theorem, that there is a single axis of rotation, still holds and we have

$$\mathbf{L} = I' \boldsymbol{\omega}, \quad (14.4.2)$$

where I' is the momentum of inertia tensor taking O' as the origin. We can find this by changing the coordinates in the definition of I_{ij} (Equation (14.3.9)). To do so we

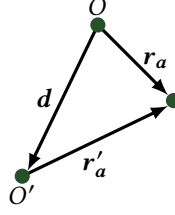


Figure 14.1: The relationship between the two origins, O and O' , and some arbitrary point.

use $\mathbf{r}_a = \mathbf{d} + \mathbf{r}'_a$ where \mathbf{r}_a is the position relative to O , \mathbf{r}'_a is the position relative to O' , and \mathbf{d} is the constant displacement of O' relative to O , all of these quantities being in the body frame, and so don't change as the object rotates. See [Figure 14.1](#).

The kinetic energy, including that of the centre of mass motion in the principle axis basis, is then given by

$$T = \frac{1}{2} \boldsymbol{\omega} \cdot I' \boldsymbol{\omega} = \frac{1}{2} (I'_1 \omega_1^2 + I'_2 \omega_2^2 + I'_3 \omega_3^2). \quad (14.4.3)$$

One common use case is when O' is a distance, l , from the centre of mass along the 3-axis. In the principle axis basis we then have

$$I'_1 = I_1 + Ml^2, \quad I'_2 = I_2 + Ml^2, \quad \text{and} \quad I'_3 = I_3. \quad (14.4.4)$$

This is the **parallel axis theorem**, which relates the moment of inertia about an axis parallel to an axis through the centre of mass with known moment of inertia.

14.5 The Eulerian Approach

Euler approached the problem of describing a rotating rigid body by considering the noninertial body frame, which rotates with the body. In this frame the principle axes are constant in time, as they too rotate with the body. Consider then some inertial frame, S_0 , which coincides instantaneously with the body frame, S . We saw previously in [Section 5.2](#) that for an instantaneous angular velocity of $\boldsymbol{\omega}$ any vector, \mathbf{A} , evolves according to

$$[\dot{\mathbf{A}}]_{S_0} = [\dot{\mathbf{A}}]_S + \boldsymbol{\omega} \times [\mathbf{A}]_S. \quad (14.5.1)$$

Note that $\mathbf{A} = [\mathbf{A}]_S = [\mathbf{A}]_{S_0}$ at the instant when the frames coincide.

Using this we have the usual equation of motion in frame S_0 that

$$[\dot{\mathbf{L}}]_{S_0} = \mathbf{G}, \quad (14.5.2)$$

where \mathbf{L} is the angular momentum and \mathbf{G} is the total external torque on the system. In the body frame we have

$$\mathbf{G} = [\dot{\mathbf{L}}]_S + \boldsymbol{\omega} \times [\mathbf{L}]_S. \quad (14.5.3)$$

Since S is the frame of the principle axes we have

$$[\mathbf{L}]_S = \mathbf{J} = (I_1 \omega_1, I_2 \omega_2, I_3 \omega_3). \quad (14.5.4)$$

Therefore we have

$$G_1 = I_1 \dot{\omega}_1 + (I_3 - I_2) \omega_2 \omega_3, \quad (14.5.5)$$

$$G_2 = I_2 \dot{\omega}_2 + (I_1 - I_3) \omega_3 \omega_1, \quad (14.5.6)$$

$$G_3 = I_3 \dot{\omega}_3 + (I_2 - I_1) \omega_1 \omega_2. \quad (14.5.7)$$

These are **Euler's equations of motion** for rigid body motion.

The problem with this approach is that \mathbf{G} is usually specified with respect to some set of axes fixed in space. This means we can't write \mathbf{G} in the body frame until after we have solved for the motion. This means that the Eulerian approach is typically only useful in the case where $\mathbf{G} = \mathbf{0}$. In which case Euler's equations of motion reduce to

$$I_1 \dot{\omega}_1 = (I_2 - I_3) \omega_2 \omega_3, \quad (14.5.8)$$

$$I_2 \dot{\omega}_2 = (I_3 - I_1) \omega_3 \omega_1, \quad (14.5.9)$$

$$I_3 \dot{\omega}_3 = (I_1 - I_2) \omega_1 \omega_2. \quad (14.5.10)$$

These apply in the body frame. Since this is noninertial it isn't necessarily true that $\boldsymbol{\omega} = \mathbf{0}$, despite the definition as the frame that rotates with the body, since we define $\boldsymbol{\omega}$ to be the rotation velocity relative to an inertial frame.

14.6 Symmetric Top

The symmetric top is perhaps the simplest example of a rotating body that isn't a sphere. The symmetric top rotates about an axis of symmetry, which we take to be I_3 , and the other two principle axes are equal, so $I_1 = I_2$. In this case Euler's equations of motion in the absence of external torque reduce to

$$I_1 \dot{\omega}_1 = (I_1 - I_3) \omega_2 \omega_3, \quad (14.6.1)$$

$$I_1 \dot{\omega}_2 = (I_3 - I_1) \omega_3 \omega_1, \quad (14.6.2)$$

$$I_3 \dot{\omega}_3 = 0. \quad (14.6.3)$$

Immediately we see from this that ω_3 is constant and we can write

$$\dot{\omega}_1 = \frac{I_1 - I_3}{I_1} \omega_3 \omega_2 = \Omega \omega_2, \quad (14.6.4)$$

$$\dot{\omega}_2 = \frac{I_3 - I_1}{I_1} \omega_3 \omega_1 = -\Omega \omega_1. \quad (14.6.5)$$

Here we define

$$\Omega := \frac{I_1 - I_3}{I_1} \omega_3, \quad (14.6.6)$$

which is constant in time in the body frame.

Differentiating the first of these with respect to time and substituting in the second we get

$$\ddot{\omega}_1 = \Omega \dot{\omega}_2 = -\Omega^2 \omega_1, \quad (14.6.7)$$

and similarly for the second

$$\ddot{\omega}_2 = -\Omega \dot{\omega}_1 = -\Omega^2 \omega_2. \quad (14.6.8)$$

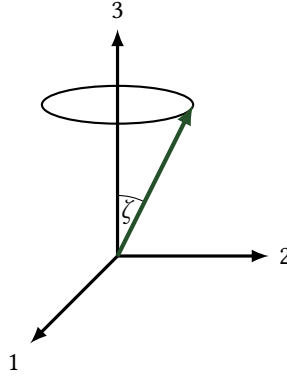


Figure 14.2: A symmetric top rotating around its axis of symmetry will have an angular momentum, ω , which precesses about the axis of symmetry tracing out a circle.

Recognising these as the equations for simple harmonic motion with frequency Ω we identify the relative phase between the two as $\pi/2$ since $\dot{\omega}_2 = -\Omega\omega_1$, so if $\omega_2 = A \cos(\Omega t)$ then $\omega_1 = A \sin(\Omega t) = A \cos(\Omega t + \pi/2)$. This corresponds to circular motion of ω in the $(1, 2)$ -plane, with ω_3 constant. This means that ω precesses about the 3-axis with frequency Ω . The precession of ω_3 traces out what we call the body cone. See Figure 14.2.

Since the motion is in a circle in the $(1, 2)$ -plane the value of $\omega_1^2 + \omega_2^2$ is constant, it's the radius of the circle squared. We also know that ω_3 is constant, and therefore ω_3^2 is constant. Combining these we see that $|\omega| = \sqrt{\omega_1^2 + \omega_2^2 + \omega_3^2}$ is constant. The intrinsic angular momentum, J , then satisfies $J^2 = I_1^2 A^2 + I_3^2 \omega_3^2 = \text{constant}$, where $A = \sqrt{\omega_1^2 + \omega_2^2}$ is the amplitude of the oscillations in the 1 and 2 directions, which is also the radius of the circle. We therefore also have $J \cdot \omega = \text{constant}$.

Since $J \cdot \omega$, $J \cdot e_3$, and $\omega \cdot e_3$ are all constant J , ω , and e_3 must make constant angles to each other. We can show that $(J \times \omega) \cdot e_3 = 0$, which shows that these three vectors are coplanar. This means that J precesses along with ω , since e_3 is fixed and J must remain in the e_3 - ω plane. Therefore J precesses with frequency Ω .

