Preprint typeset in JHEP style - HYPER VERSION

Michaelmas Term, 2004 and 2005

Classical Dynamics

University of Cambridge Part II Mathematical Tripos

Dr David Tong

Department of Applied Mathematics and Theoretical Physics,
Centre for Mathematical Sciences,
Wilberforce Road,
Cambridge, CB3 OBA, UK

http://www.damtp.cam.ac.uk/user/tong/dynamics.html
d.tong@damtp.cam.ac.uk

Recommended Books and Resources

This is a second course on classical mechanics. It assumes knowledge of the earlier lecture notes on Dynamics and Relativity.

- L. Hand and J. Finch, *Analytical Mechanics*

This very readable book covers everything in the course at the right level. It is similar to Goldstein's book in its approach but with clearer explanations, albeit at the expense of less content.

There are also three classic texts on the subject

- H. Goldstein, C. Poole and J. Safko, *Classical Mechanics*

In previous editions it was known simply as "Goldstein" and has been the canonical choice for generations of students. Although somewhat verbose, it is considered the standard reference on the subject. Goldstein died and the current, third, edition found two extra authors.

- L. Landau an E. Lifshitz, *Mechanics*

This is a gorgeous, concise and elegant summary of the course in 150 content packed pages. Landau is one of the most important physicists of the 20th century and this is the first volume in a series of ten, considered by him to be the "theoretical minimum" amount of knowledge required to embark on research in physics. In 30 years, only 43 people passed Landau's exam!

A little known fact: Landau originally co-authored this book with one of his students, Leonid Pyatigorsky. They subsequently had a falling out and the authorship was changed. There are rumours that Pyatigorsky got his own back by denouncing Landau to the Soviet authorities, resulting in his arrest.

- V. I. Arnold, *Mathematical Methods of Classical Mechanics*

Arnold presents a more modern mathematical approach to the topics of this course, making connections with the differential geometry of manifolds and forms. It kicks off with "The Universe is an Affine Space" and proceeds from there...

Contents

Acknowledgements

These notes rely heavily on the textbooks listed at the beginning and on notes from past courses given by others, in particular Anne Davis, Gary Gibbons, Robin Hudson, Michael Peskin and Neil Turok. My thanks also to Michael Efroimsky and Matt Headrick for useful comments. I am supported by the Royal Society.

1. Newton's Laws of Motion

> "So few went to hear him, and fewer understood him, that oftimes he did, for want of hearers, read to the walls. He usually stayed about half an hour; when he had no auditors he commonly returned in a quarter of that time."
>
> *Appraisal of a Cambridge lecturer in classical mechanics, circa 1690*

1.1 Introduction

The fundamental principles of classical mechanics were laid down by Galileo and Newton in the 16^{th} and 17^{th} centuries. In 1686, Newton wrote the *Principia* where he gave us three laws of motion, one law of gravity and pretended he didn't know calculus. Probably the single greatest scientific achievement in history, you might think this pretty much wraps it up for classical mechanics. And, in a sense, it does. Given a collection of particles, acted upon by a collection of forces, you have to draw a nice diagram, with the particles as points and the forces as arrows. The forces are then added up and Newton's famous "$F = ma$" is employed to figure out where the particle's velocities are heading next. All you need is enough patience and a big enough computer and you're done.

From a modern perspective this is a little unsatisfactory on several levels: it's messy and inelegant; it's hard to deal with problems that involve extended objects rather than point particles; it obscures certain features of dynamics so that concepts such as chaos theory took over 200 years to discover; and it's not at all clear what the relationship is between Newton's classical laws and quantum physics.

The purpose of this course is to resolve these issues by presenting new perspectives on Newton's ideas. We shall describe the advances that took place during the 150 years after Newton when the laws of motion were reformulated using more powerful techniques and ideas developed by some of the giants of mathematical physics: people such as Euler, Lagrange, Hamilton and Jacobi. This will give us an immediate practical advantage, allowing us to solve certain complicated problems with relative ease (the strange motion of spinning tops is a good example). But, perhaps more importantly, it will provide an elegant viewpoint from which we'll see the profound basic principles which underlie Newton's familiar laws of motion. We shall prise open "$F = ma$" to reveal the structures and symmetries that lie beneath.

Moreover, the formalisms that we'll develop here are the basis for *all* of fundamental modern physics. Every theory of Nature, from electromagnetism and general relativity, to the standard model of particle physics and more speculative pursuits such as string theory, is best described in the language we shall develop in this course. The new formalisms that we'll see here also provide the bridge between the classical world and the quantum world.

There are phenomena in Nature for which these formalisms are not particularly useful. Systems which are dissipative, for example, are not so well suited to these new techniques. But if you want to understand the dynamics of planets and stars and galaxies as they orbit and spin, or you want to understand what's happening at the LHC where protons are collided at unprecedented energies, or you want to know how electrons meld together in solids to form new states of matter, then the foundations that we'll lay in in this course are a must.

1.2 Newtonian Mechanics: A Single Particle

In the rest of this section, we'll take a flying tour through the basic ideas of classical mechanics handed down to us by Newton. More details can be found in the lectures on Dynamics and Relativity.

We'll start with a single particle.. A *particle* is defined to be an object of insignificant size. e.g. an electron, a tennis ball or a planet. Obviously the validity of this statement depends on the context: to first approximation, the earth can be treated as a particle when computing its orbit around the sun. But if you want to understand its spin, it must be treated as an extended object.

The motion of a particle of mass m at the position $\mathbf{r}$ is governed by *Newton's Second Law* $\mathbf{F} = m\mathbf{a}$ or, more precisely,

$$\mathbf{F}(\mathbf{r}, \dot{\mathbf{r}}) = \dot{\mathbf{p}} \tag{1.1}$$

where $\mathbf{F}$ is the force which, in general, can depend on both the position $\mathbf{r}$ as well as the velocity $\dot{\mathbf{r}}$ (for example, friction forces depend on $\dot{\mathbf{r}}$) and $\mathbf{p} = m\dot{\mathbf{r}}$ is the momentum. Both $\mathbf{F}$ and $\mathbf{p}$ are 3-vectors which we denote by the bold font. Equation (1.1) reduces to $\mathbf{F} = m\mathbf{a}$ if $\dot{m} = 0$. But if $m = m(t)$ (as it does, for example, in rocket science), then the form with $\dot{\mathbf{p}}$ is correct. For the rest of these lectures, we will assume that mass is constant, so $\dot{m} = 0$.

General theorems governing differential equations guarantee that, if we are given $\mathbf{r}$ and $\dot{\mathbf{r}}$ at an initial time $t = t_0$, then we can integrate equation (1.1) to determine $\mathbf{r}(t)$ for all t, at least as long as $\mathbf{F}$ remains finite. This is the goal of classical dynamics.

Newton's second law (1.1) comes with an important caveat: it holds only in an *inertial frame*. This is defined to be a frame in which a free particle with $\dot{m} = 0$ travels in a straight line, so that

$$\mathbf{r} = \mathbf{r}_0 + \mathbf{v}t \tag{1.2}$$

Newtons's first law is the statement that inertial frames exist.

An inertial frame is not unique. In fact, there are an infinite number of inertial frames. Let S be one such inertial frame. Then there are 10 linearly independent transformations $S \to S'$ such that S' is also an inertial frame, which means that if (1.2) holds in S, then it also holds in S'. These transformations are

- 3 Rotations: $\mathbf{r}' = O\mathbf{r}$ where O is a 3×3 orthogonal matrix.
- 3 Translations: $\mathbf{r}' = \mathbf{r} + \mathbf{c}$ for a constant vector $\mathbf{c}$.
- 3 Boosts: $\mathbf{r}' = \mathbf{r} + \mathbf{u}t$ for a constant velocity $\mathbf{u}$.
- 1 Time Translation: $t' = t + c$ for a constant real number c

If motion is uniform in S, then it will also be uniform in S'. These transformations make up the *Galilean Group* under which Newton's laws are invariant. They will play a particularly important role in Section 2.3 where we will see that each of these symmetries of space and time gives rise to an associated conservation law.

As a parenthetical remark, recall from special relativity that Einstein's laws of motion are invariant under Lorentz transformations which, together with translations, make up the Poincaré group. We can recover the Galilean group from the Poincaré group by pretending that the speed of light is infinite.

1.2.1 Angular Momentum

We define the *angular momentum* $\mathbf{L}$ of a particle and the *torque* $\boldsymbol{\tau}$ acting upon it as

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} \qquad \text{and} \qquad \boldsymbol{\tau} = \mathbf{r} \times \mathbf{F}$$

Note that, unlike linear momentum $\mathbf{p}$, both $\mathbf{L}$ and $\boldsymbol{\tau}$ depend on where we take the origin: we measure angular momentum with respect to a particular point.

If we take the cross-product of Newton's second law (1.1) with $\mathbf{r}$ then, using the fact that $\dot{\mathbf{r}}$ is parallel to $\mathbf{p}$, we can write $\frac{d}{dt}(\mathbf{r} \times \mathbf{p}) = \mathbf{r} \times \dot{\mathbf{p}}$. We then get a version of Newton's second law that holds for angular momentum:

$$\boldsymbol{\tau} = \dot{\mathbf{L}} \tag{1.3}$$

1.2.2 Conservation Laws

Two trivial, but nonetheless important, conservation laws follow immediately from (1.1) and (1.3). These are:

- If $\mathbf{F} = 0$ then $\mathbf{p}$ is constant throughout the motion
- If $\boldsymbol{\tau} = 0$ then $\mathbf{L}$ is constant throughout the motion

Notice that $\boldsymbol{\tau} = 0$ does not necessarily require $\mathbf{F} = 0$, but only $\mathbf{r} \times \mathbf{F} = 0$. This means that $\mathbf{F}$ must be parallel to $\mathbf{r}$. This is the definition of a *central force*. An example is given by the gravitational force that the Sun exerts on the Earth. These seemingly trivial conservation laws will gain a level of profundity in Section 2.3 where we see how they are related to the Galilean symmetries described above.

1.2.3 Energy

The concept of "force" is very 17^{th} century. A large part of these lectures will be devoted on finding different ways to formulate classical mechanics that largely do away with the notion of force, instead focussing on the more modern concept of energy. (Here modern means 19^{th} century.)

A particle has *kinetic energy* T, defined as

$$T = \frac{1}{2} m \, \dot{\mathbf{r}} \cdot \dot{\mathbf{r}}$$

We can compute how the kinetic energy changes with time. Invoking Newton's law (1.1), we have

$$\frac{dT}{dt} = \dot{\mathbf{p}} \cdot \dot{\mathbf{r}} = \mathbf{F} \cdot \dot{\mathbf{r}}$$

Suppose that the particle travels from position $\mathbf{r}_1$ at time t_1 to position $\mathbf{r}_2$ at time t_2. Then the change in kinetic energy is given by

$$T(t_2) - T(t_1) = \int_{t_1}^{t_2} \frac{dT}{dt} \, dt = \int_{t_1}^{t_2} \mathbf{F} \cdot \dot{\mathbf{r}} \, dt = \int_{\mathbf{r}_1}^{\mathbf{r}_2} \mathbf{F} \cdot d\mathbf{r} \tag{1.4}$$

The final expression, involving the integral of the force over the path, is called the *work done* by the force. This simple calculation tells us that the work done is equal to the change in kinetic energy.

From now on we will mostly focus on a very special type of force known as a *conservative* force. Such a force depends only on position $\mathbf{r}$, rather than velocity $\dot{\mathbf{r}}$, and has the additional property that the work done is independent of the path taken. In particular, for a closed path, the work done vanishes. A result from Vector Calculus tells us that

$$\oint \mathbf{F} \cdot d\mathbf{r} = 0 \qquad \Leftrightarrow \qquad \nabla \times \mathbf{F} = 0 \tag{1.5}$$

It is a deep property of flat space $\mathbf{R}^3$, proven in Vector Calculus, that (1.5) implies we may write the force as

$$\mathbf{F} = -\nabla V(\mathbf{r}) \tag{1.6}$$

for some *potential* $V(\mathbf{r})$. Systems which admit a potential of this form include gravitational, electrostatic and interatomic forces.

When we have a conservative force, we necessarily have a conservation law for energy. To see this, return to equation (1.4) which now reads

$$T(t_2) - T(t_1) = -\int_{\mathbf{r}_1}^{\mathbf{r}_2} \nabla V \cdot d\mathbf{r} = -V(t_2) + V(t_1) \tag{1.7}$$

Rearranging things, we have

$$T(t_1) + V(t_1) = T(t_2) + V(t_2) \equiv E \tag{1.8}$$

which tells us that $E = T + V$ is also a constant of motion. This is the energy.

Examples

Here are a collection of simple examples, all of which we met in our previous lectures on classical mechanics.

- The Simple Harmonic Oscillator

  This is a one-dimensional system with a force proportional to the distance x to the origin: $F(x) = -kx$. This force arises from a potential $V = \frac{1}{2}kx^2$. Since $F \neq 0$, momentum is not conserved, which is obvious because the particle oscillates backwards and forwards. And since the system lives in only one dimension, angular momentum is not defined. But energy $E = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}kx^2$ is conserved.

- The Damped Simple Harmonic Oscillator

  We now include a friction term so that $F(x, \dot{x}) = -kx - \gamma\dot{x}$. Since F is not conservative, energy is no longer conserved. This system loses energy until it comes to rest.

- A Particle Moving Under Gravity

  Consider a particle of mass m moving in 3 dimensions under the gravitational pull of a much larger particle of mass M. The force is $\mathbf{F} = -(GMm/r^2)\hat{\mathbf{r}}$ which arises from the potential $V = -GMm/r$. Again, the linear momentum $\mathbf{p}$ of the smaller particle is not conserved, but the force is both central and conservative, ensuring the particle's total energy E and the angular momentum $\mathbf{L}$ are conserved.

1.3 Newtonian Mechanics: Many Particles

It's easy to generalise the above discussion to many particles: we simply add an index to everything in sight. We'll consider N particles, labelled by the index $a = 1, \ldots, N$. We'll denote the mass and position of the a^{th} particle as m_a and $\mathbf{r}_a$ respectively. Newton's law (1.1) now gets upgraded to

$$\mathbf{F}_a = \dot{\mathbf{p}}_a \tag{1.9}$$

where $\mathbf{F}_a$ is the experienced force by the a^{th} particle. The novelty is that, in addition to external forces forces that act on all particles, there can also be interaction forces acting between various particles. We assume that the forces only act between pairs of particles, rather than between three or more particles. (This assumption holds for all the known laws of physics.) This means that, in general, we can decompose the force $\mathbf{F}_a$ acting on the a^{th} particle as

$$\mathbf{F}_a = \sum_{b \neq a} \mathbf{F}_{ab} + \mathbf{F}_a^{\text{ext}}$$

where $\mathbf{F}_{ab}$ is the force acting on the a^{th} particle due to the b^{th} particle, while $\mathbf{F}_a^{\text{ext}}$ is the external force on the a^{th} particle.

We can make reveal something interesting by summing over all N particles,

$$\sum_a \mathbf{F}_a = \sum_{a,b \text{ with } b \neq a} \mathbf{F}_{ab} + \sum_a \mathbf{F}_a^{\text{ext}} = \sum_{a<b} (\mathbf{F}_{ab} + \mathbf{F}_{ba}) + \sum_a \mathbf{F}_a^{\text{ext}} \tag{1.10}$$

where, in the second equality, we've re-written the sum to be over all pairs $a < b$. At this stage we invoke another assumption about the interaction forces $\mathbf{F}_{ab}$. We assume that these forces obey

$$\mathbf{F}_{ab} = -\mathbf{F}_{ba} \tag{1.11}$$

This assumption is important enough to be elevated to the grand sounding *Newton's third law*. It's famous mantra is: every action has an equal and opposite reaction. For

our purposes, it means that the first term in (1.10) vanishes and we're left with

$$\sum_a \mathbf{F}_a = \sum_a \mathbf{F}_a^{\rm ext} \equiv \mathbf{F}^{\rm ext} \tag{1.12}$$

To see the physics lurking behind this, we define the total mass M and the *centre of mass* $\mathbf{R}$ of the system by

$$M = \sum_a m_a \quad \text{and} \quad \mathbf{R} = \frac{1}{M}\sum_a m_a \mathbf{r}_a \tag{1.13}$$

Then using Newton's second law (1.9), and summing over all particles, we arrive at the simple formula,

$$\mathbf{F}^{\rm ext} = M\ddot{\mathbf{R}} \tag{1.14}$$

which is identical to Newton's second law (1.1) for a single particle. This is an important formula. It tells that the centre of mass of a system of particles acts just as if all the mass were concentrated there. Practically, it means that it doesn't matter if you throw a tennis ball or a very lively cat: the centre of mass of each traces the same path.