Let ζ be the angle between e_3 and ω , and τ be the angle between e_3 and J . Then

$$\cos \tau = \frac{J \cdot e_3}{|J|} = \frac{I_3 \omega_3}{\sqrt{I_1^2 A^2 + I_3^2 \omega_3^2}} = \frac{\omega_3}{\sqrt{\frac{I_1^2}{I_3^2} A^2 + \omega_3^2}}. \quad (14.6.9)$$

Similarly

$$\cos \zeta = \frac{\omega \cdot e_3}{|\omega|} = \frac{\omega_3}{\sqrt{A^2 + \omega_3^2}}. \quad (14.6.10)$$

From this we see that if $I_1 > I_3$, so $I_1^2/I_3^2 \in (0, 1)$, then $\cos \tau > \cos \zeta$, which means that $\tau < \zeta$, since \cos is a decreasing function on $[0, \pi]$, which is also the total range of angles that J and ω can make to e_3 . The reverse, that if $I_1 < I_3$, then $\tau > \zeta$, also holds.

For the special case where $I_3 = I_1 = I_2$ Euler's equations reduce further to $\dot{\omega} = 0$, which is trivially true since J and ω are parallel in this case.

We wish to translate these results into the lab frame. The vectors \mathbf{J} , $\boldsymbol{\omega}$, and \mathbf{e}_3 remain coplanar. The angular velocity, $\boldsymbol{\omega}$, traces out a cone about \mathbf{J} , called the space cone. We also know that $\boldsymbol{\omega}$ traces out a body cone about \mathbf{e}_3 . We therefore find that the body cone rolls without slipping about the outside of the space cone if $I_3 < I_1$, and inside the space cone if $I_3 > I_1$.

14.7 Asymmetric Top

For the most general case in which all three principle moments of inertia are different the general solution to Euler's equations for torque free motion requires elliptic integrals and doesn't teach us much. Instead we will focus on looking for a solution where $\boldsymbol{\omega}$ is constant in the body frame. From Euler's equations we see that this is only possible if two components of $\boldsymbol{\omega}$ are zero in the body frame. Without loss of generality we take ω_3 to be the only non-zero component of $\boldsymbol{\omega}$ and Euler's equations then give us

$$I_1 \dot{\omega}_1 = (I_2 - I_3) \omega_2 \omega_3 = 0, \quad (14.7.1)$$

$$I_2 \dot{\omega}_2 = (I_3 - I_1) \omega_3 \omega_1 = 0, \quad (14.7.2)$$

$$I_3 \dot{\omega}_3 = (I_1 - I_2) \omega_1 \omega_2 = 0. \quad (14.7.3)$$

So we see that as long as two components of $\boldsymbol{\omega}$ are zero in the body frame then $\boldsymbol{\omega}$ is constant. This means that

an asymmetric body can only rotate with constant angular velocity if the rotation axis is one of the principal axes.

Suppose we have an asymmetric top rotating such that $\boldsymbol{\omega} = \omega \mathbf{e}_3$ is constant. We can investigate the stability by seeing by considering $\omega_1 = \varepsilon_1$, $\omega_2 = \varepsilon_2$, and $\omega_3 = \omega + \varepsilon_3$, where ε_i is small compared to ω . Inserting this in Euler's equations and expanding to first order in ε_i we get

$$I_1 \dot{\varepsilon}_1 = (I_2 - I_3)(\omega + \varepsilon_3)\varepsilon_2 \approx (I_2 - I_3)\omega\varepsilon_2, \quad (14.7.4)$$

$$I_2 \dot{\varepsilon}_2 = (I_3 - I_1)(\omega + \varepsilon_3)\varepsilon_1 \approx (I_3 - I_1)\omega\varepsilon_1, \quad (14.7.5)$$

$$I_3 \dot{\varepsilon}_3 = (I_1 - I_2)\varepsilon_1\varepsilon_2 \approx 0. \quad (14.7.6)$$

Differentiating the first equation, remembering that I_i and ω are constant, then substituting in the second equation we get

$$I_1 \ddot{\varepsilon}_1 = (I_2 - I_3)\omega \dot{\varepsilon}_2 = (I_2 - I_3)\omega \frac{I_3 - I_1}{I_2} \omega \varepsilon_1, \quad (14.7.7)$$

which we rearrange to get

$$I_1 I_2 \ddot{\varepsilon}_1 = (I_2 - I_3)(I_3 - I_1)\omega^2 \varepsilon_1. \quad (14.7.8)$$

If we swap the order of the two equations then we get the same equation but for ε_2 . The general solution to this equation is

$$\varepsilon_1 = A e^{i\Omega t} + B e^{-i\Omega t}, \quad (14.7.9)$$

and similar for ε_2 , where

$$\Omega^2 = \frac{(I_1 - I_3)(I_2 - I_3)}{I_1 I_2} \omega^2. \quad (14.7.10)$$

If I_3 is either the smallest or largest moment of inertia then the right hand side is positive, meaning Ω will be real. If I_3 is the middle axis then Ω is imaginary and hence ε_1 grows exponentially as

$$\varepsilon_1 \sim B e^{|\Omega|t}. \quad (14.7.11)$$

This results in the tennis racket theorem, so called because it is easy to demonstrate by throwing a tennis racket spinning about each of its three axes.

Theorem 14.7.12 — Tennis Racket Theorem. Steady rotation about the intermediate axis is unstable, steady rotation about the major or minor axes is stable.

Fifteen

Lagrange's Method

Lagrange's method for rigid body motion is suitable for systems with a nonzero external torque. The trade-off with Euler's method is that Lagrange's method assumes that there is a point, O , which is fixed. This point is normally not the centre of mass. We will consider the moment of inertia about O , and call it I (we were calling the moment of inertia about a non-centre-of-mass point I' before).

The angular momentum about O is

$$\mathbf{L} = I\boldsymbol{\omega}. \quad (15.0.1)$$

The kinetic energy, including that due to the centre of mass motion, is

$$T = \frac{1}{2}\boldsymbol{\omega} \cdot I\boldsymbol{\omega} = \frac{1}{2}(I_1\omega_1^2 + I_2\omega_2^2 + I_3\omega_3^2) \quad (15.0.2)$$

where I_i are the eigenvalues of I , which we take to be arbitrary.

We need a set of three generalised coordinates to specify the orientation of the body, fixing O fixes three of the six degrees of freedom. The canonical choice is to use **Euler angles**, $(\vartheta, \varphi, \psi)$, which are defined by a series of rotations. There are multiple ways to define Euler angles which don't work together, we choose to use one convention, called the y convention, the only real advantage being that they are slightly more natural for the derivation of conservation laws.

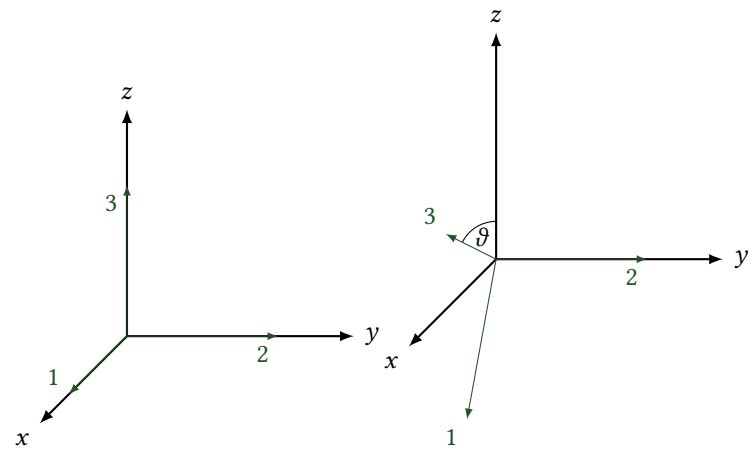
The process for defining Euler angles starts with the object in a reference state and then three rotations are made so that the object has the desired orientation, each rotation defining an angle:

1. Start with the object aligned with the body axes (those that rotate with the body) aligned with the space axes.
2. Rotate the body in the (x, z) -plane, the angle is ϑ .
3. Rotate the body about the z -space-axis defining the angle φ .
4. Rotate the body about the 3-axis (a body axis), defining the angle ψ .