Newton's third law (1.11) sometimes seems like a poor cousin of Newton's second law, $F = ma$. But the calculation above shows its importance. The result (1.14) tells us that we don't need to know about the internal constituents of any object to make progress. You can figure out the motion of the planets around the Sun without knowing about their internal geophysical dynamics because none of those internal forces matter: they all just cancel out due Newton's third law. Equally, you can figure out the motion of the planets without knowing about the constituent quarks, again because Newton's third law tells us that all the interaction forces between the quarks cancel out. In other words, Newton's third law makes science possible. It tells us that you don't have to sweat the small stuff. You can just focus on the big, important things.

1.3.1 Momentum Revisited

The *total momentum* is defined to be $\mathbf{P} = \sum_a \mathbf{p}_a$ and our version of Newton's law (1.14) can be written as $\dot{\mathbf{P}} = \mathbf{F}^{\rm ext}$. This immediately gives us a conservation law for the total linear momentum of a system of particles: $\mathbf{P}$ is constant if $\mathbf{F}^{\rm ext}$ vanishes.

Similarly, we define *total angular momentum* to be $\mathbf{L} = \sum_a \mathbf{L}_a$. Previously we showed that the change of angular momentum is equal to the torque (1.3). We now want a

version of this equation for the total angular momentum of a system of particles. If we compute the time derivative, we find

$$\dot{\mathbf{L}} = \sum_a \mathbf{r}_a \times \dot{\mathbf{p}}_a = \sum_a \mathbf{r}_a \times \left(\sum_{b \neq a} \mathbf{F}_{ab} + \mathbf{F}_a^{\rm ext} \right) = \sum_{a,b \,\text{with}\, a \neq b} \mathbf{r}_a \times \mathbf{F}_{ba} + \sum_a \mathbf{r}_a \times \mathbf{F}_a^{\rm ext}$$

The last term in this expression is the definition of *total external torque*:

$$\boldsymbol{\tau}^{\rm ext} = \sum_a \mathbf{r}_a \times \mathbf{F}_a^{\rm ext}$$

However, that still leaves us with the first term on the right hand side. To get a version of our previous result (1.3), we would need this term to vanish. But, given our assumptions so far, there's no reason for this to happen. In fact, we need yet another assumption. To see this, we again rewrite this term as a sum over pairs $a < b$,

$$\sum_{a,b \,\text{with}\, a \neq b} \mathbf{r}_a \times \mathbf{F}_{ba} = \sum_{a<b} (\mathbf{r}_a - \mathbf{r}_b) \times \mathbf{F}_{ab}$$

This will vanish if and only if the force $\mathbf{F}_{ab}$ is parallel to the line joining to two particles $(\mathbf{r}_a - \mathbf{r}_b)$. We make this assumption. Together with (1.11), this is sometimes referred to as the strong form of Newton's third law. When this condition holds, we have a $\dot{\mathbf{L}} = \boldsymbol{\tau}^{\rm ext}$, which is the version of (1.3) that holds for a system of particles. We can also immediately find a trivial statement about the conservation of total angular momentum, namely $\mathbf{L}$ is constant if $\boldsymbol{\tau}^{\rm ext} = 0$.

Most forces do indeed obey the strong form of Newton's third law, with $\mathbf{F}_{ab} = -\mathbf{F}_{ba}$ and $\mathbf{F}_{ab}$ is parallel to $(\mathbf{r}_a - \mathbf{r}_b)$. In particular, both the gravitational and electrostatic forces have this property.

However, at first glance, it appears that some forces do not obey Newton's third law. The most famous example is the Lorentz force acting on two moving particles, each carrying electric charge q. This force is given by

$$\mathbf{F}_{ab} = q\mathbf{v}_a \times \mathbf{B}_b$$

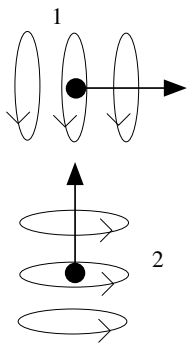


Figure 1:

where $\mathbf{v}_a$ is the velocity of the $a^{\rm th}$ particle and $\mathbf{B}_b$ is the magnetic field generated by the motion of the $b^{\rm th}$ particle. A description of both the Lorentz force law, and the way a moving charge generates a magnetic field, can be found in the lectures on Electromagnetism. For now, all we need is the direction of the magnetic field which

is shown in the figure. This depicts two particles, crossing each other in a "T" . Taking cross products, the force on particle 1 from particle 2 is schematically of the form

$$\mathbf{F}_{12} \ \sim \rightarrow \times \leftarrow = 0 \tag{1.15}$$

Importantly it vanishes! Meanwhile, the force on particle 2 from particle 1 is non-zero, and in the direction

$$\mathbf{F}_{21} \ \sim \uparrow \times \otimes \ \sim \leftarrow \tag{1.16}$$

What's going on? Above, I just bigged up Newton's third law and said it was necessary for science to happen. Does this violation mean that conservation of total linear and angular momentum is violated? Thankfully, no! We need to realise that the electromagnetic field itself carries angular momentum which restores the conservation law. Once we appreciate this, we've got a rather cheap counterexample to Newton's third law, little different from an underwater swimmer who can twist and turn and appear to violate Newton's third law if we don't take into account the momentum of the water.

1.3.2 Energy Revisited

The total kinetic energy of a system of many particles is $T = \frac{1}{2}\sum_a m_a \dot{\mathbf{r}}_a^2$. There's a useful way to write this. We decompose the position vector $\mathbf{r}_a$ as

$$\mathbf{r}_a = \mathbf{R} + \tilde{\mathbf{r}}_a$$

with $\mathbf{R}$ the centre of mass (1.13) and $\tilde{\mathbf{r}}_a$ is the distance from the centre of mass to the particle a. Then, a short calculation shows that the total kinetic energy decomposes as

$$T = \frac{1}{2}M\dot{\mathbf{R}}^2 + \frac{1}{2}\sum_a m_a \dot{\tilde{\mathbf{r}}}_a^2 \tag{1.17}$$

Here the would-be cross-term $\sum_a m_a \dot{\mathbf{R}}\dot{\tilde{\mathbf{r}}}_a$ vanishes on account of the fact that, by construction, $\sum_a m_a \tilde{\mathbf{r}}_a = 0$. We see that the kinetic energy splits up into the kinetic energy of the centre of mass, together with an *internal energy* describing how the system is moving around its centre of mass. We can now repeat the calculation that we did for a single particle and see how the total kinetic energy changes over time. We have

$$T(t_2) - T(t_1) = \sum_i \int \mathbf{F}_a^{\text{ext}} \cdot d\mathbf{r}_a + \sum_{a\neq b} \int \mathbf{F}_{ab} \cdot d\mathbf{r}_a$$

We again need to restrict ourselves to conservative forces to find a form of energy conservation. This time we need two requirements,

- Conservative external forces: $\mathbf{F}_a^{\text{ext}} = -\nabla_a V_a(\mathbf{r}_1, \ldots, \mathbf{r}_N)$
- Conservative internal forces: $\mathbf{F}_{ab} = -\nabla_a V_{ab}(\mathbf{r}_1, \ldots, \mathbf{r}_N)$

where $\nabla_a \equiv \partial/\partial \mathbf{r}_a$.

There is a natural form of these conservative forces that results in the strong version of Newton's third law, $\mathbf{F}_{ab} = -\mathbf{F}_{ba}$ together with the requirement that this is parallel to $(\mathbf{r}_a - \mathbf{r}_b)$. This follows for any internal potential satisfying $V_{ab} = V_{ba}$ with

$$V_{ab}(\mathbf{r}_1, \ldots \mathbf{r}_N) = V_{ab}(|\mathbf{r}_a - \mathbf{r}_b|) \tag{1.18}$$

Note, in particular, that V_{ab} depends only on the distance between the a^{th} and b^{th} particles, and doesn't care less where the other particles sit. We also insist on a similar restriction for the external forces, $V_a(\mathbf{r}_1, \ldots, \mathbf{r}_N) = V_a(\mathbf{r}_i a$, so that the force on particle a does not depend on the positions of the other particles. Then, following the steps we took in the single particle case, we can define the *total potential energy* $V = \sum_a V_i + \sum_{a<b} V_{ab}$ and we can show that $E = T + V$ is conserved.

2. The Lagrangian Formalism

> When I was in high school, my physics teacher called me down one day after class and said, "You look bored, I want to tell you something interesting". Then he told me something I have always found fascinating. Every time the subject comes up I work on it.
>
> *Richard Feynman*

Feynman's teacher told him about the "Principle of Least Action", one of the most profound results in physics.

2.1 The Principle of Least Action

First, let's get our notation right. Part of the power of the Lagrangian formulation over the Newtonian approach is that it does away with vectors in favour of more general coordinates. We start by doing this trivially. We rewrite the positions of N particles with coordinates $\mathbf{r}_i$ as x^A where $A = 1, \ldots 3N$. Then Newton's equations read

$$\dot{p}_A = -\frac{\partial V}{\partial x^A} \tag{2.1}$$

where $p_A = m_A \dot{x}^A$. The number of *degrees of freedom* of the system is said to be $3N$. These parameterise a $3N$-dimensional space known as the *configuration space* $\mathcal{C}$. Each point in $\mathcal{C}$ specifies a configuration of the system (i.e. the positions of all N particles). Time evolution gives rise to a curve in $\mathcal{C}$.

The Lagrangian

Define the *Lagrangian* to be a function of the positions x^A and the velocities $\dot{x}^A$ of all the particles, given by

$$L(x^A, \dot{x}^A) = T(\dot{x}^A) - V(x^A) \tag{2.2}$$

where $T = \frac{1}{2}\sum_A m_A (\dot{x}^A)^2$ is the kinetic energy, and $V(x^A)$ is the potential energy. Note the all-important minus sign between T and V! This is what distinguishes the Lagrangian from the energy (at least in these simplest examples.)

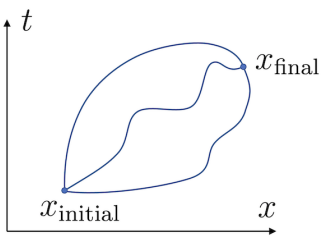


Figure 2:

The principle of least action invites us to expand our horizons. Rather than thinking only of the trajectory that obeys Newton's equations of motion, we will instead consider all possible paths at once. We will only stipulate that the path has some fixed beginning and some fixed end,

$$x^A(t_i) = x^A_{\text{initial}} \qquad \text{and} \qquad x^A(t_f) = x^A_{\text{final}}$$

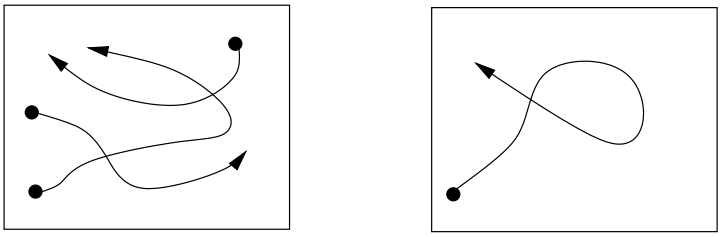


Figure 3: The path of particles in real space (on the left) and in configuration space (on the right).

Of all these possible paths, only one is the true path taken by the system. The question is: which one?

To answer this, we assign a number to each path. This number is called the *action* S defined as

$$S[x^A(t)] = \int_{t_i}^{t_f} L(x^A(t), \dot{x}^A(t))\, dt$$

The action is a functional (i.e. a function of the path which is itself a function). The principle of least action is the following result:

Theorem (Principle of Least Action): The actual path taken by the system is an extremum of S.

Proof: Consider varying a given path slightly, so

$$x^A(t) \to x^A(t) + \delta x^A(t)$$

where we fix the end points of the path by demanding $\delta x^A(t_i) = \delta x^A(t_f) = 0$. Then the change in the action is

$$\begin{aligned}
\delta S &= \delta \left[\int_{t_i}^{t_f} L\, dt \right] \\
&= \int_{t_i}^{t_f} \delta L\, dt \\
&= \int_{t_i}^{t_f} \left(\frac{\partial L}{\partial x^A} \delta x^A + \frac{\partial L}{\partial \dot{x}^A} \delta \dot{x}^A \right) dt
\end{aligned}$$

At this point we integrate the second term by parts to get

$$\delta S = \int_{t_i}^{t_f} \left(\frac{\partial L}{\partial x^A} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}^A} \right) \right) \delta x^A\, dt + \left[\frac{\partial L}{\partial \dot{x}^A} \delta x^A \right]_{t_i}^{t_f}$$

But the final term vanishes since we have fixed the end points of the path so $\delta x^A(t_i) = \delta x^A(t_f) = 0$. The requirement that the action is an extremum says that $\delta S = 0$ for all changes in the path $\delta x^A(t)$. We see that this holds if and only if

$$\frac{\partial L}{\partial x^A} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^A}\right) = 0 \qquad \text{for each } A = 1, \ldots 3N \tag{2.3}$$

These are known as the *Euler-Lagrange equations* (or sometimes just as Lagrange's equations). To finish the proof, we need only show that Lagrange's equations are equivalent to Newton's. From the definition of the Lagrangian (2.2), we have $\partial L/\partial x^A = -\partial V/\partial x^A$, while $\partial L/\partial \dot{x}^A = p_A$. It's then easy to see that equations (2.3) are indeed equivalent to (2.1). □

Some remarks on this important result:

- The principle of *least* action is a slight misnomer. The proof only requires that $\delta S = 0$, and does not specify whether it is a maxima or minima of S. Since $L = T - V$, we can always increase S by taking a very fast, wiggly path with $T \gg 0$, so the true path is never a maximum. However, it may be either a minimum or a saddle point. So "Principle of stationary action" would be a more accurate, but less catchy, name. It is sometimes called "Hamilton's principle".

- The Lagrangian L is not unique. We may make the transformation

$$L' = \alpha L \tag{2.4}$$

  with $\alpha \in \mathbf{R}$ then the equations of motion remain unchanged. This is also true if we add a total derivative to the Lagrangian,

$$L' = L + \frac{df}{dt} \tag{2.5}$$

  for any function f. To see that the last statement is true, we could either plug L' into Lagrange's equations or, alternatively note that the action changes only by a constant under this transformation and so its stationary points remain the same.

- The action has dimension $[S] = ML^2T^{-1}$. This is the same as the dimension of angular momentum. More importantly, it is the same as the dimension of Planck's constant $\hbar$. This last statement underlies a beautiful generalisation of the action principle to quantum mechanics, due to Feynman, in which the particle takes *all paths* with some probability determined by $S/\hbar$. We will describe this in Section 4.8.

As an aside, neither of the transformation (2.4) and (2.5) leave the physics invariant in the quantum world. The first transformation (2.4) rescales Planck's constant, while the second transformation (2.5) is related to rather subtle and interesting topological effects that will be discussed in the lectures on Gauge Theory.

- *All* the fundamental laws of physics can be written in terms of an action principle. This includes electromagnetism, general relativity, the standard model of particle physics, and attempts to go beyond the known laws of physics such as string theory. For example, (nearly) everything we know about the universe is captured in the Lagrangian

$$L = \sqrt{g}\left(R - \tfrac{1}{2}F_{\mu\nu}F^{\mu\nu} + \bar{\psi}\not{D}\psi + |\mathcal{D}h|^2 + h\bar{\psi}\psi - V(h)\right)$$

  where the terms carry the names of Einstein, Maxwell (or Yang and Mills), Dirac, Higgs and Yukawa respectively, and describe gravity, the forces of nature (electromagnetism and the nuclear forces), the dynamics of particles like electrons and quarks and the dynamics of the Higgs boson. If you want to understand what the terms in this equation really mean, then you can find explanations in the lectures on General Relativity, Electromagnetism, and Quantum Field Theory.

Back to more down-to-earth pursuits, there are three very important reasons for working with the Euler-Lagrange equations rather than Newton's. The first is that Euler-Lagrange equations hold in any coordinate system, while Newton's are restricted to an inertial frame. The second is that the Lagrangian approach reveals a deep connection between symmetries and conservation laws. And the third is the ease with which we can deal with constraints in the Lagrangian system. In the rest of this section, we look at each of these in turn, illustrated with plenty of examples.

2.2 Generalised Coordinates

We shall now show that Lagrange's equations hold in any coordinate system. In fact, this follows immediately from the action principle, which is a statement about paths and not about coordinates. But here we shall be a little more pedestrian in order to explain exactly what we mean by changing coordinates, and why it's useful. Let

$$q^i = q^i(x^1, \ldots, x^{3N}, t) \qquad \text{with } i = 1, \ldots, 3N$$

where we've included the possibility of using a coordinate system which changes with time t. Then, by the chain rule, we can write

$$\dot{q}^i = \frac{dq^i}{dt} = \frac{\partial q^i}{\partial x^A}\dot{x}^A + \frac{\partial q^i}{\partial t}$$

In this equation, and for the rest of this course, we're using the summation convention in which repeated indices are summed over. To be a good coordinate system, we should be able to invert the relationship so that $x^A = x^A(q^i, t)$ which we can do as long as we have $\det\big(\partial x^A/\partial q^i\big) \neq 0$. Then we have,

$$\dot{x}^A = \frac{\partial x^A}{\partial q^i}\dot{q}^i + \frac{\partial x^A}{\partial t} \tag{2.6}$$

Now we can look at what becomes of the Lagrangian $L(x^A, \dot{x}^A)$ when we substitute in $x^A = x^A(q^i, t)$. Using (2.6) we have

$$\frac{\partial L}{\partial q^i} = \frac{\partial L}{\partial x^A}\frac{\partial x^A}{\partial q^i} + \frac{\partial L}{\partial \dot{x}^A}\left(\frac{\partial^2 x^A}{\partial q^i \partial q^j}\dot{q}^j + \frac{\partial^2 x^A}{\partial q^i \partial t}\right) \tag{2.7}$$

while

$$\frac{\partial L}{\partial \dot{q}^i} = \frac{\partial L}{\partial \dot{x}^A}\frac{\partial \dot{x}^A}{\partial \dot{q}^i} = \frac{\partial L}{\partial \dot{x}^A}\frac{\partial x^A}{\partial q^i} \tag{2.8}$$

where, in the second equality, we "cancelled the dots" to write $\partial \dot{x}^A/\partial \dot{q}^i = \partial x^A/\partial q^i$ which we can prove by differentiating the expression (2.6) with respect to $\dot{q}^i$. Taking the time derivative of (2.8) gives us

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^A}\right)\frac{\partial x^A}{\partial q^i} + \frac{\partial L}{\partial \dot{x}^A}\left(\frac{\partial^2 x^A}{\partial q^i \partial q^j}\dot{q}^j + \frac{\partial^2 x^A}{\partial q^i \partial t}\right) \tag{2.9}$$

So combining (2.7) with (2.9) we find

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} = \left[\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^A}\right) - \frac{\partial L}{\partial x^A}\right]\frac{\partial x^A}{\partial q^i} \tag{2.10}$$

Equation (2.10) is our final result. We see that if the Euler-Lagrange' equation is solved in the x^A coordinate system, so that [...] on the RHS vanishes, then it is also solved in the q^i coordinate system,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} = 0 \tag{2.11}$$

Conversely, if (2.11) holds, then the Euler-Lagrange equation is also satisfied in the x^A coordinate system as long as our choice of coordinates is invertible, ensuring that $\det\big(\partial x^A/\partial q^i\big) \neq 0$.

The upshot is that the Euler-Lagrange equations hold in *any* coordinate system. This is in contrast to Newton's equations which are only valid in an inertial frame.

In our original Cartesian coordinates, the momentum was related to the velocity by

$$p_A = \frac{\partial L}{\partial \dot{x}^A}$$

Now we have a more general set of coordinates at our disposal. Nonetheless, the analogous quantity plays an important role. We define

$$p_i = \frac{\partial L}{\partial \dot{q}^i} \tag{2.12}$$

This is known as the *generalised momentum* conjugate to q^i. Note that our notation has become a little subtle here: p_A and p_i can be very different object, but are distinguished only by the letter we've chosen to label the index. In practice, this won't lead to much ambiguity when solving specific problems. In terms of the generalised momentum p_i, the Euler-Lagrange equations become

$$\frac{\partial p_i}{\partial t} - \frac{\partial L}{\partial q^i} = 0 \tag{2.13}$$

We'll meet many examples of generalised coordinates in these lectures. But to get us going, here are two examples that we've seen before, but now phrased in the Lagrangian language .

An Example: Polar Coordinates

Consider a particle moving in $\mathbf{R}^3$ with Cartesian coordinates $\mathbf{x} = (x, y, z)$ and a general potential $V(\mathbf{x})$. The Lagrangian is

$$L = \frac{1}{2}m\dot{\mathbf{x}}^2 - V(\mathbf{x})$$

We can work instead in spherical polar coordinates, defined by

$$x = r\sin\theta\cos\phi \quad , \quad y = r\sin\theta\sin\phi \quad , \quad z = r\cos\theta$$

where $0 \leq \theta \leq \pi$ and $0 \leq \phi < 2\pi$. We can substitute this change of variables directly into the Lagrangian to get

$$L = \tfrac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\,\dot{\phi}^2) - V(r, \theta, \phi) \tag{2.14}$$

The generalised momenta (2.12) are

$$p_r = m\dot{r} \quad , \quad p_\theta = mr^2\dot{\theta} \quad , \quad p_\phi = mr^2\sin^2\theta\dot{\phi}$$

We can use this form of the Lagrangian to immediately derive the equations of motion in polar coordinates. We'll see a number of uses for this as we proceed.

2.2.1 Rotating Coordinate Systems

Consider a free particle with Lagrangian given by

$$L = \tfrac{1}{2}m\dot{\mathbf{x}}^2$$

with $\mathbf{x} = (x, y, z)$. Now measure the motion of the particle with respect to a coordinate system which is rotating with angular velocity $\boldsymbol{\omega} = (0, 0, \omega)$ about the z axis. If $\mathbf{x}' = (x', y', z')$ are the coordinates in the rotating system, we have the relationship

$$x' = x\cos\omega t + y\sin\omega t \quad , \quad y' = y\cos\omega t - x\sin\omega t \quad , \quad z' = z \tag{2.15}$$

Then we can substitute these expressions into the Lagrangian to find L in terms of the rotating coordinates,

$$L = \tfrac{1}{2}m[(\dot{x}' - \omega y')^2 + (\dot{y}' + \omega x')^2 + \dot{z}^2] = \tfrac{1}{2}m(\dot{\mathbf{x}}' + \boldsymbol{\omega}\times\mathbf{x}')^2 \tag{2.16}$$

In this rotating frame, we can use Lagrange's equations to derive the equations of motion. Taking derivatives, we have

$$\frac{\partial L}{\partial \mathbf{x}'} = m(\dot{\mathbf{x}}'\times\boldsymbol{\omega} - \boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{x}')) \quad \text{and} \quad \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\mathbf{x}}'}\right) = m(\ddot{\mathbf{x}}' + \boldsymbol{\omega}\times\dot{\mathbf{x}}')$$

so Lagrange's equation reads

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\mathbf{x}}'}\right) - \frac{\partial L}{\partial \mathbf{x}'} = 0 \quad \Rightarrow \quad m\ddot{\mathbf{x}}' = -m\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{x}') - 2m\boldsymbol{\omega}\times\dot{\mathbf{x}}' \tag{2.17}$$

The first and second terms on the right-hand side are the centrifugal and Coriolis forces respectively: they are examples of "fictitious forces", so called because they're a consequence of the reference frame, rather than any interaction.

The centrifugal force $\mathbf{F}_{\rm cent} = -m\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{x}')$ points outwards in the plane perpendicular to $\boldsymbol{\omega}$ with magnitude $m\omega^2|\mathbf{x}'_\perp| = m|\mathbf{v}_\perp|^2/|\mathbf{x}'_\perp|$ where $_\perp$ denotes the projection perpendicular to $\boldsymbol{\omega}$.

The Coriolis force $\mathbf{F}_{\rm cor} = -2m\boldsymbol{\omega}\times\dot{\mathbf{x}}'$ is responsible for the large scale circulation of oceans and the atmosphere. For a particle travelling on the surface of the rotating earth, the direction of the Coriolis force is drawn in figure 4. We see that a particle thrown in the northern hemisphere will be seen to rotate in a clockwise direction; a particle thrown in the southern hemisphere rotates in an anti-clockwise direction. For a particle moving along the equator, the Coriolis force points directly upwards and competes with gravity.

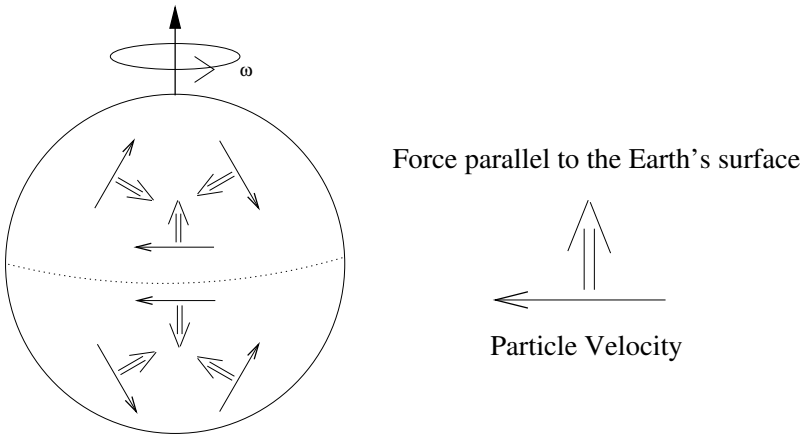


Figure 4: In the northern hemisphere, a particle is deflected in a clockwise direction; in the southern hemisphere in an anti-clockwise direction.

We previously derived the expressions for the centrifugal and Coriolis forces in our earlier course on classical mechanics. There we introduced a time-dependent basis of vectors, appropriate to describe what happens in the rotating frame and, after some manipulations, derived the equation of motion (2.17). The derivation using the Lagrangian is somewhat easier, largely because we have to think less: we just plug in the coordinate substitution (2.15) and go, with all the subtleties of which frame we're working in are dealt by the Lagrangian.

2.2.2 Joseph-Louis Lagrange (1736-1813)

> "His voice is very feeble, at least in that he does not become heated; he has a very pronounced Italian accent and pronounces the s like z ... The students, of whom the majority are incapable of appreciating him, give him little welcome, but the professors make amends for it."
>
> *Fourier analysis of Lagrange*

Lagrange started off life studying law but changed his mind and turned to mathematics after reading a book on optics by Halley (of comet fame). Despite being mostly self-taught, by the age of 19 he was a professor in his home town of Turin.

He stayed in Italy, somewhat secluded, for the next 11 years although he communicated often with Euler and, in 1766, moved to Berlin to take up Euler's recently vacated position. It was there he did his famous work on mechanics and the calculus of variations that we've seen above. In 1787 he moved once again, now to Paris. He was just in time for the French revolution and only survived a law ordering the arrest of

all foreigners after the intervention of the chemist Lavoisier who was a rather powerful political figure. (One year later, Lavoisier lost his power, followed quickly by his head.)

Lagrange published his collected works on mechanics in 1788 in a book called "Mechanique Analytique". He considered the work to be pure mathematics and boasts in the introduction that it contains no figures, thereby putting the anal in analytique.

2.3 Symmetries and Conservation Laws

When approaching any problem, the first thing that you should do is look for conserved quantities. One of the advantages of the Lagrangian approach is that conservation laws are intimately tied to symmetries. The link, as we now explain, is through an important and beautiful theorem due to Emmy Noether.

Let's start with a definition. A function $F(q^i, \dot{q}^i, t)$ of the coordinates, their time derivatives and (possibly) time t is called a *conserved quantity* if the total time derivative vanishes

$$\frac{dF}{dt} = \sum_{j=1}^{n} \left(\frac{\partial F}{\partial q^j}\dot{q}^j + \frac{\partial F}{\partial \dot{q}^j}\ddot{q}^j \right) + \frac{\partial F}{\partial t} = 0$$

whenever $q^i(t)$ satisfy Lagrange's equations. Conserved quantities are also called *constants of motion*, because F remains constant along the path followed by the system.

We'll meet many conserved quantities in these lectures (and, indeed, in subsequent lectures). Here are two of the most common and the most important:

Conservation of (Generalised) Momentum

Suppose that the Lagrangian L depends only on the time derivative of some coordinate, but not on the coordinate itself, meaning that $\partial L/\partial q^j = 0$ for some q^j. Then q^j is said to be *ignorable*. (Sometimes q^j is said to be *cyclic*.) Then the corresponding generalised momentum,

$$p_j = \frac{\partial L}{\partial \dot{q}^j} \tag{2.18}$$

is conserved.

The proof of this statement is straightforward. We have

$$\frac{dp_j}{dt} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^j}\right) = \frac{\partial L}{\partial q^j} = 0$$

where, in the second equality, we have used the Euler-Lagrange equation (2.11).

If our chosen coordinates q^i coincide with our original Cartesian coordinates x^A, then the conservation law described above is very intuitive. Suppose that a particle moves in $\mathbf{R}^2$ in a potential $V(\mathbf{x})$, with $\mathbf{x} = (x, y)$. Then if $V(\mathbf{x})$ is independent of one coordinate – say y – then the momentum in that direction, p_y, will be conserved. This is simply because the force is $\mathbf{F} = -\nabla V$ and if $\partial V/\partial y = 0$ then there's no force in the y-direction.

However, the conservation law above is more general since it applies to generalised coordinates, not just Cartesian coordinates. Here is another simple example to illustrate the basic idea. We'll stick with a particle moving in $\mathbf{R}^2$, but now we work in polar coordinates, defined by

$$x = r\cos\phi \quad \text{and} \quad y = r\sin\phi$$

Following our discussion of generalised coordinates in Section 2.2, we can simply substitute this into the Lagrangian to get

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) - V(\mathbf{x}) = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\phi}^2) - V(r, \phi)$$

Suppose now that the potential is rotationally invariant, so $\partial V/\partial\phi = 0$. Then the result above tells us that the generalised momentum p_ϕ is conserved, where

$$p_\phi = \frac{\partial L}{\partial\dot{\phi}} = mr^2\dot{\phi}$$

But this is something very familiar: it is just the angular momentum about the origin.

Conservation of Energy

Our second example arises when L does not depend explicitly on time t, so that $\partial L/\partial t = 0$. In this case, we have the constant of motion

$$E = \dot{q}^j\frac{\partial L}{\partial\dot{q}^j} - L \tag{2.19}$$

where, as always, we're employing the summation convention in the first term. As the name suggests, this constant of motion is identified as the energy of the system.