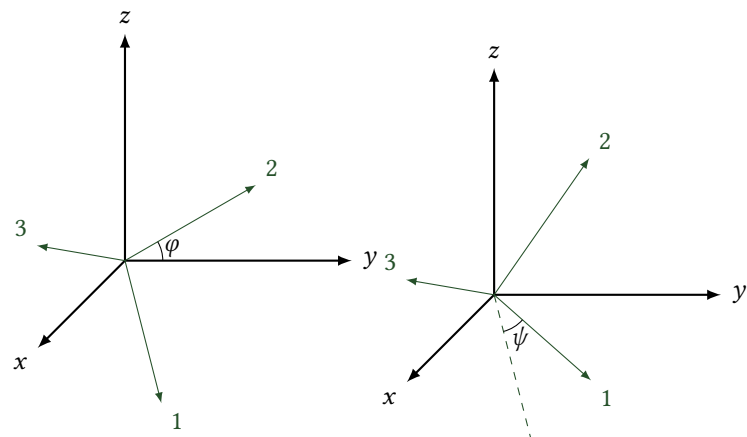
These steps are drawn in [Figure 15.1](#).

15.1 The Lagrangian

We want to find a Lagrangian, \mathcal{L} , where the generalised coordinates are the Euler angles as defined in the previous section. The potential term is straight forward if



(a) Start with the body and space axes aligned. (b) Rotate through angle ϑ about the y -axis.



(c) Rotate through angle φ about the z axis. (d) Rotate through angle ψ about the 3-axis.

Figure 15.1: How to define the rotation of a set of body axes relative to a set of space axes using Euler angles and the y -convention adopted in this course.

we restrict ourselves to tops under gravity, the potential is simply

$$V = Mgl \cos \vartheta \quad (15.1.1)$$

where l is the distance from the fixed point to the centre of mass and ϑ is the Euler angle.

The kinetic energy is not so straight forward. We know that in general

$$T = \frac{1}{2} \boldsymbol{\omega} \cdot I \boldsymbol{\omega} = \frac{1}{2} (I_1 \omega_1^2 + I_2 \omega_2^2 + I_3 \omega_3^2). \quad (15.1.2)$$

We therefore require $\boldsymbol{\omega}$ as a function of $\vartheta, \varphi, \psi, \dot{\vartheta}, \dot{\varphi}$, and $\dot{\psi}$.

Define the axes \mathbf{e}_i in the body frame as usual, and also an intermediate set of axes, \mathbf{e}'_i , which are the axes after the second rotation, before the third rotation. Since the third rotation is about the 3-axis we have that $\mathbf{e}_3 = \mathbf{e}'_3$. Looking down the 3-axis we see that

$$\mathbf{e}_1 \cdot \mathbf{e}'_1 = \cos \psi, \quad (15.1.3)$$

$$\mathbf{e}_1 \cdot \mathbf{e}'_2 = \sin \psi, \quad (15.1.4)$$

$$\mathbf{e}_2 \cdot \mathbf{e}'_1 = -\sin \psi, \quad (15.1.5)$$

$$\mathbf{e}_2 \cdot \mathbf{e}'_2 = \cos \psi, \quad (15.1.6)$$

since this is simply a rotation in the plane normal to \mathbf{e}_3 by angle ψ .

We will first find the components of $\boldsymbol{\omega}$ in the primed basis set. From the definition of the Euler angles we can identify that $\dot{\vartheta}$ is the rate of rotation about the 2'-axis, $\dot{\varphi}$ is the rate of rotation about the z-axis, and $\dot{\psi}$ is the rate of rotation about the 3-axis. Therefore we have

$$\boldsymbol{\omega} = \dot{\vartheta} \mathbf{e}'_2 + \dot{\varphi} \hat{\mathbf{z}} + \dot{\psi} \mathbf{e}'_3. \quad (15.1.7)$$

The 2', 3', and z-axes are not orthogonal, so this doesn't correspond directly to the components of $\boldsymbol{\omega}$ in some orthonormal basis. We can decompose $\hat{\mathbf{z}}$ as

$$\hat{\mathbf{z}} = (\hat{\mathbf{z}} \cdot \mathbf{e}'_3) \mathbf{e}'_3 + (\hat{\mathbf{z}} \cdot \mathbf{e}'_1) \mathbf{e}'_1 \cos \vartheta - \mathbf{e}'_1 \sin \vartheta. \quad (15.1.8)$$

Doing so we get

$$\boldsymbol{\omega} = -\dot{\varphi} \sin \vartheta \mathbf{e}'_1 + \dot{\vartheta} \mathbf{e}'_2 + (\dot{\psi} + \dot{\varphi} \cos \vartheta) \mathbf{e}'_3 = \sum_i \omega'_i \mathbf{e}'_i. \quad (15.1.9)$$

Finally we can move to the unprimed basis using $\mathbf{e}'_3 = \mathbf{e}_3$ and hence $\mathbf{e}'_3 \cdot \mathbf{e}_1 = \mathbf{e}'_3 \cdot \mathbf{e}_2 = 0$ so

$$\omega_1 = \mathbf{e}_1 \cdot \boldsymbol{\omega} = -\dot{\varphi} \sin \vartheta \cos \psi + \dot{\vartheta} \sin \psi, \quad (15.1.10)$$

$$\omega_2 = \mathbf{e}_2 \cdot \boldsymbol{\omega} = \dot{\varphi} \sin \vartheta \sin \psi + \dot{\vartheta} \cos \psi, \quad (15.1.11)$$

$$\omega_3 = \dot{\varphi} \cos \vartheta + \dot{\psi}. \quad (15.1.12)$$

This then gives the kinetic energy as

$$T = \frac{1}{2} (I_1 \omega_1^2 + I_2 \omega_2^2 + I_3 \omega_3^2) \quad (15.1.13)$$

$$= \frac{1}{2} I_1 (-\dot{\varphi} \sin \vartheta \cos \psi + \dot{\vartheta} \sin \psi)^2 \quad (15.1.14)$$

$$+ \frac{1}{2} I_2 (\dot{\varphi} \sin \vartheta \sin \psi + \dot{\vartheta} \cos \psi)^2 \quad (15.1.15)$$

$$+ \frac{1}{2} I_3 (\dot{\varphi} \cos \vartheta + \dot{\psi})^2. \quad (15.1.16)$$

Notice that this is quadratic in the velocities.

This simplifies if the 3-axis, which passes through the fixed pivot point, is also a symmetry axis of the body, which is to say that $I_1 = I_3 = A$, denoting $I_3 = C$ and expanding we then have

$$T = \frac{1}{2} [A(\omega_1^2 + \omega_2^2) + C\omega_3^2] \quad (15.1.17)$$

$$= \frac{1}{2} [A(\omega_1'^2 + \omega_2'^2) + C\omega_3^2] \quad (15.1.18)$$

$$= \frac{1}{2} [A(\dot{\vartheta}^2 + \dot{\varphi}^2 \sin^2 \vartheta) + C(\dot{\varphi} \cos \vartheta + \dot{\psi})^2]. \quad (15.1.19)$$

Here we have used the fact that the 1 and 2 axes differ from the 1' and 2' axes by a rotation in the plane and therefore lengths are left unchanged, in particular $\omega_1^2 + \omega_2^2 = \omega_1'^2 + \omega_2'^2$. This means that the final rotation was not necessary in this case. We could also have argued that the final result must be independent of ψ by symmetry in this case and then evaluated our earlier result for $\psi = 0$ to achieve the same result.

The Lagrangian is then

$$\mathcal{L} = \frac{1}{2} [A(\dot{\vartheta}^2 + \dot{\varphi}^2 \sin^2 \vartheta) + C(\dot{\varphi} \cos \vartheta + \dot{\psi})^2] - Mgl \cos \vartheta. \quad (15.1.20)$$

15.2 Conservation Laws and Equations of Motion

Now that we have the Lagrangian we go through the usual procedure of deriving the conserved quantities and then the equations of motions. First we note that \mathcal{L} doesn't explicitly depend on ψ , and so

$$p_\psi := \frac{\partial \mathcal{L}}{\partial \dot{\psi}} = C(\dot{\psi} + \dot{\varphi} \cos \vartheta) = I_3 \omega_3 \quad (15.2.1)$$

is conserved. The reason for this conservation law is the symmetry of the system under rotations of the body about the 3-axis. Since the centre of mass is on the 3-axis the gravitational torque cannot change ω_3 . By convention we define the conserved quantity

$$n = \omega_3 = \frac{p_\psi}{I_3} = \dot{\psi} + \dot{\varphi} \cos \vartheta \quad (15.2.2)$$

which we call the **spin**.