The proof that E is conserved is again of the plug-it-in-and-see variety. We have

$$\begin{aligned}\frac{dE}{dt} &= \ddot{q}^j\frac{\partial L}{\partial\dot{q}^j} + \dot{q}^j\frac{d}{dt}\left(\frac{\partial L}{\partial\dot{q}^j}\right) - \frac{\partial L}{\partial q^j}\dot{q}^j - \frac{\partial L}{\partial\dot{q}^j}\ddot{q}^j \\ &= \left[\frac{d}{dt}\left(\frac{\partial L}{\partial\dot{q}^j}\right) - \frac{\partial L}{\partial q^j}\right]\dot{q}^j\end{aligned}$$

where, by assumption, a would-be $\partial L/\partial t$ term does not appear in the first line. We see that the final expression in square brackets vanishes whenever the Euler-Lagrange equations (2.11) hold.

As these lectures progress, we'll get more intuition for why E should be identified as energy. But, for now, we can just look at our simplest example of a particle moving in the presence of a potential,

$$L = \frac{1}{2}m\dot{\mathbf{x}}^2 - V(\mathbf{x}) \quad \Rightarrow \quad E = m\dot{\mathbf{x}}^2 - L = \frac{1}{2}m\dot{\mathbf{x}}^2 + V(\mathbf{x})$$

We see that we do indeed recover the familiar expression for energy.

One final comment: when the energy is written as a function of q^i and p_i, it is known as the Hamiltonian of the system and denoted as H. We'll learn more about this in Section 4.

2.3.1 Noether's Theorem

We now turn to Noether's theorem, which relates symmetries of the action to conservation laws. To understand this, we first need to define more carefully what we mean by a symmetry.

Roughly speaking, a symmetry is a way to change the dynamical variables so that the value of the Lagrangian remains unchanged. To quantify this, we first introduce the idea of a one-parameter family of maps

$$q^i(t) \to q^i(s,t) \qquad\qquad s \in \mathbf{R}$$

such that $q^i(0,t) = q^i(t)$. Then this transformation is said to be a continuous *symmetry* of the Lagrangian L if

$$\frac{\partial}{\partial s} L(q^i(s,t), \dot{q}^i(s,t), t) = 0$$

Noether's theorem states that, for each such continuous symmetry, there exists a conserved quantity.

Proof of Noether's Theorem: The prof is again straightforward, once you know what you're looking for. Differentiating the Lagrangian with respect to the parameter s, we have

$$\frac{\partial L}{\partial s} = \frac{\partial L}{\partial q^i}\frac{\partial q^i}{\partial s} + \frac{\partial L}{\partial \dot{q}^i}\frac{\partial \dot{q}^i}{\partial s}$$

Evaluating this at $s = 0$, and using the fact that the parameter s labels a symmetry, we have

$$\begin{aligned} 0 = \left.\frac{\partial L}{\partial s}\right|_{s=0} &= \left.\frac{\partial L}{\partial q^i}\,\frac{\partial q^i}{\partial s}\right|_{s=0} + \left.\frac{\partial L}{\partial \dot{q}^i}\,\frac{\partial \dot{q}^i}{\partial s}\right|_{s=0} \\ &= \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right)\left.\frac{\partial q^i}{\partial s}\right|_{s=0} + \left.\frac{\partial L}{\partial \dot{q}^i}\,\frac{\partial \dot{q}^i}{\partial s}\right|_{s=0} \qquad \text{(By Lagrange)} \\ &= \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\left.\frac{\partial q^i}{\partial s}\right|_{s=0}\right) \end{aligned} \tag{2.20}$$

This tells us that the quantity

$$Q = \left.\frac{\partial L}{\partial \dot{q}^i}\frac{\partial q^i}{\partial s}\right|_{s=0} \tag{2.21}$$

is constant for all time. This is our conserved quantity. □

Noether's theorem links two rather different ideas: symmetry and conservation. Moreover, the theorem is constructive. It doesn't just tell us there there is some conserved quantity associated to a symmetry. It gives us an explicit expression for it, namely (2.21). Therein lies its power.

Below, we will explain how Noether's theorem relates to the conservation of momentum, angular momentum, and energy. At heart, these results are only mildly different from the stories of conservation of (generalised) momentum and energy that we've just seen, albeit highlighting the symmetry aspect somewhat more.

However, as we proceed deeper into the laws of physics, we will see more intricate and subtle examples of Noether's theorem at play. In fact, it turns out that all conservation laws that we know of in the Standard Model and General Relativity are related to symmetries through Noether's theorem. This includes the conservation of electric charge and the conservation of particles such as protons and neutrons, collectively known as baryons.

Finally, a couple of caveats. First, Noether's theorem necessarily holds for *continuous* symmetries, meaning those that are parameterised by a parameter $s \in \mathbf{R}$. There are also *discrete* symmetries in nature which don't depend on a continuous parameter. For example, many theories are invariant under reflection, also known as *parity*, which acts as $\mathbf{r}_i \to -\mathbf{r}_i$. These types of symmetries do not give rise to conservation laws in classical physics (although they do in quantum physics).

The second caveat is that although all known conservation laws in Nature arise, via Noether's theorem, from conservation laws, it's not quite true that this is the only way in which conservation laws can arise. There is a second way in which things can be conserved, associated to topology. We won't mention this in these lectures, but you can read more in the lectures on Solitons.

With these comments and caveats in place, let's now turn to some simple applications of Noether's theorem.

Homogeneity of Space and Momentum

Consider the closed system of N particles discussed in Section 1 with Lagrangian

$$L = \tfrac{1}{2}\sum_{i=1}^{N} m_i \dot{\mathbf{r}}_i^2 - V(|\mathbf{r}_i - \mathbf{r}_j|) \tag{2.22}$$

This Lagrangian has the symmetry of translation. This means that the Lagrangian is unchanged for any $\mathbf{r}_i \to \mathbf{r}_i + s\mathbf{n}$ for any vector $\mathbf{n}$ and for any real number s, so that

$$L(\mathbf{r}_i, \dot{\mathbf{r}}_i, t) = L(\mathbf{r}_i + s\mathbf{n}, \dot{\mathbf{r}}_i, t)$$

This is the statement that space is homogeneous and a translation of the system by $s\mathbf{n}$ does nothing to the equations of motion. These translations are elements of the Galilean group that we met in section 1.2. From Noether's theorem, we can compute the conserved quantity associated with translations. It is

$$Q = \sum_i \frac{\partial L}{\partial \dot{\mathbf{r}}_i} \cdot \mathbf{n} = \sum_i \mathbf{p}_i \cdot \mathbf{n} \tag{2.23}$$

which we recognise as the the total linear momentum in the direction $\mathbf{n}$. Since this holds for all $\mathbf{n}$, we conclude that $\sum_i \mathbf{p}_i$ is conserved. But this is very familiar. It is simply the conservation of total linear momentum. To summarise

$$\begin{aligned} \text{Homogeneity of Space} &\Rightarrow \text{Translation Invariance of } L \\ &\Rightarrow \text{Conservation of Total Linear Momentum} \end{aligned}$$

This statement should be intuitively clear. One point in space is much the same as any other. So why would a system of particles speed up to get over there, when here is just as good? This manifests itself as conservation of linear momentum.

Isotropy of Space and Angular Momentum

The isotropy of space is the statement that a closed system, described by the Lagrangian (2.22), is invariant under rotations around an axis $\hat{\mathbf{n}}$, so all $\mathbf{r}_i \to \mathbf{r}'_i$ are rotated by the same amount. To work out the corresponding conserved quantities it will suffice to work with the infinitesimal form of the rotations

$$\mathbf{r}_i \to \mathbf{r}_i + \delta\mathbf{r}_i = \mathbf{r}_i + \alpha\hat{\mathbf{n}} \times \mathbf{r}_i \tag{2.24}$$

where α is considered infinitesimal. To see that this is indeed a rotation, you could calculate the length of the vector and notice that it's preserved to linear order in α. Then we have

$$L(\mathbf{r}_i, \dot{\mathbf{r}}_i) = L(\mathbf{r}_i + \alpha\hat{\mathbf{n}} \times \mathbf{r}_i, \dot{\mathbf{r}}_i + \alpha\hat{\mathbf{n}} \times \dot{\mathbf{r}}_i)$$

which gives rise to the conserved quantity

$$\sum_i \frac{\partial L}{\partial \dot{\mathbf{r}}_i} \cdot (\hat{\mathbf{n}} \times \mathbf{r}_i) = \sum_i \hat{\mathbf{n}} \cdot (\mathbf{r}_i \times \mathbf{p}_i) = \hat{\mathbf{n}} \cdot \mathbf{L}$$

This is the component of the total angular momentum in the direction $\hat{\mathbf{n}}$. Since the vector $\hat{\mathbf{n}}$ is arbitrary, we get the result

$$\begin{aligned} \text{Isotropy of Space} &\Rightarrow \text{Rotational Invariance of } L \\ &\Rightarrow \text{Conservation of Total Angular Momentum} \end{aligned}$$

Homogeneity of Time and Energy

Finally, we turn to the symmetry of time translation. The laws of physics are the same today as they were yesterday. Indeed, as far as we can tell, the laws of physics are the same today as they were during the big bang. In mathematical language, this means L is invariant under $t \to t + s$ or, in other words, $\partial L/\partial t = 0$.

But we already saw in (2.19) that any system with $\partial L/\partial t = 0$ has a conserved energy

$$E = \dot{q}^i(\partial L/\partial \dot{q}^i) - L$$

This is telling us that the existence of a conserved quantity which we call energy can be traced to the homogeneous passage of time. Putting this together with our previous result, we see that

Time is to Energy as Space is to Momentum

Recall from the lectures on Special Relativity that energy and 3-momentum fit together to form a 4-vector which rotates under spacetime transformations. Here we see that the link between energy-momentum and time-space exists even in the non-relativistic framework of Newtonian physics. You don't have to be Einstein to see it. You just have to be Emmy Noether.

2.4 Lagrangians for the Laws of Physics

Having developed these tools, we now apply them to a few examples. The examples that we've chosen in this section all come from the fundamental laws of physics: gravity, electromagnetism, special relativity and, as we will see, a kind of warm-up example for general relativity. If, instead, you're more interested in mechanical kinds of problems, like pendulum, or beads sliding on wires, then you will have to wait until Section 2.5 where we first develop the machinery of constraints.

2.4.1 Gravity: The Two Body Problem

We start by looking afresh at something that we studied in some detail in our previous course: two objects interacting through a central force.

The most famous example of this type is the gravitational interaction between two particles. Applied to objects in our solar system, this leads to the elliptic orbits of planets and the hyperbolic orbits of comets. Our first task is to frame this famous physics problem in the Lagrangian setting.

We take the two particles to have masses m_1 and m_2, and positions $\mathbf{r}_1$ and $\mathbf{r}_2$. We start by rewriting the Lagrangian in terms of the centre of mass $\mathbf{R}$, defined by $(m_1 + m_2)\mathbf{R} = m_1\mathbf{r}_1 + m_2\mathbf{r}_2$, and the separation $\mathbf{r}_{12} = \mathbf{r}_1 - \mathbf{r}_2$. For now, we will work with an arbitrary potential of the form $V(|\mathbf{r}_{12}|)$

$$L = \tfrac{1}{2}m_1\dot{\mathbf{r}}_1^2 + \tfrac{1}{2}m_2\dot{\mathbf{r}}_2^2 - V(|\mathbf{r}_{12}| = \tfrac{1}{2}(m_1 + m_2)\dot{\mathbf{R}}^2 + \tfrac{1}{2}\mu\dot{\mathbf{r}}_{12}^2 - V(|\mathbf{r}_{12}|) \qquad (2.25)$$

where

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

is the *reduced mass*. The Lagrangian splits into a piece describing the centre of mass $\mathbf{R}$ and a piece describing the separation. We'd already anticipated that this would happen back in (1.17). From now on we neglect the centre of mass piece and focus on the separation.

We know from Noether's theorem that, because the Lagrangian has rotational symmetry, the angular momentum

$$\mathbf{J} = \mathbf{r}_{12} \times \mathbf{p}_{12}$$

is conserved, where $\mathbf{p}_{12}$ is the momentum conjugate to $\mathbf{r}_{12}$. (In Section 1, and in most other courses, we called the angular momentum vector $\mathbf{L}$. But in the world of

Lagrangians, the letter L has already been taken. The letter $\mathbf{J}$ is our fall back option.) Since $\mathbf{J}$ is perpendicular to $\mathbf{r}_{12}$, the motion of the orbit must lie in a plane perpendicular to $\mathbf{J}$. This allows us to turn the 3d problem described by the Lagrangian (2.25) into a 2d problem. Using polar coordinates (r, ϕ) in the plane, where $r = |\mathbf{r}_{12}|$, the Lagrangian becomes

$$L = \tfrac{1}{2}\mu(\dot{r}^2 + r^2\dot{\phi}^2) - V(r) \tag{2.26}$$

To make further progress, notice that ϕ is ignorable so we can once again invoke Noether's theorem. This time it tells us that the magnitude of angular momentum

$$J = \mu r^2 \dot{\phi} \tag{2.27}$$

is conserved. To figure out the motion we calculate Lagrange's equation for r from (2.26)

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{r}}\right) - \frac{\partial L}{\partial r} = \mu\ddot{r} - \mu r\dot{\phi}^2 + \frac{\partial V}{\partial r} = 0$$

We then eliminate $\dot{\phi}$ from this equation by writing it in terms of the constant J to get a differential equation for the orbit purely in terms of r,

$$\mu\ddot{r} = -\frac{\partial}{\partial r}V_{\text{eff}}(r)$$

where the effective potential is given by

$$V_{\text{eff}}(r) = V(r) + \frac{J^2}{2\mu r^2}$$

The last term is known as the *angular momentum barrier*.

There is an important point here: we must substitute the conserved quantity J into the equations of motion. If you substitute J for $\dot{\phi}$ directly into the Lagrangian (2.26), you will derive an equation that looks like the one above, but you'll get a minus sign wrong! This is because the Euler-Lagrange equations are derived under the assumption that r and ϕ are independent and this is no longer the case if you impose at the Lagrangian level that r and ϕ are related by (2.27).

If, however, we really want a Lagrangian that describes the dynamics for some fixed angular momentum J, then we can simply work with the effective potential,

$$L = \frac{1}{2}\mu\dot{r}^2 - V_{\text{eff}}(r)$$

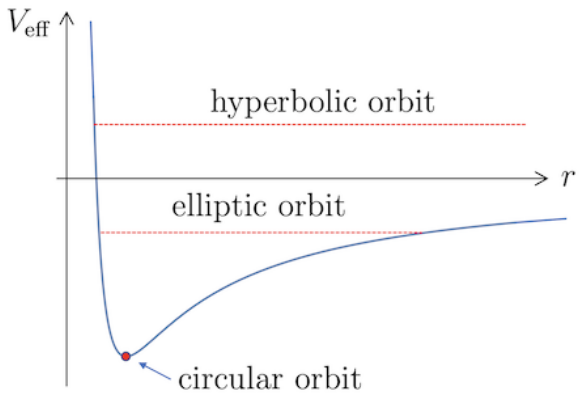


Figure 5: The effective potential for two bodies interacting gravitationally.

The next step is to note that, because $\partial L/\partial t = 0$, energy is conserved. This means that

$$E = \tfrac{1}{2}\mu\dot{r}^2 + V_{\text{eff}}(r) \tag{2.28}$$

is a constant of motion. We can invert this to get an expression for the time in terms of the distance,

$$t - t_0 = \sqrt{\frac{\mu}{2}} \int \frac{dr}{\sqrt{E - V_{\text{eff}}(r)}} \tag{2.29}$$

If we succeed in writing the solution to a problem in terms of an integral like this then we say we've "reduced the problem to quadrature". It's kind of a cute way of saying we can't do the integral. But at least we have an expression for the solution that we can play with or, if all else fails, we can simply plot numerically.

elliptic orbit circular orbit hyperbolic orbit

Up to this point the analysis is for an arbitrary potential $V(r)$. At this point let's specialise to the case of two bodies interacting gravitationally with

$$V(r) = -\frac{Gm_1m_2}{r} \tag{2.30}$$

where G is Newton's constant. For this potential, the different solutions were studied in our previous Dynamics and Relativity course, where Kepler's laws were derived. The orbits fall into two categories: elliptic if $E < 0$ and hyperbolic if $E > 0$ as shown in figure 5.

It's worth noting the methodology we used to solve this problem. We started with 6 degrees of freedom describing the positions of two particles. Translational invariance allowed us to separate out the centre of mass, reducing this to 3 degrees of freedom describing the distance between the particles. We then used rotational invariance to deduce the existence of the conserved angular momentum $\mathbf{J}$, using the conservation of the direction of $\mathbf{J}$ to reduce to 2 degrees of freedom (r and ϕ), and conservation of the magnitude $|\mathbf{J}|$ to reduce to a single variable r. Finally the fact that the Lagrangian is independent of time meant that E is also conserved and this allowed us to solve the problem. This nicely illustrates just how important symmetries, and their attendant conservation laws, are in helping us solve problems.

Finally, this is a good time to point out that the Lagrangian approach isn't a magic bullet that allows us to solve problems without any work. Pretty much all the steps described above, including the methodology of focussing on conservation laws, follow the analysis of our earlier Dynamics and Relativity course. The Lagrangian merely allows us to capture the equations of motion in a compact form, and then identify the conservation laws through symmetries. Arguably, for this example, this isn't a huge advantage. But it really comes into its own as examples get more complicated, where the Lagrangian is sometimes significantly simpler than the equations of motion.

2.4.2 Gravity Again: The Restricted Three Body Problem

As our next example, we stick with the gravitational theme, but this time we up the ante and consider three particles instead of two. Now the problem becomes more challenging. In fact, in general the problem of three or more particles interacting through gravity does not have a closed solution and we must resort to numerical methods. However, it turns out that, in particular limits we can make progress with approximations. That is what we'll focus on here.

We'll take our particles to have masses m_1, m_2 and m_3. The limit of interest here is when $m_3 \ll m_1, m_2$. Then it is a good approximation to first solve for the motion of m_1 and m_2 interacting alone, and then subsequently solve for the motion of m_3 in the time dependent potential set up by m_1 and m_2. We will now see how this works.

For simplicity, let's assume m_1 and m_2 are in a circular orbit with, the angular coordinate increasing as $\phi = \omega t$. We saw in the previous section that the circular orbit occurs for $\partial V_{\text{eff}}/\partial r = 0$, from which we get an expression relating the angular velocity of the orbit to the distance

$$\omega^2 = \frac{G(m_1 + m_2)}{r^3}$$

which is a special case of Kepler's third law. Let's further assume that m_3 moves in the same plane as m_1 and m_2. (This is pretty good assumption for the sun-earth-moon system.) To solve for the motion of m_3 in this background, we use our ability to change coordinates. We will go to a frame that rotates with m_1 and m_2, with the centre of mass at the origin.

We saw in Section 2.2.1 how to write down the Lagrangian in a rotating frame. We pick our frame so that the first particle, with mass m_1, is a distance $r\mu/m_1$ from the origin, while the second is a distance $r\mu/m_2$ from the origin in the opposite direction. We can then write down the Lagrangian for the third particle, following (2.16)

$$L = \tfrac{1}{2}m_3\left[(\dot{x}-\omega y)^2 + (\dot{y}+\omega x)^2\right] - V$$

Here V is the gravitational potential for the third particle interacting with the first two

$$V = -\frac{Gm_1m_3}{r_{13}} - \frac{Gm_2m_3}{r_{23}}$$

where the separations are given by

$$r_{13}^2 = (x + r\mu/m_1)^2 + y^2 \quad , \quad r_{23}^2 = (x - r\mu/m_2)^2 + y^2 \tag{2.31}$$

Be aware that x and y are the dynamical coordinates in this system, while r is the fixed separation between m_1 and m_2. The equations of motion arising from L are

$$\begin{aligned} m_3\ddot{x} &= 2m_3\omega\dot{y} + m_3\omega^2 x - \frac{\partial V}{\partial x} \\ m_3\ddot{y} &= -2m_3\omega\dot{x} + m_3\omega^2 y - \frac{\partial V}{\partial y} \end{aligned} \tag{2.32}$$

The full solutions to these equations are interesting and complicated. In fact, in 1889, Poincaré studied the restricted three-body system and discovered the concept of chaos in dynamical systems for the first time. (There is a cute story in which Poincaré's discovery led him to win 2,500 krona and lose 3,500 krona). Here we'll be less ambitious and try to find solutions of the form $\dot{x} = \dot{y} = 0$. This is where the third body sits stationary to the other two and the whole system rotates together. Physically, this arises because the centrifugal force of the third body exactly cancels its gravitational force. The equations we have to solve are

$$m_3\omega^2 x = \frac{\partial V}{\partial x} = Gm_1m_3\frac{x + r\mu/m_1}{r_{13}^3} + Gm_2m_3\frac{x - r\mu/m_2}{r_{23}^3} \tag{2.33}$$

$$m_3\omega^2 y = \frac{\partial V}{\partial y} = Gm_1m_3\frac{y}{r_{13}^3} + Gm_2m_3\frac{y}{r_{23}^3} \tag{2.34}$$

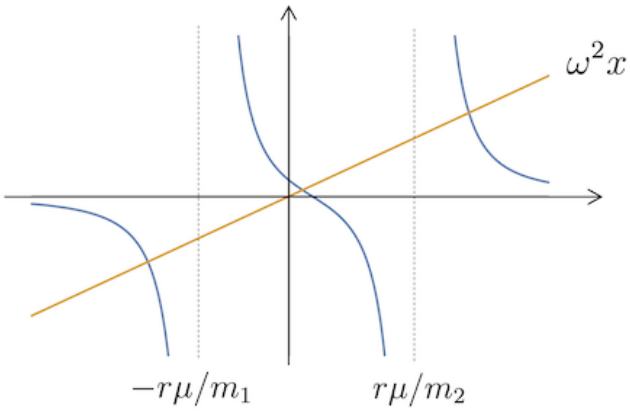


Figure 6: The three solutions sitting on $y = 0$.

As we'll now show, there are five solutions to these equations. First, suppose that $y = 0$ so that the particle with mass m_3 sits on the same line as those with masses m_1 and m_2. Then we have to solve the algebraic equation

$$\omega^2 x = Gm_1 \frac{x + r\mu/m_1}{|x + r\mu/m_1|^3} + Gm_2 \frac{x - r\mu/m_2}{|x - r\mu/m_2|^3} \tag{2.35}$$

There are three solutions to this equation. To see this, we can plot the left-hand and right-hand sides of these equations, as shown in Figure 6. This shows that there is one solutions sit in each of the regimes,

$$x < -\frac{r\mu}{m_1} \quad , \quad -\frac{r\mu}{m_1} < x < \frac{r\mu}{m_2} \quad , \quad x > \frac{r\mu}{m_2} \tag{2.36}$$

Next, we look for solutions with $y \neq 0$. From (2.34) we have

$$\frac{Gm_2}{r_{23}^3} = \omega^2 - \frac{Gm_1}{r_{13}^3}$$

which we can substitute into (2.33). After a little algebra, we find the condition for solutions to be

$$\omega^2 = \frac{G(m_1 + m_2)}{r_{13}^3} = \frac{G(m_1 + m_2)}{r_{23}^3} \tag{2.37}$$

which means that we must have $r_{13} = r_{23} = r$. This means that the three particles form an equilateral triangle. There are two such points.

The five stationary points are shown in Figure 7. They are known as *Lagrange points*. The three solutions (2.36) that have $y = 0$ are called L_1, L_2 and L_3. Those with $y \neq 0$, that solve (2.37), are called L_4 and L_5.

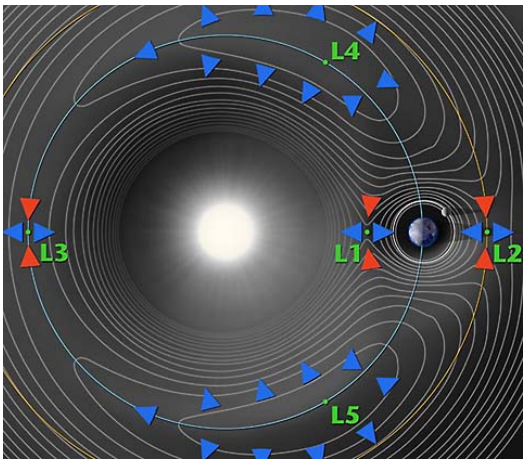


Figure 7: The five Lagrange points in the Earth-Sun system. (This figure was made as part of the educational outreach programme for the WMAP satellite.)

For the Earth-Sun system, we make good use of the Lagrange points L_1 and L_2 to place satellites in orbits that are a constant distance from the Sun, and hence at a constant temperature. Solar observatories sit at L_1, between us and the Sun. Meanwhile, more cosmologically minded satellites sit at L_2. In the past, this was home to WMAP and PLANCK, important missions that measured the cosmic microwave background radiation in exquisite detail. (You can read more about what they found in the lectures on Cosmology.) The James Webb Telescope is one of the current inhabitants of L_2. We'll return to the Lagrange points in Section 2.6.3 where we analyse their stability.

The Lagrange point L_3 hasn't yet found any use in science, but it has found plenty of use in science fiction. It's a popular location for authors to place a planet, hidden at all times from the Earth by the Sun.

2.4.3 A Particle in an Electromagnetic Field

We turn now to electromagnetism. Of course, if we're just interested in the Coulomb force between particles then we've already done the work in the previous section because, famously, the Coulomb force has the same inverse-square form as the force of gravity. Here, instead, we'll think of electromagnetism as an external force acting on a particle.

Our goal is to write down the Lagrangian that describes a particle carrying electric charge q moving in a background electric field $\mathbf{E}$ and magnetic field $\mathbf{B}$. As we now explain, there's actually a small surprise waiting for us in this Lagrangian.

Recall from our course on Electromagnetism that the electric field $\mathbf{E}$ and the magnetic field $\mathbf{B}$ can be written in terms of a vector potential $\mathbf{A}(\mathbf{x}, t)$ and a scalar potential $\phi(\mathbf{x}, t)$,

$$\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t} \quad \text{and} \quad \mathbf{B} = \nabla \times \mathbf{A} \tag{2.38}$$

We will show that the force on a particle of electric charge q is captured by the Lagrangian

$$L = \frac{1}{2}m\dot{\mathbf{x}}^2 - q\phi + q\dot{\mathbf{x}} \cdot \mathbf{A} \tag{2.39}$$

The surprise is that this Lagrangian depends on the potentials ϕ and $\mathbf{A}$, rather than the electric and magnetic field themselves. When we first encounter these potentials in electromagnetism, they're nothing more than a useful trick to allow us to solve for two of the four Maxwell equations. However, as we look at more advanced theories, both ϕ and $\mathbf{A}$ become more and more indispensable. The Lagrangian approach to electromagnetism is the first place we see this: it's not possible to write down a Lagrangian in terms of $\mathbf{E}$ and $\mathbf{B}$. Instead, we're obliged to use the potentials ϕ and $\mathbf{A}$.

Next, we compute the canonical momentum of the particle. It is

$$\mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{x}}} = m\dot{\mathbf{x}} + q\mathbf{A} \tag{2.40}$$

Here there's another minor surprise. The canonical momentum $\mathbf{p}$ is *not* the usual mass times velocity. It is altered in the presence of a magnetic field. We'll say a little more about this below.

From here, it's straightforward to derive the Euler-Lagrange equation. We have

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\mathbf{x}}}\right) - \frac{\partial L}{\partial \mathbf{x}} = \frac{d}{dt}\left(m\dot{\mathbf{x}} + q\mathbf{A}\right) + q\nabla\phi - q\nabla(\dot{\mathbf{x}} \cdot \mathbf{A}) = 0$$

To disentangle this, it's useful to work with indices $a, b = 1, 2, 3$ on the Cartesian coordinates, and rewrite the equation of motion as

$$m\ddot{x}^a = -q\left(\frac{\partial \phi}{\partial x^a} + \frac{\partial A_a}{\partial t}\right) + q\left(\frac{\partial A_b}{\partial x^a} - \frac{\partial A_a}{\partial x^b}\right)\dot{x}^b$$

Happily, the potentials ϕ and $\mathbf{A}$ have arranged themselves back into the combinations $\mathbf{E}$ and $\mathbf{B}$ in this expression. This is simplest to see if we rewrite out original definition (2.38) in index form, so it reads

$$E_a = -\frac{\partial \phi}{\partial x^a} - \frac{\partial A_a}{\partial t} \quad \text{and} \quad B_c = \epsilon_{cab}\frac{\partial A_b}{\partial x^a} \quad \Rightarrow \quad \frac{\partial A_b}{\partial x^a} - \frac{\partial A_a}{\partial x^b} = \epsilon_{abc}B_b$$

The equation of motion then becomes

$$m\ddot{x}^a = eE_a + q\epsilon_{abc}\dot{x}^b B_c$$

or, reverting to vector notation,

$$m\ddot{\mathbf{x}} = q\left(\mathbf{E} + \dot{\mathbf{x}} \times \mathbf{B}\right) \tag{2.41}$$

This is the Lorentz force law that we met in our previous course on Classical Mechanics.

We saw from the beginning that the Lagrangian formulation works is well adapted to conservative forces that can be written in terms of a potential. This means that it's no surprise to see the potential ϕ in the Lagrangian (2.39) giving rise to the electric part of the Lorentz force law (2.4.3). But we also mentioned that the Lagrangian formalism is no good at dealing with friction-type forces which are velocity dependent, like $F = -k\dot{x}$. Here we see that there are other velocity-dependent forces, such as that arising from a magnetic field, that can be nicely captured in the Lagrangian framework. Part of the reason for this is that the magnetic part of the Lorentz force law does no work, so energy remains conserved. But, in addition, there is the subtlety that we already mentioned that the Lagrangian necessarily depends on $\mathbf{A}$ and cannot be expressed in terms of the original magnetic field $\mathbf{B}$, even though the final equation of motion only depends on $\nabla \times \mathbf{A} = \mathbf{B}$.

The fact that the Lagrangian (2.39) depends on ϕ and $\mathbf{A}$ might ring alarm bells. This is because neither ϕ nor $\mathbf{A}$ are uniquely defined. We're always at liberty to change these by transformations of the form

$$\phi \to \phi - \frac{\partial \chi}{\partial t} \quad , \quad \mathbf{A} \to \mathbf{A} + \nabla\chi \tag{2.42}$$

for any function $\chi(\mathbf{x}, t)$. This is known as a *gauge transformation*. Using the definitions (2.38), it's straightforward to show that the electric and magnetic fields $\mathbf{E}$ and $\mathbf{B}$ are unchanged under gauge transformations. In fact, it should be the case that nothing physical depends on the "choice of gauge", meaning the choice of the function χ.

We've already seen that the equation of motion (2.4.3) derived from our Lagrangian (2.39) is independent of gauge transformations, since it depends only on $\mathbf{E}$ and $\mathbf{B}$. But we might be nervous about the fact that the Lagrangian itself does depend on χ. We can, however, alleviate these nerves. Under a gauge transformation (3.66), we have

$$L \to L + q\frac{\partial \chi}{\partial t} + q\dot{\mathbf{x}} \cdot \nabla\chi = L + q\frac{d\chi}{dt}$$

The Lagrangian changes by a total derivative. This means that the action $S = \int dt\ L$ changes only by a constant which explains why the equations of motion are independent of the function χ.