The Lagrangian is also independent of φ , and so

$$p_\varphi := \frac{\partial \mathcal{L}}{\partial \dot{\varphi}} = A\dot{\varphi} \sin^2 \vartheta + C(\dot{\psi} + \dot{\varphi} \cos \vartheta) \cos \vartheta = A\dot{\varphi} \sin^2 \vartheta + Cn \cos \vartheta \quad (15.2.3)$$

is a conserved quantity. This corresponds to symmetry of the system under rotations about the z-axis. We can identify that $p_\varphi = L_z$, is the z component of the orbital angular momentum.

The Lagrangian *does* depend explicitly on ϑ . The conjugate momentum is

$$p_\vartheta := \frac{\partial \mathcal{L}}{\partial \dot{\vartheta}} = A\dot{\vartheta}. \quad (15.2.4)$$

The Euler–Lagrange equation for ϑ is

$$\dot{p}_\vartheta = \frac{\partial \mathcal{L}}{\partial \vartheta} \quad (15.2.5)$$

which gives

$$A\ddot{\vartheta} = A\dot{\varphi}^2 \sin \vartheta \cos \vartheta - C(\dot{\psi} + \dot{\varphi} \cos \vartheta)\dot{\varphi} \sin \vartheta + Mgl \sin \vartheta. \quad (15.2.6)$$

This can be written as

$$A\ddot{\vartheta} = A\dot{\varphi}^2 \sin \vartheta \cos \vartheta - Cn\dot{\varphi} \sin \vartheta + Mgl \sin \vartheta. \quad (15.2.7)$$

The Lagrangian doesn't depend explicitly on time, hence the energy function,

$$h = \sum_{i=1}^3 \dot{q}_i p_i - \mathcal{L} \quad (15.2.8)$$

is conserved. Since the kinetic energy is quadratic in the velocity the value of the energy function coincides with the total energy $E = T + V$. We can therefore compute $T + V$ instead of h giving

$$E = \frac{A}{2}(\dot{\vartheta}^2 + \dot{\varphi}^2 \sin^2 \vartheta) + \frac{C}{2}n^2 + Mgl \cos \vartheta. \quad (15.2.9)$$

We can write this as

$$\dot{\vartheta}^2 + \dot{\varphi}^2 \sin^2 \vartheta = \frac{2E - Cn^2}{A} - \frac{2Mgl \cos \vartheta}{A}. \quad (15.2.10)$$

Substituting in the orbital angular momentum we can write this as

$$\dot{\vartheta}^2 = \frac{2E - Cn^2}{A} - \frac{2Mgl \cos \vartheta}{A} - \frac{(L_z - Cn \cos \vartheta)^2}{A^2 \sin^2 \vartheta}. \quad (15.2.11)$$

After accounting for conservation laws we have only a single nontrivial degree of freedom, ϑ , described by a nonlinear first order ODE.

15.3 The Zone Function

Since $\dot{\vartheta}$ must be real we require that $\dot{\vartheta}^2 > 0$. For given value of n , L_z , and E this will define a zone of allowed values of the quantity $u := \cos \vartheta$. Writing $\dot{u} = -\dot{\vartheta} \sin \vartheta$ we find that

$$\dot{u}^2 = \left(\frac{2E - Cn^2}{A} \right) (1 - u^2) - \frac{(L_z - Cnu)^2}{A^2} - \left(\frac{2Mgl}{A} \right) u(1 - u^2) = f(u). \quad (15.3.1)$$

This is cubic in u . We will therefore find an interval of values of u for which $f(u)$ is positive and in which the initial value of $f(u)$ lies.

15.4 Steady Precession

Conservation of energy gives us

$$\frac{1}{2}A\dot{\vartheta}^2 = E - \frac{Cn^2}{2} - Mgl \cos \vartheta - \frac{(L_z - Cn \cos \vartheta)^2}{2A \sin^2 \vartheta} = E' - U(\vartheta). \quad (15.4.1)$$

Here we have defined $E' := E - Cn^2/2$, and we define $U(\vartheta)$ to account for the rest of the equation. We can identify this as the energy equation for a particle of 'mass' A in a one dimensional potential, $U(\vartheta)$, with total energy E' . $U(\vartheta)$ becomes very large as ϑ tends to zero, and $U(\vartheta)$ is bounded below.

If $\dot{\vartheta} = 0$ then we call this **steady precession**. This occurs when

$$E' = \min U(L_z, n) \quad (15.4.2)$$

and

$$\vartheta = \min \vartheta(L_z, n). \quad (15.4.3)$$

We then have $\vartheta = \text{constant}$ and since both L_z and n are conserved, and related to the degrees of freedom by

$$L_z = A\dot{\varphi} \sin^2 \vartheta + Cn \cos \vartheta, \quad \text{and} \quad n = \dot{\psi} + \dot{\varphi} \cos \vartheta. \quad (15.4.4)$$

This means that both $\dot{\psi}$ and $\dot{\varphi}$ must also be constant.

We find $\vartheta_0 = \min \vartheta$ by finding where $du/d\vartheta = 0$. This gives

$$\sin \vartheta_0 [\dot{\varphi}^2 A \cos \vartheta_0 - \dot{\varphi} Cn + Mgl] = 0. \quad (15.4.5)$$

We can also find the same result from the Lagrange equation for ϑ by noticing that $\dot{\vartheta} = 0$ implies $\ddot{\vartheta} = 0$.

There are two solutions corresponding to two different physical actions:

- If $\sin \vartheta_0 = 0$ then the top spins upright, we call this a **sleeping top**.
- If $\vartheta_0 \neq 0$, then the term in square brackets above must be zero, which we can solve to give

$$\dot{\varphi} = \frac{Cn \pm \sqrt{C^2 n^2 - 4AMgl \cos \vartheta_0}}{2A \cos \vartheta_0}. \quad (15.4.6)$$

If $\cos \vartheta_0 > 0$, which corresponds to a top *not* hanging below its pivot, so, for example, spinning on a table, then steady precession requires

$$C^2 n^2 \geq 4AMgl \cos \vartheta_0. \quad (15.4.7)$$

This gives a minimum speed for the top to be spinning to have steady precession.

15.5 Fast Top

For a sufficiently fast spinning top we can expand the equation for $\dot{\varphi}$ in steady state precession to give

$$\dot{\varphi} = \frac{Cn \pm (C^2 n^2 - 4AMgl \cos \vartheta_0)^{1/2}}{2A \cos \vartheta_0} \quad (15.5.1)$$

$$\approx \frac{Cn}{2A \cos \vartheta_0} \left[1 \pm \left(1 - \frac{2AMgl \cos \vartheta_0}{C^2 n^2} \right) \right]. \quad (15.5.2)$$

There are two possible roots to this equation:

– Slow precession due to gravity, which has

$$\dot{\vartheta} = \frac{Mgl}{Cn}. \quad (15.5.3)$$

+ Free precession where

$$\dot{\varphi} = \frac{Cn}{A \cos \vartheta_0}. \quad (15.5.4)$$

These two limiting cases can also be checked by other methods. For example, for slow precession we can balance the gravitational torque such that \dot{L} is constant. For free precession we can obtain the same result from Euler's equations of motion with zero torque.

For a real gyroscope the free precession is damped by air resistance and the slow precession dominates.

15.6 Nutation

Keeping L_z fixed and increasing the energy, E' , such that E' is slightly above $\min U$ we get oscillations on top of the steady precession, we call this **nutation**. The path traced by the 3-axis will trace out either a wave or loops on the surface of a sphere while staying within a zone of values of ϑ . The path depends on the relative frequencies of the φ and ϑ oscillations. For small amplitudes there is simple harmonic motion in ϑ with frequency

$$\Omega_{\text{nut}} = \sqrt{\frac{U''(\vartheta_0)}{A}}. \quad (15.6.1)$$

For large enough spin,

$$Cn \gg \sqrt{4AMgl \cos \vartheta_0} \quad (15.6.2)$$

we have $\Omega_{\text{nut}} \gg \dot{\varphi}$ and there are many oscillations per precession period.

15.7 Sleeping Top

The sleeping top is the case of $\vartheta_0 = 0$. Lagrange's equation of motion is

$$A\ddot{\vartheta} = A\dot{\varphi}^2 \sin \vartheta \cos \vartheta - Cn\dot{\varphi} \sin \vartheta + Mgl \sin \vartheta. \quad (15.7.1)$$

Expanding this for small ϑ we get

$$A\ddot{\vartheta} = (A\dot{\varphi}^2 - Cn\dot{\varphi} + Mgl)\vartheta. \quad (15.7.2)$$

One solution to this is $\vartheta = 0$.

The top starts of upright and $L_z = Cn$. Since both L_z and Cn are conserved this means that

$$Cn = L_z = A\dot{\varphi} \sin^2 \vartheta + Cn \cos \vartheta. \quad (15.7.3)$$

Expanding this for small ϑ we get

$$\frac{1}{2}Cn\vartheta^2 = A\dot{\varphi}\vartheta^2. \quad (15.7.4)$$

The sleeping top then has

$$\dot{\vartheta} = \frac{Cn}{2A}. \quad (15.7.5)$$

This is the frequency at which a sleeping top precesses when perturbed by an infinitesimal distance from the vertical. We then have

$$A\ddot{\vartheta} = \left(Mgl - \frac{C^2 n^2}{4A} \right) \vartheta. \quad (15.7.6)$$

If $n^2 > 4MglA/C^2$ then we have simple harmonic motion and the motion of the sleeping top is stable. If the spin is too low however then any perturbation will grow exponentially and so the top is unstable.

We see this in real life when spinning a top. It will spin upright for a long time slowly decreasing in spin speed due to friction and at some point will rapidly start to veer away from upright before falling over.

Part IV

Oscillations and Lagrange Multipliers

Sixteen

Oscillations

16.1 Static Equilibrium

Consider a system of D degrees of freedom in a time-independent conservative potential, that is

$$V = V(\{q\}, \{\dot{q}\}, t) = V(\{q\}), \quad \text{with } \{q\} = (q_1, \dots, q_D). \quad (16.1.1)$$

Static equilibrium, that is no motion, is only possible if all of the generalised forces, Q_i , vanish together for some set of coordinates, $\{q\}_{\text{eq}}$. This set of coordinates defines an equilibrium position.

We can redefine our coordinates, $\{q\} \rightarrow \{q\} - \{q\}_{\text{eq}}$, such that the equilibrium position is the origin. We then have

$$Q_j(\{0\}) = - \left(\frac{\partial V(\{q\})}{\partial q_j} \right)_{\{q\}=0} = 0 \quad (16.1.2)$$

for all j . Expanding V about this point terms linear in $\{q\}$ vanish and we are left with

$$V(\{q\}) = V(\{0\}) + \frac{1}{2} \sum_i \sum_j \frac{\partial^2 V}{\partial q_i \partial q_j} q_i q_j + O(\{q\}^3). \quad (16.1.3)$$

The first term here is just a constant, which we will neglect. We then write

$$V(\{q\}) = \frac{1}{2} q_i V_{ij} q_j \quad (16.1.4)$$

where we now use the summation convention and we define

$$V_{ij} := \left(\frac{\partial^2 V}{\partial q_i \partial q_j} \right)_{\{q\}=0}. \quad (16.1.5)$$

This is a constant $D \times D$ matrix of second derivatives evaluated at the origin. This matrix is symmetric (for sufficiently nice V). If $\{q\} = 0$ is a local minimum then V_{ij} will be positive semi-definite, that is the eigenvalues will all be nonnegative, this ensures that we cannot decrease the potential energy by changing the value of any of $\{q\}$, and hence it is indeed a minimum.

16.2 The Lagrangian

We will consider only systems where the kinetic energy is time independent and quadratic in the velocities. We then have

$$T = T(\{q\}, \{\dot{q}\}, t) = \frac{1}{2} \dot{q}_i T_{ij} \dot{q}_j. \quad (16.2.1)$$

Here T_{ij} is a matrix depending, in general, on the coordinates, $\{q\}$.

For small displacements about the equilibrium we can take T_{ij} to be an approximately constant matrix. It is convenient to choose T_{ij} to be symmetric. This is possible since we can always symmetrise any asymmetric matrix, \tilde{T}_{ij} , by writing $(\tilde{T}_{ij} + \tilde{T}_{ji})/2$, without changing the value of T . Kinetic energies are typically positive and so T_{ij} is positive definite, that is all its eigenvalues are positive.

For small displacements the Lagrangian is

$$\mathcal{L} = T - V = \frac{1}{2} (\dot{q}_i T_{ij} \dot{q}_j - q_i V_{ij} q_j). \quad (16.2.2)$$

Lagrange's equations then give

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) = \frac{\partial \mathcal{L}}{\partial q_i} \implies T_{ij} \ddot{q}_j = -V_{ij} q_j. \quad (16.2.3)$$

These are linear equations, so we say that the system has been linearised by the assumption that the displacements are small. We can then solve the resulting equations with matrix methods.

We define **normal modes** to be the solutions to the equations of motion in which the generalised coordinates all oscillate with the same angular frequency, ω :

$$q_j = a_j \cos(\omega t + \varepsilon_j). \quad (16.2.4)$$

The equation of motion is then

$$\omega^2 T_{ij} a_j + V_{ij} a_j = 0. \quad (16.2.5)$$

Setting $\lambda := \omega^2$ we get

$$(V_{ij} - \lambda T_{ij}) a_j = 0. \quad (16.2.6)$$

This is a sort of generalised eigenvalue equation, the basic version being $(M_{ij} - \lambda \delta_{ij}) a_j = 0$. The process of solving this equation is very similar to the eigenvalue equation:

1. Find the eigenvalues, $\lambda^{(i)}$, by solving

$$\det(V - \lambda T) = 0. \quad (16.2.7)$$

2. For a given $\lambda^{(k)}$ write

$$(V_{ij} - \lambda^{(k)} T_{ij}) e_j^{(k)} = 0. \quad (16.2.8)$$

This defines an (unnormalised) eigenvector, $e^{(k)}$, for the k th mode which has j th component $e_j^{(k)}$. In general these eigenvectors will not be orthogonal.

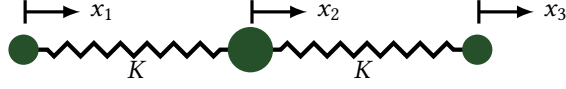


Figure 16.1: Triatomic molecule.

16.3 Triatomic Molecule

Consider a triatomic molecule, such as H_2O or CO_2 , with one central atom of mass M and two outer atoms of mass m . We can model the bonds as being springs with spring constant K . See Figure 16.1.

We define the coordinates x_i as the displacements of the atoms, as shown in Figure 16.1. The kinetic energy is then

$$T = \frac{1}{2}m(\dot{x}_1^2 + \dot{x}_3^2) + \frac{1}{2}M\dot{x}_2^2. \quad (16.3.1)$$

The potential is simply the energy stored in the springs, which is related to the extension of the springs, $x_1 - x_2$ and $x_3 - x_2$:

$$V = \frac{1}{2}K(x_1 - x_2)^2 + \frac{1}{2}K(x_3 - x_2)^2 = \frac{1}{2}K(x_1^2 + 2x_2^2 + x_3^2 - 2x_1x_2 - 2x_3x_2). \quad (16.3.2)$$

Choosing x_i as our generalised coordinates we find the matrices for T and V are

$$T_{ij} = \begin{pmatrix} m & 0 & 0 \\ 0 & M & 0 \\ 0 & 0 & m \end{pmatrix}, \quad \text{and} \quad V_{ij} = \begin{pmatrix} K & -K & 0 \\ -K & 2K & -K \\ 0 & -K & K \end{pmatrix}. \quad (16.3.3)$$

Note that we have chosen V_{ij} to be symmetric, there are other nonsymmetric matrices that would give the same result, for example

$$\begin{pmatrix} K & -2K & 0 \\ 0 & 2K & 0 \\ 0 & -2K & K \end{pmatrix}. \quad (16.3.4)$$

The generalised eigenvalue problem is then to solve

$$\begin{vmatrix} K - \lambda m & -K & 0 \\ -K & 2K - \lambda M & -K \\ 0 & -K & K - \lambda m \end{vmatrix} = 0. \quad (16.3.5)$$