Note, however, that the canonical momentum (2.40) does depend on the choice of gauge. This means that the momentum $\mathbf{p}$ can't have a physical meaning because it takes different values depending on which potentials you choose to describe the electric and magnetic fields. Meanwhile, the usual momentum $m\dot{\mathbf{x}}$, sometimes called the mechanical momentum in this context, is physical but isn't conjugate to the position $\mathbf{x}$. At this point, this may sound like nothing more than a pedantic mathematical observation, but it has many interesting consequences that arise when we describe the quantum mechanics of particles in magnetic fields.

An Example: A Particle in a Uniform Magnetic Field

We can build some intuition for some of the peculiarities of the Lagrangian by looking at the simple example of a particle moving in a uniform magnetic field $\mathbf{B} = (0, 0, B)$. First, let's remind ourselves of the physics, and we'll then see what this looks like from the Lagrangian perspective. The Lorentz force law reads

$$\ddot{x} = -qB\dot{y} \quad , \quad \ddot{y} = qB\dot{x} \quad , \quad \ddot{z} = 0 \tag{2.43}$$

The particle travels at some constant speed in the z-direction. This is boring and we'll ignore it. More interesting is the motion in the (x, y)-plane. The equations are solved by

$$x = x_0 + R\sin(\omega t) \quad \text{and} \quad y = y_0 + R\cos(\omega t) \tag{2.44}$$

Here x_0, y_0 and R are all integration constants, while the frequency ω is given by

$$\omega = \frac{qB}{m}$$

This is known as the *cyclotron frequency*. We see that the particle goes in circles in the (x, y)-plane.

Now let's return to the Lagrangian. Obviously we will ultimately end up with the same equations of motion (2.43) and solution (2.44). But our interest here is what happens in the intermediate steps. In particular, to build the Lagrangian we much first find a vector potential $\mathbf{A}$ such that $\mathbf{B} = \nabla \times \mathbf{A}$. There are many choices, related by gauge transformations. For example, we could pick

$$\mathbf{A} = (-By, 0, 0) \tag{2.45}$$

At first glance, there's something a little disconcerting about this. Our magnetic field is uniform, so it looks as if our system has translational invariance in all directions. But the choice of the vector potential (2.45) is not translationally invariant in the y-direction. In fact, it's simple to convince yourself that there is no choice of $\mathbf{A}$ that respects all the symmetric of $\mathbf{B}$.

Given the relationship between translational invariance and momentum conservation that may raise some alarm bells. It means that choosing (2.45) has the consequence that p_y is *not* conserved, even though the magnetic field itself is translationally invariant in the y-direction.

From (2.40), the momenta $\mathbf{p}$ are

$$p_x = m\dot{x} - qBy \quad \text{and} \quad p_y = m\dot{y}$$

The fact that $\mathbf{A}$ in (2.45) depends on y means that p_y is not conserved. But that's ok! After all, we've seen that the particles go in circles (2.44), so it's certainly true that $\dot{y}$ is not a constant of motion. In fact, the lack of p_y conservation turns out to be a feature, not a bug. This is because, as we've seen above, the solution is that the particle travels in circles. But $\mathbf{A}$ is independent of x, so p_x is conserved. But $p_x = m\dot{x} + qA_x$ and gets the additional contribution from A_x. So while it's certainly true that $\dot{x}$ is not a constant of motion, it's simple to check that the canonical momentum p_x is a constant of motion when evaluated on the circle solution (2.44).

The upshot is that the need to write the Lagrangian in terms of $\mathbf{A}$, rather than $\mathbf{B}$, tallies nicely with Noether's theorem. A naive application of Noether's theorem might suggest that momentum is conserved when the magnetic field is translationally invariant. In fact, things are a little more subtle.

Lagrangians for Fields

In these lectures, we restrict ourselves to the classical dynamics of particles. But, ultimately, the laws of physics do not describe particles: they describe fields. One of the most important aspects of the principle of least action is that it generalises very easily to fields.

For particles, the dynamical degrees of freedom are things like the position $\mathbf{x}(t)$ which, as the notation shows, depend on time. Correspondingly, the action is an integral over time so that it ascribes a number to each particle trajectory. Fields, however, are functions of space and time. The most familiar examples are the electric and magnetic fields $\mathbf{E}(\mathbf{x}, t)$ and $\mathbf{B}(\mathbf{x}, t)$. Now we want the action to ascribe a number to every possible field configuration. This means that the action should be an integral over time and space.

We won't describe the action principle for fields in these lectures. But, in passing, I will just mention that it is very straightforward. In particular, the Maxwell equations

follow from the action

$$S = \int dt d^3x \ \left(\frac{\epsilon_0}{2}\mathbf{E}^2 - \frac{1}{2\mu_0}\mathbf{B}^2 \right)$$

where ϵ_0 and μ_0 are the constants of nature with pretentious names that describe the strength of the electric and magnetic forces respectively. Importantly, the action above should be varied with respect to the potentials ϕ and $\mathbf{A}$, rather than the electric and magnetic fields themselves. This then reproduces the Maxwell equations. All of this will be explained in great detail in Section 5 of the lectures on Electromagnetism.

2.4.4 Special Relativity

Our focus in these lectures is very much on Newtonian mechanics. But, as we know, when the speed of particles is some fraction of the speed of light c, we must replace Newtonian mechanics with the framework of special relativity. In this section, we explain how to describe relativistic particles using Lagrangians and actions. I should warn you that I won't recap much of special relativity here: you might want to refresh a few things from the lectures on Dynamics and Relativity. (Or from the lectures on Electromagnetism which contain a section essentially identical to this one.)

We can start by writing down a naive action principle for a free, relativistic particle. (For once, we'll write down actions rather than Lagrangians; the reason for this will become clear as we proceed.) We can consider

$$S[\mathbf{x}(t)] = -mc^2 \int dt \ \sqrt{1 - \frac{\dot{\mathbf{x}}^2}{c^2}} \tag{2.46}$$

This looks promising for a few reasons. First, if we Taylor expand the action we get

$$S[\mathbf{x}(t)] = \int dt \left[-mc^2 + \frac{1}{2}m\dot{\mathbf{x}}^2 + \ldots \right]$$

The first term is just a constant and doesn't affect the equation of motion, while the second term is the familiar action for a free, non-relativistic particle. The ... are then corrections that kick in as the particle gets close to the speed of light. Next, we can look at the canonical momentum. It is

$$\mathbf{p}(t) = \frac{\partial S}{\partial \dot{\mathbf{x}}(t)} = m\gamma\dot{\mathbf{x}}(t) \quad \text{with} \quad \gamma = \sqrt{\frac{1}{1 - \dot{\mathbf{x}}^2/c^2}}$$

This is indeed the right momentum for a relativistic particle, replete with the familiar gamma factor. This means that we get the right equation of motion for a free particle, namely $d\mathbf{p}/dt = 0$. We could now think about adding further forces to the action, such as the coupling to electric and magnetic fields that we described in Section 2.4.3. We will again see that the action (2.46) gives the right answer.

All of this means that there's a lot to like about the simple action (2.46). It gives the right answers. However, there's also something more than a little unsatisfactory about it, because it doesn't seem to embrace the spirit of special relativity. Recall that the essence of special relativity is that there is a symmetry, captured by Lorentz transformations, that relate different observers. Moreover, we've been bragging in this section about how well suited the Lagrangian approach is to symmetries. Yet there's no hint of the Lorentz symmetry in the action (2.46). This is because the action is written for an observer in a very particular reference frame, with a very particular choice of the time coordinate t.

It would be nice to do better and write down an action principle that holds for observers in any reference frame. And we can, but it's going to take a little bit of work.

First, we can see why it might be tricky to write down an action that holds in any reference frame. We know from special relativity that the position $\mathbf{x}$ of a particle sits in a 4-vector X^μ with $\mu = 0, 1, 2, 3$ with components

$$X^\mu = (ct, \mathbf{x})$$

Ideally we want to write down an action that depends on the trajectory X^μ that the particle makes in spacetime. To do this, we first need to parameterise the worldline of the particle. We'll call this parameter σ and think of the trajectory of the particle as tracing out a worldline $X^\mu(\sigma)$ in spacetime.

But this means that the degrees of freedom for our relativistic particle are very different from those of a non-relativistic particle. For a non-relativistic particle, the action is a function of $\mathbf{x}(t)$. But for a relativistic particle, the action is a function of $\mathbf{x}(\sigma)$ and $t(\sigma)$ for some parameter σ. This means that we have one more degree of freedom, $t(\sigma)$, for the relativistic particle. In addition, we've got to introduce this new parameter σ that describes where we are on the worldline. Yet, despite these differences, the relativistic action should reduce to the non-relativistic action when the speed of a particle is small. How can we achieve all these things?

The answer to this is rather nice. It's simplest if we just write down the action and then explore its properties. The better action for a relativistic particle is a functional of the trajectory $X^\mu(\sigma)$, given by

$$S[X^\mu(\sigma)] = -mc \int_{\sigma_1}^{\sigma_2} d\sigma \, \sqrt{\eta_{\mu\nu} \frac{dX^\mu}{d\sigma} \frac{dX^\nu}{d\sigma}} \tag{2.47}$$

Here $\eta_{\mu\nu}$ is the Minkowski metric which we take to have signature

$$\eta_{\mu\nu} = \text{diag}(+1, -1, -1, -1)$$

There are a number of interesting properties of (2.47). First, it is manifestly invariant under Lorentz transformations. These are symmetries of the form

$$X^\mu \to \Lambda^\mu{}_\nu X^\nu \quad \text{with} \quad \Lambda^\rho{}_\mu \eta_{\rho\sigma} \Lambda^\sigma{}_\nu = \eta_{\mu\nu}$$

and you can see by inspection that this leaves the action (2.47) unchanged. For this reason, (2.47) is sometimes known as the *covariant action*.

The action S is actually closely related to something familiar from the world of special relativity: it is proportional to the proper time experienced by the particle. Recall that a particle moving along a worldline $X^\mu(\sigma)$, experience a proper time

$$\tau(\sigma) = \frac{1}{c}\int_\sigma^0 d\sigma' \, \sqrt{\eta_{\mu\nu}\frac{dX^\mu}{d\sigma'}\frac{dX^\nu}{d\sigma'}} \tag{2.48}$$

In special relativity, the proper time is maximised by a particle that does *not* accelerate. This fact is famous from the twin paradox where the dull stay-at-home twin ages fastest. Here it sits nicely with the fact that the proper time is identified with the action and hence is extremised on solutions to the equations of motion.

In addition to Lorentz invariance, the action (2.47) has a second symmetry of a very different kind, and this is the key to understanding its properties. This second symmetry is *reparameterisation invariance*. So far, we didn't say how we picked the parameterisation σ of the worldline. Suppose that we made a different choice, and parameterised the worldline by a different coordinate $\tilde{\sigma}$, related to the first parameterisation by a monotonic function $\tilde{\sigma}(\sigma)$. Then we could equally as well construct an action $\tilde{S}$ using this new parameter, given by

$$\tilde{S} = -mc\int_{\tilde{\sigma}_1}^{\tilde{\sigma}_2} d\tilde{\sigma} \, \sqrt{\eta_{\mu\nu}\frac{dX^\mu}{d\tilde{\sigma}}\frac{dX^\nu}{d\tilde{\sigma}}}$$

We might worry that this different parameterisation will give different equations of motion. Happily this is not the case because the two actions are, in fact, identical. This follows from a quick calculation,

$$\tilde{S} = -mc\int_{\sigma_1}^{\sigma_2} d\sigma \, \frac{d\tilde{\sigma}}{d\sigma}\sqrt{\eta_{\mu\nu}\frac{dX^\mu}{d\sigma}\frac{dX^\nu}{d\sigma}\left(\frac{d\sigma}{d\tilde{\sigma}}\right)^2} = S$$

We see that the action takes the same form regardless of our choice of parameterisation. Although we've called this a "symmetry", it's not a symmetry in the same sense as Lorentz transformations. In particular, reparameterisation does not generate new solutions from old ones. Instead, it is a redundancy in the way we describe the system. It is similar to the gauge "symmetry" of electromagnetism which, despite the name, is also a redundancy rather than a symmetry.

Reparameterisation invariance has a number of consequences. The first is that it explains why the action (2.47) has only three degrees of freedom, even though it is a function of four variables $X^\mu(\sigma)$. This is because one of the degrees of freedom X^μ is not physical. Suppose that you solve the equation of motion to find a trajectory $X^\mu(\sigma)$. In most dynamical systems, each of these four functions would tell you something about the physical trajectory. But, for us, reparameterisation invariance means that there is no actual information in the value of σ. To find the physical path, we should eliminate σ to find the relationship between the X^μ. And this kills one degree of freedom.

We can see this most clearly by making a cunning choice for the parameter σ that parameterises the worldline. Suppose that we choose σ to coincide with the time t for some intertial observer: $\sigma = t$. Then $dX^0/d\sigma = c$ and the action (2.47) then becomes

$$S = -mc^2 \int_{t_1}^{t_2} dt \sqrt{1 - \frac{\dot{\mathbf{x}}^2}{c^2}} \tag{2.49}$$

where here $\dot{\mathbf{x}} = d\mathbf{x}/dt$. But this is the action (2.46) that we started this section with. So our two actions (2.47) and (2.46) are indeed equivalent, but each has different advantages. The action (2.46) makes it clear that we are dealing with a system with three degrees of freedom $\mathbf{x}$, but Lorentz invariance is hidden. Meanwhile the action (2.47) has manifest Lorentz invariance, but at the cost of introducing more degrees of freedom than are physical. But, as we've seen above, the reparameterisation invariance of the action allows us to remove the time degree of freedom and return to (2.46).

There's yet another manifestation of reparameterisation invariance. To see this, we compute the canonical momentum associated to X^μ,

$$P_\mu = \frac{\partial L}{\partial \dot{X}^\mu} = -mc \frac{1}{\sqrt{\dot{X}^\nu \dot{X}_\nu}} \dot{X}_\mu$$

where here $\dot{X}^\mu = \partial X^\mu/\partial\sigma$. You can check that this coincides with the usual definition of 4-momentum $P^\mu = m dX^\mu/d\tau$. (This follows from the fact that the proper time τ defined by (2.48) obeys $d\tau/d\sigma = L/mc^2$ with L the Lagrangian.) It's a familiar result

from special relativity that these momenta are not all independent, but obey

$$P^\mu P_\mu = m^2 c^2 \tag{2.50}$$

While this result is familiar in special relativity, it's rather surprising from the perspective of Lagrangian mechanics. In all the other examples we've met in these lectures, the canonical momenta are independent of each other. But not here. This novel feature can be traced to the existence of reparameterisation invariance, meaning that there was a redundancy in our original description. Indeed, whenever theories have such a redundancy there will be some constraint analogous to (2.50). As one final comment, note that if we expand out (2.50), we have

$$(P^0)^2 = \mathbf{p}^2 + m^2 c^2$$

In particular, we see that we must have $P^0 \neq 0$. This is important. There's nothing that tells us that we must have $\mathbf{p} \neq 0$. The particle is quite able to just sit still in space if it wants. But $P^0 \neq 0$ tells us that the particle is obliged to move in the time direction. Physically, this again reflects the fact that the action (2.47) has only three degrees of freedom, not four. Physiologically, this is why you get old.

So far, our covariant action (2.47) only describes a free, relativistic particle. But it's straightforward to couple it to relativistic forces. The most obvious such force is electromagnetism and here there is a nice interplay with relativity. It turns out that the potentials ϕ and $\mathbf{A}$ that we introduced in (2.38) sit nicely in a 4-vector $A^\mu = (\phi/c, \mathbf{A})$, known as a *gauge field*. If the particle carries electric charge q, then it couples to the electromagnetic field through the action

$$S[X^\mu(\sigma)] = \int_{\sigma_1}^{\sigma_2} d\sigma \ \left[-mc\sqrt{\eta_{\mu\nu}\frac{dX^\mu}{d\sigma}\frac{dX^\nu}{d\sigma}} - qA_\mu(X)\frac{dX^\mu}{d\sigma} \right] \tag{2.51}$$

If we again pick the worldline parameter σ to coincide with the time of some inertial observer, $\sigma = t$, then the calculation that previously led us to the action (2.49) will now include the additional terms of the electromagnetic action (2.39). You can find more details on this in the lectures on Electromagnetism.

2.4.5 A Brief Look at Curved Geometry

We now turn to a slightly more abstract, but extremely useful Lagrangian. We will consider a particle moving in N-dimensional space, and introduce generalised coordinates q^i, with $i = 1, \ldots, N$. Furthermore, we'll consider a particle with no potential energy, only kinetic energy. The novelty is that the particle is moving a on general, curved space rather than flat space that we've considered so far. How do we describe this?

There is a whole mathematical framework, known as differential geometry, to describe curved spaces. (It is covered in some detail in the lectures on General Relativity.) Here we will be very brief. The key idea is to introduce a function that describes the distance between two nearby points. We consider the point labelled by coordinates q^i and a nearby point labelled by coordinates $q^i + \delta q^i$. The the distance δs between them is

$$\delta s = \sqrt{g_{ij}(q)\, \delta q^i\, \delta q^j} \tag{2.52}$$

Here, we have introduced a matrix of functions, $g_{ij}(q)$ known as the *metric*. In principle, each element of this matrix can be a function of all N coordinates q_k, $k = 1, \ldots, N$. The metric is a symmetric matrix, so that $g_{ij} = g_{ji}$. (One way of seeing this is that there's no point including an anti-symmetric piece of g_{ij} because it won't contribute to the distance (2.52) anyway.) We will assume that $\det(g_{ij}) \neq 0$, so that the inverse matrix $(g^{-1})^{ij}$ exists, with $(g^{-1})^{ij} g_{jk} = \delta^i{}_k$.

Suppose that the particle is moving on flat, Euclidean space $\mathbf{R}^2$, with Cartesian coordinates $q^i = (x, y)$, Then the metric is simply $g_{ij} = \delta_{ij}$ and the distance above is nothing more than Pythagoras' theorem

$$\delta s = \sqrt{\delta x^2 + \delta y^2}$$

The more general form of the metric $g_{ij}(q)$ allows us to describe curved space. Note, however, that just because $g_{ij}(q)$ depends on q doesn't mean that we necessarily have a curved space. We may just be working in some other coordinates. For example, if we work with flat $\mathbf{R}^2$ in polar coordinates, so that $q^i = (r, \phi)$, then

$$g_{ij} = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix} \quad \Rightarrow \quad \delta s = \sqrt{\delta r^2 + r^2 \delta\phi^2}$$

which is the way we measure distances in polar coordinates.

With this preamble in place, we can now write down the Lagrangian for a non-relativistic particle moving on this (possibly) curved space. It takes the rather natural form

$$L = \frac{1}{2} g_{ij}(q)\, \dot{q}^i \dot{q}^j \tag{2.53}$$

Deriving the equations of motion from this Lagrangian is an exercise in keeping index notation straight. We have

$$\frac{\partial L}{\partial q^k} = \frac{1}{2} \frac{\partial g_{ij}}{\partial q^k} \dot{q}^i \dot{q}^j \quad \text{and} \quad \frac{\partial L}{\partial \dot{q}^k} = g_{ik} \dot{q}^i$$

The Euler-Lagrange equations are then

$$0 = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^k}\right) - \frac{\partial L}{\partial k} = g_{ik}\ddot{q}^i + \frac{\partial g_{ik}}{\partial q^j}\dot{q}^i\dot{q}^j - \frac{1}{2}\frac{\partial g_{ij}}{\partial q^k}\dot{q}^i\dot{q}^j$$
$$= g_{ik}\ddot{q}^i + \frac{1}{2}\left(\frac{\partial g_{ik}}{\partial q^j} + \frac{\partial g_{jk}}{\partial q^i} - \frac{\partial g_{ij}}{\partial q^k}\right)\dot{q}^i\dot{q}^j$$

where, to go from the first to the second line, we've the fact that the $\partial g/\partial q$ terms are contracted with $\dot{q}^i\dot{q}^j$ which is symmetric in i and j indices. The upshot is that the equation of motion can be written as

$$\ddot{q}^i + \Gamma^i_{jk}\dot{q}^j\dot{q}^k = 0 \quad \text{where} \quad \Gamma^i_{jk} = \frac{1}{2}(g^{-1})^{il}\left(\frac{\partial g_{jl}}{\partial q^k} + \frac{\partial g_{kl}}{\partial q^j} - \frac{\partial g_{jk}}{\partial q^l}\right) \tag{2.54}$$

In the language of differential geometry, the object Γ^i_{jk} is known as *Christoffel symbol* (or sometimes as the *Levi-Civita connection*), while the equation of motion (2.54) is called the *geodesic equation*.

In general relativity, there is a natural generalisation of the geodesic equation (2.54) to describe a particle moving in curved spacetime, rather than just curved space. This, it turns out, is the correct description of gravity.

The Lagrangian (2.53) describes a particle moving in curved space. There are also natural field theory version of this Lagrangian in which q^i are functions of both space and time. These Lagrangians have the uninspiring name of *sigma models*. They crop up all over the place, in condensed matter physics, particle physics, and even string theory.

2.5 Constraints

There is one last advantage of the Lagrangian formulation that we now explain. In the above examples, we worked with particles roaming freely throughout space, with their motion only affected by various potentials. But in many examples in Newtonian mechanics, there are various constraint forces at play that can't be described by potentials. These are things like the tension of ropes, and normal forces applied by surfaces. In the Lagrangian formulation, we don't have to worry about such things. In this section, we'll show why.

An Example: The Pendulum

We can illustrate the basic by looking at the simply pendulum of length l. This has a single dynamical degree of freedom θ, the angle the pendulum makes with the vertical.

The position of the mass m in the plane is described by two cartesian coordinates x and y subject to a constraint $x^2+y^2=l^2$. We can easily solve this constraint and parameterise the general solution by an angle θ,

$$x = l\sin\theta \quad \text{and} \quad y = l\cos\theta \tag{2.55}$$

Employing the Newtonian method to solve this system, we introduce the tension T as shown in the diagram and resolve the force vectors to find,

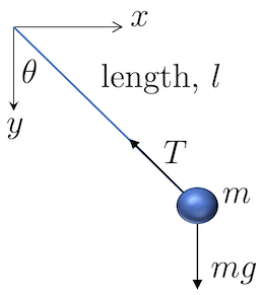


Figure 8:

$$m\ddot{x} = -\frac{Tx}{l} \qquad , \qquad m\ddot{y} = mg - \frac{Ty}{l} \tag{2.56}$$

To determine the motion of the system, we we substitute in the constrained coordinates (2.55) and easily find

$$\ddot{\theta} = -\frac{g}{l}\sin\theta \qquad , \qquad T = ml\dot{\theta}^2 + mg\cos\theta \tag{2.57}$$

While this example was pretty straightforward to solve using Newtonian methods, things get rapidly harder when we consider more complicated constraints (and we'll see plenty presently). Moreover, you may have noticed that half of the work of the calculation went into computing the tension T. On occasion we'll be interested in this. For example, we might want to know how fast we can spin the pendulum before it breaks. But often we won't care about these constraint forces, but will only want to know the motion of the pendulum itself. In this case it seems like a waste of effort to go through the motions of computing T. We'll now see how we can avoid this extra work in the Lagrangian formulation. First, we define what we mean by constraints more rigorously.

2.5.1 Holonomic Constraints

We'll take as our starting point N particles, each of which can move freely in $\mathbf{R}^3$. This means that our configuration space is parameterised by Cartesian coordinates x^A, with $A = 1, \ldots, 3N$. This isn't the most general starting point, but it will serve our purposes and the generalisations should hopefully be obvious.

Holonomic Constraints are relationships between the coordinates of the form

$$f_\alpha(x^A, t) = 0 \qquad \alpha = 1, \ldots, 3N - n \tag{2.58}$$

In general the constraints can be time dependent and our notation above allows for this. Holonomic constraints can be solved in terms of n *generalised coordinates* q^i, $i = 1, \ldots n$. So

$$x^A = x^A(q^1, \ldots, q^n) \tag{2.59}$$

The system is said to have n degrees of freedom. For the pendulum example above, the system has a single degree of freedom, $q = \theta$.

Now let's see how the Lagrangian formulation deals with constraints of this form. We introduce $3N - n$ new dynamical degrees of freedom, $\lambda^\alpha(t)$. These are called *Lagrange multipliers*. Note that each sits on the same footing as the original $x^A(t)$ in the sense that the Lagrange multipliers are also functions of time. We now define a new Lagrangian

$$L' = L(x^A, \dot{x}^A) + \lambda^\alpha f_\alpha(x^A, t) \tag{2.60}$$

where, in the second term, the repeated index means that we're summing over all $\alpha = 3N - n$. We now treat λ^α like new coordinates. Since L' doesn't depend on $\dot{\lambda}^\alpha$, Lagrange's equations for λ^α are

$$\frac{\partial L'}{\partial \lambda^\alpha} = f_\alpha(x^A, t) = 0 \tag{2.61}$$

But this is precisely the constraints that we want. It tells us that we are indeed solving a problem with the constraints imposed. However, this comes with a price, because the Euler-Lagrange equation for x^A is now

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^A}\right) - \frac{\partial L}{\partial x^A} = \lambda^\alpha \frac{\partial f_\alpha}{\partial x^A} \tag{2.62}$$

The left-hand-side is the equation of motion for the unconstrained system. The right-hand-side is the novel and captures the constraint forces in the system. We can now solve these equations as we did in the Newtonian formulation.

The Pendulum Example Again

The Lagrangian for the pendulum is given by that of a free particle moving in the plane, augmented by the Lagrange multiplier term for the constraints. It is

$$L' = \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2) + mgy + \tfrac{1}{2}\lambda(x^2 + y^2 - l^2)$$

From this we can calculate the two equations of motion for x and y,

$$m\ddot{x} = \lambda x \qquad \text{and} \qquad \ddot{y} = mg + \lambda y$$

while the equation of motion for λ reproduces the constraint $x^2 + y^2 - l^2 = 0$. Comparing with the Newtonian approach (2.56), we find that the Lagrange multiplier λ is proportional to the tension: $\lambda = -T/l$.

So we see that we can easily incorporate constraint forces into the Lagrangian setup using Lagrange multipliers. But the big news is that we don't have to! Often we don't care about the tension T or other constraint forces, but only want to know what the generalised coordinates q^i are doing. In this case we have the following useful theorem

Theorem: For constrained systems, we may derive the equations of motion directly in generalised coordinates q^i

$$L[q^i, \dot{q}^i, t] = L[x^A(q^i, t), \dot{x}^A(q^i, \dot{q}^i, t)] \tag{2.63}$$

Proof: Let's work with $L' = L + \lambda^\alpha f_\alpha$ and change coordinates to

$$x_A \to \begin{cases} q^i & i = 1, \ldots, n \\ f_\alpha & \alpha = 1, \ldots 3N - n \end{cases} \tag{2.64}$$

Here, the novelty is that we're treating the constraints $f_\alpha(x^A, t)$ as coordinates of the original system. We know from our discussion in Section 2.2 that the Euler-Lagrange equations take the same form in these new coordinates. In particular, we may look at the equations for q^i,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} = \lambda_\alpha \frac{\partial f_\alpha}{\partial q^i}$$

But, by definition, $\partial f_\alpha / \partial q^i = 0$. So we are left with the Euler-Lagrange equations purely in terms of q^i, with no sign of the constraint forces,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} = 0 \quad i = 1, \ldots, n \tag{2.65}$$

This means that if we are only interested in the dynamics of the generalised coordinates q^i, we may ignore the Lagrange multipliers and work entirely with the unconstrained Lagrangian $L(q^i, \dot{q}^i, t)$ defined in (2.63) where we just substitute in $x_A = x_A(q^i, t)$. The resulting equations of motion are then given by (2.65). □

The Pendulum Example for the Last Time

Let's see how this works in the simple example of the pendulum. We can parameterise the constraints in terms of the generalised coordinate θ so that $x = l\sin\theta$ and $y = l\cos\theta$. We now substitute this directly into the Lagrangian for a particle moving in the plane under the effect of gravity, to get

$$L = \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2) + mgy = \tfrac{1}{2}ml^2\dot{\theta}^2 + mgl\cos\theta \tag{2.66}$$

From this we may derive Lagrange's equations using the coordinate θ directly

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right) - \frac{\partial L}{\partial \theta} = ml^2\ddot{\theta} + mgl\sin\theta = 0 \tag{2.67}$$

which indeed reproduces the equation of motion for the pendulum (2.57). Note that, as promised, we haven't calculated the tension T using this method. This has the advantage that we've needed to do less work. If we need to figure out the tension, we have to go back to the more laborious Lagrange multiplier method.

Non-Holonomic Constraints

Before we turn to various examples of holonomic constraints, it's worth flagging up that there are plenty of other kinds of constraints that can't be written in holonomic form $f_\alpha(x^A, t) = 0$. Systems with non-holonomic constraints may be straightforward to solve, or they may be more complicated. But there's no general theory that covers then and you'll need to develop different methods for each problem.

Here's a simple example of a non-holonomic constraint. Consider a particle moving under gravity, constrained to move on the outside of a sphere of radius R. This means that the position of the particle is constrained by an inequality, $x^2 + y^2 + z^2 \geq R^2$, rather than an equality, which is a kind of non-holonomic constraint. We might, for example, start the particle perched on the top of the sphere, give it a slight nudge, and want to know when it falls off. It's straightforward to solve this: the particle falls off when the normal, contact force vanishes. But the theory of holonomic constraints won't help you.

Here's a second example. A coin, or radius R, is balanced on its edge, and rolls along the ground. The coin's position in the plane is described by Cartesian coordinates x and y. But the coin has two, further degrees of freedom. This is the angle θ that it makes with the x-axis, and the angle ϕ that some chosen point on the rim makes with the ground. If the coin rolls without slipping, then the velocity of the rim$v_{\rm rim} = R\dot{\phi}$. So, in terms of our four coordinates, we have the constraint

$$\dot{x} = R\dot{\phi}\sin\theta \quad \text{and} \quad \dot{y} = R\dot{\phi}\cos\theta$$

Now we have a velocity-dependent constraint. It is non-holonomic because it can't be integrated to give a constraint of the form $f(x, y, \theta, \phi) = 0$. Again, this particular example isn't hard to solve. (Indeed, we looked at very similar problems in the lectures on Dynamics and Relativity.)

2.5.2 A Bead on a Rotating Hoop

Take a hoop of radius a that rotates about the vertical axis with frequency ω. A bead of mass m is threaded on the hoop and moves without friction, as shown in the figure. We want to determine the motion of this bead.