This gives us the characteristic polynomial

$$\lambda(K - \lambda m)(\lambda m M - (M + 2m)K) = 0. \quad (16.3.6)$$

This has roots

$$\lambda^{(1)} = 0, \quad \lambda^{(2)} = \frac{K}{m}, \quad \text{and} \quad \lambda^{(3)} = \frac{K(2m + M)}{Mm}. \quad (16.3.7)$$

To find the eigenvectors, $\mathbf{e}^{(k)}$ we solve

$$(V_{ij} - \lambda^{(k)} T_{ij})e_j^{(k)} = 0. \quad (16.3.8)$$

We find that

$$\mathbf{e}^{(1)} = (1, 1, 1), \quad \mathbf{e}^{(2)} = (1, 0, -1), \quad \text{and} \quad \mathbf{e}^{(3)} = (1, -2m/M, 1). \quad (16.3.9)$$

Writing $\omega_i = \sqrt{\lambda^{(i)}}$ we can interpret these three eigenvectors as describing three linearly independent modes of vibration:

1. The first is uniform motion of the whole molecule, all three atoms move the same amount and the oscillation frequency is $\omega_1 = 0$, since there is no restoring force as the molecule is free to move.
2. The second is oscillation at frequency $\omega_2 = \sqrt{K/m}$, which corresponds to the two outer atoms moving π out of phase with the middle atom stationary.
3. The third is oscillation at frequency $\omega_3 = \sqrt{K(2m+M)/(mM)}$, which corresponds to the two outer atoms moving in the same direction and the central atom moving in the opposite direction.

The general solution describing the motion is then a linear combination of these three distinct types of motion:

$$\mathbf{x}(t) = A_1 \mathbf{e}^{(1)} \cos(\omega_1 t + \varphi_1) + A_2 \mathbf{e}^{(2)} \cos(\omega_2 t + \varphi_2) + A_3 \mathbf{e}^{(3)} \cos(\omega_3 t + \varphi_3). \quad (16.3.10)$$

The amplitudes, A_i , and phases, φ_i , are set by the initial conditions, giving 6 constants having solved 3 second order differential equations.

In this case since $\omega_1 = 0$ the cosine factor is just a constant. In addition A_1 can be arbitrarily large. Taking a limit we find that

$$\lim_{\omega_1 \rightarrow 0} A_1 \cos(\omega_1 t + \varphi_1) = A + vt \quad (16.3.11)$$

for some constants A and v . This means we should really write the solution as

$$\mathbf{x}(t) = (A + vt) \mathbf{e}^{(1)} + A_2 \mathbf{e}^{(2)} \cos(\omega_2 t + \varphi_2) + A_3 \mathbf{e}^{(3)} \cos(\omega_3 t + \varphi_3). \quad (16.3.12)$$

To avoid this we usually neglect zero modes where the whole system moves together by simply restricting V_{ij} to be positive definite and treat the motion of the whole system separately.

16.4 Simultaneous Diagonalisation

The matrix T is positive definite. We will show that there is a coordinate transformation such that in the transformed coordinates both T and V , which are symmetric, are simultaneously diagonalised. We start by applying an orthogonal transformation, C , to T and V such that T is diagonalised, defining

$$T' := C^T T C, \quad \text{and} \quad V' := C^T V C. \quad (16.4.1)$$

This is always possible since T is symmetric. We now have

$$T' = \begin{pmatrix} T_1 & & & \\ & T_2 & & \\ & & \ddots & \\ & & & T_D \end{pmatrix} \quad (16.4.2)$$

Since T is positive definite and T_i are the eigenvalues of T all of T_i are positive, and hence $\sqrt{T_i}$ are real. This means we can define the real matrix

$$K := \begin{pmatrix} 1/\sqrt{T_1} & & & \\ & 1/\sqrt{T_2} & & \\ & & \ddots & \\ & & & 1/\sqrt{T_D} \end{pmatrix} \quad (16.4.3)$$

We can then apply a transform with this matrix defining

$$T'' = K^T C^T T C K, \quad \text{and} \quad V'' = K^T C^T V C K. \quad (16.4.4)$$

The way we have defined K means that $T'' = K^T T' K = K T' K = I$. Also V'' is still symmetric.

We then apply a final orthogonal transformation, D , to diagonalise V'' , which is possible since it is symmetric. This has no effect on T'' since $T'' = I$ so $D^T T'' D = D^T I D = D^T D = D^{-1} D = I$, since D is orthogonal. We therefore define $T''' = D T'' D = I$ and $V''' = D^T V'' D$ which gives a diagonal V''' :

$$V''' = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_D \end{pmatrix}. \quad (16.4.5)$$

Defining $M := C K D$ we have

$$T''' = M^T T M = I, \quad \text{and} \quad V''' = M^T V M. \quad (16.4.6)$$

We can now interpret this as a coordinate transformation:

$$q_i \rightarrow \tilde{q}_i, \quad \text{with} \quad q_i = M_{ij} \tilde{q}_j. \quad (16.4.7)$$

This is invertible and so $\{\tilde{q}\}$ are a valid set of generalised coordinates. In general this is *not* an orthogonal transformation, unless $T_i = 1$ for all i .

Dropping the tilde and working exclusively in this simultaneously diagonalised basis we have $T_{ij} = \delta_{ij}$ and $V_{ij} = \lambda_i \delta_{ij}$. The Lagrangian is then

$$\mathcal{L} = \frac{1}{2} \dot{q}_i T_{ij} \dot{q}_j - \frac{1}{2} q_i V_{ij} q_j = \sum_i \left[\frac{1}{2} \dot{q}_i^2 - \frac{1}{2} \lambda_i q_i^2 \right]. \quad (16.4.8)$$

We now stop using the summation convention. This form of the Lagrangian is called the **normal form** and the transformed coordinates the **normal coordinates**. The Euler–Lagrange equations are now very simple:

$$\ddot{q}_i = -\lambda_i q_i. \quad (16.4.9)$$

Which is simple harmonic motion with the general solution

$$q_i(t) = A_i \cos(\omega_i t + \varphi_i). \quad (16.4.10)$$

Where as usual $\omega_i = \sqrt{\lambda_i}$, and A_i and φ_i are constants of integration fixed by the initial conditions.

We don't need to sum over all of the modes any more. Each coordinate executes simple harmonic motion independent of the other coordinates. The effort in solving the motion is moved to finding the transformation that diagonalises both matrices simultaneously which allows us to relate the transformed coordinates, which may not have an obvious physical meaning, to a set of more easily understandable generalised coordinates.

When solving problems with this method we don't typically find the matrices C , K , and D separately, instead we identify that $M_{ij} = e_i^{(j)}$. This shows that the two methods of solving the equations are almost identical since both involve solving the generalised eigenvalue equation $(V_{ij} - \lambda T_{ij})a_j = 0$, and finding the eigenvectors, $\mathbf{e}^{(k)}$. The only difference is in how we separate out the individual components of the motion.

Seventeen

Lagrange Multipliers

Recall that a holonomic constraint is one which can be written in the form $f(\{q\}, t) = 0$. So far we have been accounting for these by using them to reduce the number of degrees of freedom, and hence the number of generalised coordinates in a problem. There is another method in which we can include them directly in the Lagrangian. This is the method of Lagrange multipliers. This method can also be used in many other optimisation procedures involving minimisation of functionals.

Consider a system of two coordinates, x and y , which are related by the holonomic constraint $f(x, y) = 0$ for some function f . We construct the Lagrangian, $\mathcal{L} = T - V$, as normal using both coordinates treating them as independent. We find that the change in action for a change in the coordinates, $x \rightarrow x + \delta x$ and $y \rightarrow y + \delta y$ is

$$\delta S = \int \left(\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}} \right) - \frac{\partial \mathcal{L}}{\partial x} \right) \delta x + \left(\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{y}} \right) - \frac{\partial \mathcal{L}}{\partial y} \right) \delta y dt. \quad (17.0.1)$$

If x and y are independent coordinates then both brackets would have to vanish independently. However, we are assuming that they aren't independent.

Now consider the Lagrangian

$$\mathcal{L}'(x, y, t) = \mathcal{L}(x, y) + \lambda f(x, y). \quad (17.0.2)$$

Note that this has the same numeric value, since $f(x, y) = 0$, but the functional form is different, which we know can be important for functional calculus. The change in the action can then be written as

$$\delta S' = \int \left(\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}} \right) - \frac{\partial \mathcal{L}}{\partial x} + \lambda \frac{\partial f}{\partial x} \right) \delta x + \left(\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{y}} \right) - \frac{\partial \mathcal{L}}{\partial y} + \lambda \frac{\partial f}{\partial y} \right) \delta y dt. \quad (17.0.3)$$

Since λ is arbitrary we can choose it such that each term must be zero.

In general if we have n (not necessarily independent) coordinates and m constraints, of the form

$$f_\alpha(q_1, \dots, q_n, t) = 0 \quad (17.0.4)$$

for $\alpha = 1, \dots, m$, then we define the modified Lagrangian

$$\mathcal{L}'(q_1, \dots, q_n, \lambda_1, \dots, \lambda_m) = \mathcal{L}(q_1, \dots, q_n) + \sum_{\alpha} \lambda_{\alpha} f_{\alpha}(q_1, \dots, q_n, t). \quad (17.0.5)$$

We then take all of q_i and λ_{α} as our coordinates and form the Euler–Lagrange equations for them. Since there is no $\dot{\lambda}_{\alpha}$ dependence in \mathcal{L}' the Euler–Lagrange equations for λ_{α} immediately return the constraint $f_{\alpha}(\{q\}) = 0$.

17.1 Holonomic Example

Consider Atwood's machine from [Section 9.2](#). Previously we treated the constraint by reducing the number of degrees of freedom. Now we shall do so with Lagrange multipliers. Using the notation of [Section 9.2](#) we know that the kinetic energy and potential are

$$T = \frac{1}{2}M\dot{x}^2 + \frac{1}{2}m\dot{y}^2, \quad \text{and} \quad V = Mgx + mgy. \quad (17.1.1)$$

Our holonomic constraint is $x + y + \ell = 2h$, which when written in the required form is $f(x, y) = x + y + \ell - 2h = 0$.

The Lagrangian accounting for this constraint is

$$\mathcal{L} = \frac{1}{2}M\dot{x}^2 + \frac{1}{2}m\dot{y}^2 - Mgx - mgy + \lambda(x + y + \ell - 2h). \quad (17.1.2)$$

We then take x , y , and λ as our coordinates. The Euler–Lagrange equations are then

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}} \right) = \frac{d}{dt} (M\dot{x}) = M\ddot{x} = \frac{\partial \mathcal{L}}{\partial x} = -Mg + \lambda, \quad (17.1.3)$$

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{y}} \right) = \frac{d}{dt} (m\dot{y}) = m\ddot{y} = \frac{\partial \mathcal{L}}{\partial y} = -mg + \lambda, \quad (17.1.4)$$

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{\lambda}} \right) = 0 = \frac{\partial \mathcal{L}}{\partial \lambda} = x + y + \ell - 2h. \quad (17.1.5)$$

Notice that we cannot drop the constant terms in order for this method to work, since we need them to ensure that we get the full constraint back. Differentiating the last of these equations twice, which is just the constraint, we get $\ddot{x} = -\ddot{y}$. We can subtract the second from the first to eliminate λ giving

$$M\ddot{x} + Mg = -m\ddot{x} + mg. \quad (17.1.6)$$

This gives the result

$$\ddot{x} = -\frac{(M - m)g}{M + m}. \quad (17.1.7)$$

This is the same solution we found using the other method.

Notice that if we missed the constraint we would have found $\ddot{x} = \ddot{y} = -g$, which corresponds to both masses being free. This pattern, of missing constraints giving the free case, is fairly common.

17.2 Non-Holonomic Example

Consider a hanging chain of length ℓ and density ρ hanging between two points which are the same height and are a distance $2a$ apart, with $2a < D$, the length of the chain. The chain will hang in the shape which minimises the potential energy, which is the same as maximising the action, since we are looking for a stationary solution so the kinetic energy is zero.

The total energy is given by integrating along the chain:

$$V = \rho g \int y \, ds = \rho g \int_{-a}^a y \sqrt{1 + y'^2} \, dx. \quad (17.2.1)$$

Here we take $x = 0$ as the midpoint between the two attachment points.

The constraint that we need to introduce is the constant length of the chain. This can be written in integral form as

$$\int ds = \int \sqrt{1 + y'^2} \, dx = D \quad (17.2.2)$$

This gives us the functional

$$I = \int_{-a}^a \left[\rho g y \sqrt{1 + y'^2} + \lambda (\sqrt{1 + y'^2} - D) \right] dx. \quad (17.2.3)$$

Since there is no explicit dependence on x we know that the following is constant:

$$y' \frac{\partial F}{\partial y'} - F = -\frac{\rho g y + \lambda}{\sqrt{1 + y'^2}} = C. \quad (17.2.4)$$

The solution to this is

$$y(x) = \frac{C}{\rho g} \cosh \left(\frac{\rho g x}{C} \right) - \frac{\lambda}{\rho g}. \quad (17.2.5)$$

Requiring $y(\pm a) = 0$ we get

$$\lambda = C \cosh \left(\frac{\rho g a}{C} \right). \quad (17.2.6)$$

We then find C by requiring that the total length be D , which gives

$$D = \int_{-a}^a \sqrt{1 + y'^2} \, dx = \int_{-a}^a \cosh \left(\frac{\rho g x}{C} \right) dx, \quad (17.2.7)$$

here we have used

$$\frac{d}{dx} \cosh x = \sinh x, \quad \text{and} \quad 1 + \sinh^2 x = \cosh^2 x. \quad (17.2.8)$$

Appendices

A

Differential Equations

Many of the equations of motion we will derive are second order linear ordinary differential equations of the form

$$ay''(x) + by'(x) + cy(x) = f(x) \quad (\text{A.0.1})$$

for some constants a , b , and c , and functions y and f . There is a general procedure for solving such equations.

1. First define the homogenous equation

$$ay''(x) + by'(x) + cy(x) = 0. \quad (\text{A.0.2})$$

This is important since if y_1 is a solution to the homogeneous equation and y_p is a solution to the inhomogeneous equation then $y_p + \lambda y_1$ with λ being a constant, is a solution to the original inhomogeneous equation also. The most general solution is the solution to the inhomogeneous equation and a linear combination of solutions to the homogenous equation. The number of solutions to the homogenous equation is equal to the order of the equation, so for our purposes 2.

2. Solve the homogenous equation. This is usually done by making a guess and substituting into the homogenous equation and solving for any unknowns in the guess. The most common guess is of the form e^{mx} . If we make this guess after substitution we have

$$(am^2 + bm + c)e^{mx} = 0. \quad (\text{A.0.3})$$

Since $e^{mx} \neq 0$ we must therefore have that the quadratic in m is zero. This gives us two values for m , m_1 and m_2 . If $m_1 \neq m_2$ then we define

$$y_1(x) = e^{m_1 x}, \quad \text{and} \quad y_2(x) = e^{m_2 x}. \quad (\text{A.0.4})$$

If $m_1 = m_2$ then instead define

$$y_1(x) = e^{m_1 x}, \quad \text{and} \quad y_2(x) = xe^{m_1 x}. \quad (\text{A.0.5})$$

Occasionally other guesses, such as trig functions, make more sense, particularly if there are no first derivatives in the equation.

Table A.1: Common guesses to solve the inhomogeneous differential equation $ay''(x) + by'(x) + cy(x) = f(x)$ for different functions f .

Form of f	Suggestion for y_p
Constant, k	Constant k/c
Polynomial in x	Polynomial in x
$ke^{\alpha x}$	$Ae^{\alpha x}$
$\sin(\alpha x)$ and/or $\cos(\alpha x)$	$A \sin(\alpha x) + B \cos(\alpha x)$

3. Returning to the inhomogeneous equation we now need to find a solution, y_p , of course if the original equation was inhomogeneous then this step can be skipped. We again do this by guessing. The exact guess depends on the form of f but some common choices are given in [Table A.1](#). Once you have chosen a function substitute it in and see if it is valid.

4. If you have no boundary conditions then you are done. The solution is

$$y(x) = y_p(x) + \sum_i \lambda_i y_i(x) \quad (\text{A.0.6})$$

where y_i are solutions to the homogeneous equation. If you have boundary conditions then substitute them into this equation and solve for λ_i .

■ **Example A.0.7** Suppose we wish to solve

$$y''(x) + 4y'(x) + 4y(x) = \sin(3x), \quad (\text{A.0.8})$$

with the boundary conditions $y(0) = 10$ and $y'(0) = 0$.
We start by defining the homogeneous equation

$$y''(x) + 4y'(x) + 4y(x) = 0. \quad (\text{A.0.9})$$

We propose that e^{mx} is a solution to this. Substituting this in we find

$$(m^2 + 4m^2 + 4)e^{mx} = 0. \quad (\text{A.0.10})$$

The quadratic has one repeated root $m = -2$ and so our solutions to the homogenous equation are

$$y_1(x) = e^{-2x}, \quad \text{and} \quad y_2(x) = xe^{-2x}. \quad (\text{A.0.11})$$

In this case $f(x) = \sin(3x)$ and so we try a solution of the form

$$y_p(x) = A \sin(3x) + B \cos(3x). \quad (\text{A.0.12})$$

Substituting this into our original equation we have

$$\begin{aligned} & -9A \sin(3x) - 9B \cos(3x) + 12A \cos(3x) - 12B \sin(3x) \\ & + 4A \sin(3x) + 4B \cos(3x) \\ & = -(5A + 12B) \sin(3x) + (12A - 5B) \cos(3x). \end{aligned} \quad (\text{A.0.13})$$

In order for this to be equal to $\sin(3x)$ we require that

$$5A + 12B = -1, \quad \text{and} \quad 12A - 5B = 0. \quad (\text{A.0.14})$$

Solving these simultaneous equations gives us

$$A = -\frac{5}{169}, \quad \text{and} \quad B = -\frac{12}{169}. \quad (\text{A.0.15})$$

Hence our solution is

$$y_p(x) = -\frac{5}{169} \sin(3x) - \frac{12}{169} \cos(3x). \quad (\text{A.0.16})$$

The general solution is then

$$y(x) = -\frac{5}{169} \sin(3x) - \frac{12}{169} \cos(3x) + \lambda_1 e^{-2x} + \lambda_2 x e^{-2x}. \quad (\text{A.0.17})$$

Substituting for our first initial condition, $y(0) = 10$, we have

$$y(0) = -\frac{12}{169} + \lambda_1 = 10 \implies \lambda_1 = \frac{1702}{169}. \quad (\text{A.0.18})$$

The second initial condition, $y'(0) = 0$, gives us


$$y'(0) = -\frac{15}{169} - 2\frac{1702}{169} - 2\lambda_2 = 0 \implies \lambda_2 = -\frac{263}{26}. \quad (\text{A.0.19})$$

Hence our final solution is

$$y(x) = -\frac{5}{169} \sin(3x) - \frac{12}{169} \cos(3x) + \frac{1702}{169} e^{-2x} - \frac{263}{26} x e^{-2x}. \quad (\text{A.0.20})$$

B

Electromagnetism

 For more details see the notes for electromagnetism.

B.1 Maxwell's Equations

Maxwell's equations are, in SI units

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}, \quad (\text{M1})$$

$$\nabla \cdot \mathbf{B} = 0, \quad (\text{M2})$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \quad (\text{M3})$$

$$\nabla \times \mathbf{B} = \mu_0 \left(\mathbf{j} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right). \quad (\text{M4})$$

Here \mathbf{E} and \mathbf{B} are the electric and magnetic fields, ρ is the electric charge density, \mathbf{j} is the electric current density, ϵ_0 and μ_0 are the permittivity and permeability of free space, also known as the electric and magnetic constants. The value of ϵ_0 and μ_0 , and the speed of light, $c = 1/\sqrt{\epsilon_0\mu_0}$, determine the system of units used to measure \mathbf{E} , \mathbf{B} , ρ , and \mathbf{j} .

The force on a particle of charge e in an electromagnetic field is given by

$$\mathbf{F} = e(\mathbf{E} + \dot{\mathbf{r}} \times \mathbf{B}) \quad (\text{B.1.1})$$

where \mathbf{r} is the position of the particle. This force is velocity dependent, and therefore cannot be treated with a simple $T - V$ Lagrangian (unless $\mathbf{B} = \mathbf{0}$).

B.2 Potentials

It is a well known fact of vector calculus that a divergence free vector field, such as \mathbf{B} , can be written as the curl of a vector field:

$$\nabla \cdot \mathbf{B} = 0 \iff \exists \mathbf{A} \text{ such that } \mathbf{B} = \nabla \times \mathbf{A}. \quad (\text{B.2.1})$$

In the case of \mathbf{B} being the magnetic field we call \mathbf{A} the magnetic vector potential.

Writing M3 in terms of the vector potential we have

$$\nabla \times \mathbf{E} = -\frac{\partial}{\partial t}(\nabla \times \mathbf{A}) \implies \nabla \times \left(\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} \right) = 0. \quad (\text{B.2.2})$$

A standard result of vector calculus is that any curl free field, \mathbf{G} , such that $\nabla \times \mathbf{G} = 0$, can be written as the gradient of a scalar potential, φ :

$$\nabla \times \mathbf{G} = 0 \iff \exists \varphi \text{ such that } \mathbf{G} = -\nabla \varphi \quad (\text{B.2.3})$$

where the negative sign is a convention. From [Equation \(B.2.2\)](#) we see that

$$\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} \quad (\text{B.2.4})$$

is curl free, and so there exists a scalar field, φ , which we call the scalar electric potential, such that

$$\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} = -\nabla \varphi \implies \mathbf{E} = -\nabla \varphi - \frac{\partial \mathbf{A}}{\partial t}. \quad (\text{B.2.5})$$

We can use the potential to rewrite Maxwell's equations as

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}, \quad (\text{B.2.6})$$

$$\mathbf{B} = \nabla \times \mathbf{A}, \quad (\text{B.2.7})$$

$$\mathbf{E} = -\nabla \varphi - \frac{\partial \mathbf{A}}{\partial t}, \quad (\text{B.2.8})$$

$$\nabla \times \mathbf{B} = \mu_0 \left(\mathbf{j} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right). \quad (\text{B.2.9})$$

Bibliography

- [1] Matt Parker, *The bubble that breaks maths*. 05/11/2021 <https://www.youtube.com/watch?v=310m4VrSzb8> (accessed 08/11/2021)

Acronyms

C

CoM: centre of momentum [18](#)

E

EoM: equation of motion [25](#)

N

N1: Newton's first law [3](#)

N2: Newton's second law [4](#)

N3: Newton's third law [4](#)

P

PA: principle axes [78](#)

Index

A

acceleration, 3
action, 50
Atwood's machine, 36
Atwood's monkey, 37

B

body frame, 78

C

calculus of variations, 43
cancellation of dots, 28
canonical momentum, *see* generalised momentum
canonical quantisation, 72
central force, 8
central potential, 8
centre of mass, 14
centre of mass frame, 18
centre of momentum frame, 18
centrifugal force, 20
commutator, 72
commuting derivatives rule, 28
conservative force, 6
constant of motion, 56
Coriolis force, 20
cyclic coordinate, 56

D

degree of freedom, 21

E

effective potential, 10
energy function, 56
equation of motion, 25
equivalent Lagrangians, 53
Euler's equation, 44
Euler–Lagrange equation, 44

Euler–Lagrange equations, 33

F

fictitious force, 20
first integral, 42
functional, 43

G

Galilean transformation, 17
generalised coordinate, 25
generalised displacement, 26
generalised force, 26
generalised momentum, 55
generalised velocity, 26
global symmetry, 59

H

Hamilton's equations, 70
Hamilton's principle, 50
Hamiltonian, 69
homogeneous, 59

I

ignorable coordinate, 56
inertia tensor, 77
inertial frame, 4
integrable, 42
intrinsic angular momentum, 18
isotropic, 60

K

kinetic energy, 6

L

Lagrange's equations of motion, 31
Lorentz force, 64

M

moment of force, 5

momentum
 angular, 5
 linear, 3

N

Newton's first law, 3
Newton's second law, 4
Newton's third law, 4
Noether's theorem, 42
nonholonomic constraints, 23

P

path integral, 73
Poisson bracket, 71
principle axes, 78
principle moments of inertia, 78

R

rigid body, 21
rolling conditions, 22

S

simple pendulum, 39
spherical pendulum, 40

T

torque, 5
total energy, 7

V

velocity, 3
virtual displacement, 26

W

work done, 5