This is an example of a system with a time dependent holonomic constraint. There is a single degree of freedom $\psi(t)$, the angle the bead makes with the vertical. In terms of Cartesian coordinates (x, y, z) the position of the bead is

$$\begin{aligned} x &= a\sin\psi\cos\omega t \\ y &= a\sin\psi\sin\omega t \\ z &= a - a\cos\psi \end{aligned}$$

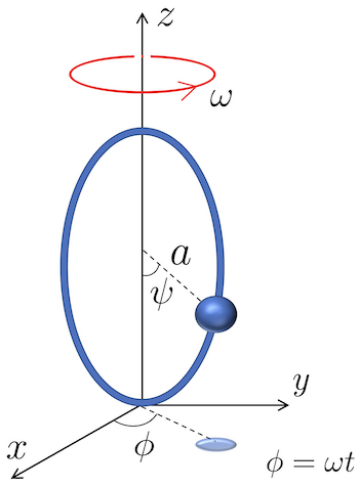


Figure 9:

To determine the Lagrangian in terms of the generalised coordinate ψ, we simply substitute these expressions into the Lagrangian for the free particle. For the kinetic energy T we have

$$T = \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) = \tfrac{1}{2}ma^2[\dot{\psi}^2 + \omega^2\sin^2\psi]$$

while the potential energy V is given by (ignoring an overall constant)

$$V = mgz = -mga\cos\psi$$

This means that our Lagrangian, as a function of ψ and $\dot{\psi}$, is given by

$$L = ma^2\left(\tfrac{1}{2}\dot{\psi}^2 - V_{\rm eff}\right)$$

where the effective potential is

$$V_{\rm eff} = \frac{1}{ma^2}\left(-mga\cos\psi - \tfrac{1}{2}ma^2\omega^2\sin^2\psi\right)$$

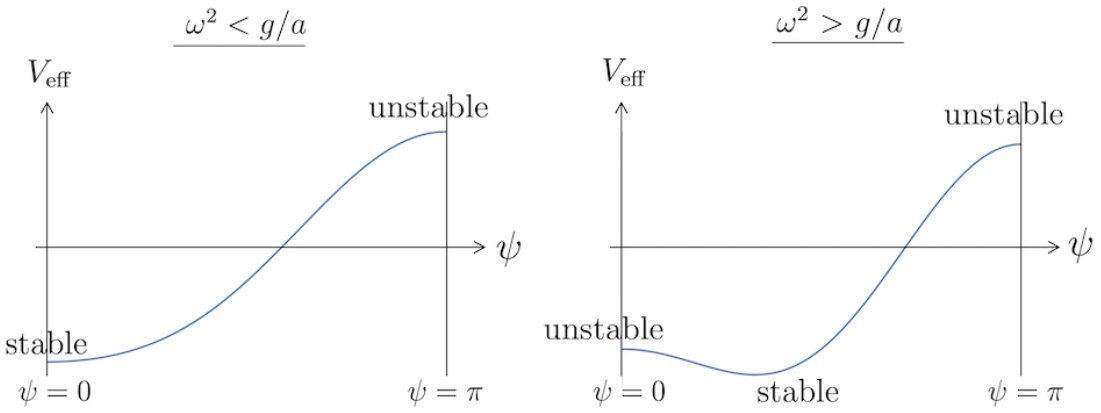


Figure 10: The effective potential for the bead depends on how fast the hoop is rotating

We can now derive the equations of motion for the bead simply from Lagrange's equations which read

$$\ddot{\psi} = -\frac{\partial V_{\text{eff}}}{\partial \psi}$$

We can look for stationary solutions of these equations in which the bead doesn't move, i.e solutions of the form $\ddot{\psi} = \dot{\psi} = 0$. From the equation of motion, we must solve $\partial V_{\text{eff}}/\partial\psi = 0$ to find that the bead can remain stationary at points satisfying

$$g \sin\psi = a\omega^2 \sin\psi \cos\psi$$

The number of solutions depends on how fast the hoop is spinning. There are always two such solutions, given by $\psi = 0$ and $\psi = \pi$ while, if the hoop is spinning fast enough, so that $\omega^2 \geq g/a$, then a third stationary point exists at $\cos\psi = g/a\omega^2$. Which of these stationary points is stable depends on whether $V_{\text{eff}}(\psi)$ has a local minimum (stable) or maximum (unstable). This in turn depends on the value of ω.

The effective potential V_{eff} is drawn in Figure 10 two values of ω, the first below the critical value, the second above. For all values of ω the bead perched at the top of the hoop at $\psi = \pi$ is unstable. For $\omega^2 < g/a$, the stable equilibrium point is at the bottom of the hoop at $\psi = 0$. But, as the hoop spins faster, at some point this position becomes unstable. For $\omega^2 > g/a$, the bead is still in equilibrium at $\psi = 0$, but it's an unstable equilibrium. Instead, a new minimum of V_{eff} appears at $\cos\psi = g/a\omega^2$ and this is the stable point.

2.5.3 The Double Pendulum

A double pendulum is drawn in the figure to the right. It consists of two particles of mass m_1 and m_2, connected by light rods of length l_1 and l_2. For the first particle, the kinetic energy T_1 and the potential energy V_1 are the same as for a simple pendulum (2.66)

$$T_1 = \tfrac{1}{2}m_1 l_1^2 \dot{\theta}_1^2 \qquad \text{and} \qquad V_1 = -m_1 g l_1 \cos\theta_1$$

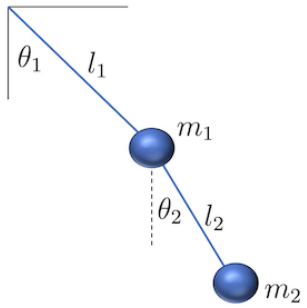


Figure 11:

For the second particle, things are a little more involved. Consider the position of the second particle in the plane of the pendulum swings. We'll give use Cartesian coordinates (x, y), with y pointing downwards, and take the origin to be the pivot of the first pendulum. We then have

$$x_2 = l_1 \sin\theta_1 + l_2 \sin\theta_2 \qquad \text{and} \qquad y_2 = l_1 \cos\theta_1 + l_2 \cos\theta_2$$

We can substitute into the kinetic energy for the second particle

$$T_2 = \tfrac{1}{2}m_2(\dot{x}^2 + \dot{y}^2) = \tfrac{1}{2}m_2\left(l_1^2\dot{\theta}_1^2 + l_2^2\dot{\theta}_2^2 + 2l_1 l_2 \cos(\theta_1 - \theta_2)\dot{\theta}_1\dot{\theta}_2\right)$$

while the potential energy is given by

$$V_2 = -m_2 g y_2 = -m_2 g\left(l_1 \cos\theta_1 + l_2 \cos\theta_2\right)$$

The Lagrangian is given by the sum of the kinetic energies, minus the sum of the potential energies: it is

$$\begin{aligned} L = {} & \tfrac{1}{2}(m_1 + m_2)l_1^2\dot{\theta}_1^2 + \tfrac{1}{2}m_2 l_2^2\dot{\theta}_2^2 + m_2 l_1 l_2 \cos(\theta_1 - \theta_2)\dot{\theta}_1\dot{\theta}_2 \\ & + (m_1 + m_2)g l_1 \cos\theta_1 + m_2 g l_2 \cos\theta_2 \end{aligned} \tag{2.68}$$

The equations of motion now follow in the usual way by computing the two Euler-Lagrange equations, one for θ_1 and one for θ_2. It turns out that the solutions to these equations are complicate and, above a certain energy, the motion is chaotic. We'll return briefly to the double pendulum in Section 2.6 where we look at the decidedly non-chaotic motion when the energy is very small.

2.5.4 The Spherical Pendulum

A spherical pendulum consists of a heavy particle, attached to a light, rigid rod that is free to rotate in three dimensions, as shown in the figure. The system has two degrees of freedom which cover the range

$$0 \leq \theta < \pi \qquad \text{and} \qquad 0 \leq \phi < 2\pi$$

In terms of cartesian coordinates, we have

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

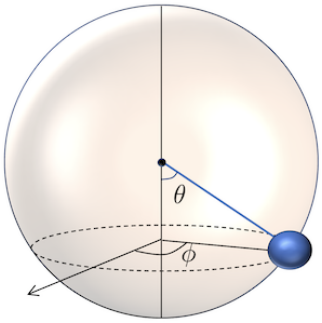


Figure 12:

We substitute these constraints into the Lagrangian for a free particle to get

$$L = \frac{1}{2}m(\dot{x}^2+\dot{y}^2+\dot{z}^2) - mgz = \frac{1}{2}ml^2(\dot{\theta}^2+\dot{\phi}^2\sin^2\theta) + mgl\cos\theta$$

The first thing to note is that the coordinate ϕ is ignorable. We know from our discussion in Section 2.3 that the corresponding conserved quantity is is given by

$$J = \frac{\partial L}{\partial\dot{\phi}} = ml^2\dot{\phi}\sin^2\theta \tag{2.69}$$

This is the component of angular momentum in the ϕ direction. Meanwhile, the equation of motion for θ follows from the Euler-Lagrange equation. It is

$$ml^2\ddot{\theta} = ml^2\dot{\phi}^2\sin\theta\cos\theta - mgl\sin\theta$$

We can substitute $\dot{\phi}$ for the constant J in this expression to get an equation entirely in terms of θ which we chose to write as

$$\ddot{\theta} = -\frac{\partial V_{\rm eff}}{\partial\theta} \tag{2.70}$$

where the effective potential is defined to be

$$V_{\rm eff}(\theta) = -\frac{g}{l}\cos\theta + \frac{J^2}{2m^2l^4}\frac{1}{\sin^2\theta}$$

Let me reiterate a warning that we gave already when discussing the two body problem: do *not* substitute $J = \mu r^2\dot{\phi}$ directly into the Lagrangian – you will get a minus sign wrong! You must substitute it into the equations of motion.

As well as the conservation of angular momentum J, we also have $\partial L/\partial t = 0$ so energy is conserved. This is given by

$$E = \tfrac{1}{2}\dot{\theta}^2 + V_{\rm eff}(\theta) \tag{2.71}$$

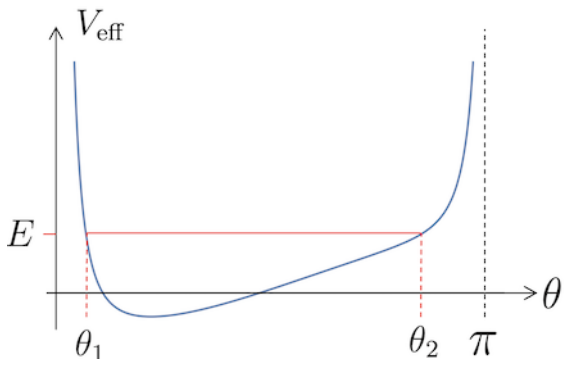


Figure 13: The effective potential for the spherical pendulum.

where E is a constant. In fact we can invert this equation for E to solve for θ in terms of an integral

$$t - t_0 = \frac{1}{\sqrt{2}} \int \frac{d\theta}{\sqrt{E - V_{\text{eff}}(\theta)}}$$

Once we have an expression for $\theta(t)$ we can solve for $\phi(t)$ using the expression for J,

$$\phi = \int \frac{J}{ml^2} \frac{1}{\sin^2\theta} dt = \frac{J}{\sqrt{2}ml^2} \int \frac{1}{\sqrt{E - V_{\text{eff}}(\theta)}} \frac{1}{\sin^2\theta} d\theta$$

which gives us $\phi = \phi(\theta) = \phi(t)$.

To get more of a handle on what these solutions look like, we plot the function V_{eff} in Figure 13. For a given energy E, the particle is restricted to the region $V_{\text{eff}} \leq E$ (a fact which follows from (2.71)). So from the figure we see that the motion is pinned between two points θ_1 and θ_2. If we draw the motion of the pendulum in real space, it must therefore look something like the figure on the right, in which the bob oscillates between the two extremes: $\theta_1 \leq \theta \leq \theta_2$.

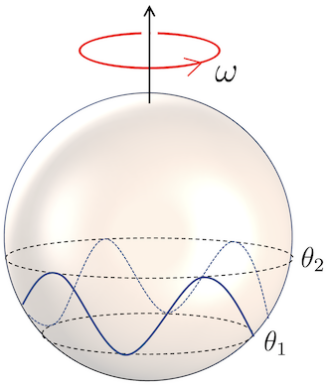


Figure 14:

Note that we could make more progress in understanding the motion of the spherical pendulum than for the double pendulum, where we essentially stopped after writing down the Lagrangian. The reason for this is the existence of two conservation laws for the spherical pendulum, energy and angular momentum, compared to just one, the energy, for the double pendulum.

There is a stable orbit which lies between the two extremal points at $\theta = \theta_0$, corresponding to the minimum of $V_{\rm eff}$. This occurs if we balance the angular momentum J and the energy E just right. We can look at small oscillations around this point by expanding $\theta = \theta_0 + \delta\theta$. Substituting into the equation of motion (2.70), we have

$$\delta\ddot{\theta} = -\left(\frac{\partial^2 V_{\rm eff}}{\partial\theta^2}\bigg|_{\theta=\theta_0}\right)\delta\theta + \mathcal{O}(\delta\theta^2)$$

so small oscillations about $\theta = \theta_0$ have frequency $\omega^2 = (\partial^2 V_{\rm eff}/\partial\theta^2)$ evaluated at $\theta = \theta_0$.

2.6 Small Oscillations and Stability

> "Physics is that subset of human experience which can be reduced to coupled harmonic oscillators"
>
> *Michael Peskin*

The harmonic oscillator is by far the most important system in all of physics. Partly, this is because we can solve it. And partly it's because a large number of other systems can, in certain circumstances, be made to look like a bunch of harmonic oscillators. Indeed, the art of physics often boils down to figuring out how to make things look like harmonic oscillators.

In this section we'll see one reason why the simple harmonic oscillator is so important. We will study the motion of systems close to equilibrium and see that the dynamics is described by n *decoupled* simple harmonic oscillators, each ringing at a different frequency.

Let's start with a single degree of freedom $x(t)$. We've already seen several examples where we get an equation of the form

$$\ddot{x} = f(x)$$

An *equilibrium point*, $x = x_0$ satisfies $f(x_0) = 0$. This means that if we start with the initial conditions

$$x = x_0 \qquad \text{and} \qquad \dot{x} = 0$$

then the system will stay there forever. But what if we start slightly away from $x = x_0$? To analyse this, we write

$$x(t) = x_0 + \eta(t)$$

where η is assumed to be small so that we can Taylor expand $f(x)$ to find

$$\ddot{\eta} = f'(x_0)\,\eta + \mathcal{O}(\eta^2) \tag{2.72}$$

We neglect the terms quadratic in η and higher. There are three possible behaviours of this system

- $f'(x_0) < 0$: In this case the restoring force sends us back to $\eta = 0$ and the solution is

  $$\eta(t) = A\cos(\omega(t - t_0))$$

  where A and t_0 are integration constants, while $\omega^2 = -f'(x_0)$. The system undergoes stable oscillations about $x = x_0$ at frequency ω.

- $f'(x_0) > 0$: In this case, the force pushes us away from equilibrium and the solution is

  $$\eta(t) = Ae^{\lambda t} + Be^{-\lambda t}$$

  where A and B are integration constants, while $\lambda^2 = f'(x_0)$. In this case, there is a very special initial condition $A = 0$ such that $x \to x_0$ at late times. But for generic initial conditions, η gets rapidly large and the approximation that η is small breaks down. We say the system has a *linear instability*.

- $f'(x_0) = 0$: In this case, it was a bad idea to neglect the quadratic terms in (2.72) since they are where the physics lies. This means that we have to go back and do more work.

We now generalise this discussion to N degrees of freedom with equations of motion of the form,

$$\ddot{q}^i = f^i(q^1, \ldots, q^N) \qquad i = 1, \ldots, N \tag{2.73}$$

An equilibrium point q_0^i must satisfy $f_i(q_0^1, \ldots, q_0^N) = 0$ for all $i = 1 \ldots, N$. Consider small perturbations away from the equilibrium point

$$q^i(t) = q_0^i + \eta^i(t) \tag{2.74}$$

where, again, we take the η^i to be small so that we can Taylor expand the f^i, and neglect the quadratic terms and higher. We have

$$\ddot{\eta}^i \approx \left.\frac{\partial f^i}{\partial q^j}\right|_{q=q_0} \eta^j \tag{2.75}$$

where the sum over $j = 1, \ldots, N$ is implicit. It's useful to write this in matrix form. We define the vector $\boldsymbol{\eta}$ and the $N \times N$ matrix F as

$$\boldsymbol{\eta} = \begin{pmatrix} \eta^1 \\ \vdots \\ \eta^N \end{pmatrix} \quad , \quad F = \begin{pmatrix} \partial f^1/\partial q^1 & \ldots & \partial f^1/\partial q^N \\ \vdots & & \vdots \\ \partial f^N/\partial q^1 & \ldots & \partial f^N/\partial q^N \end{pmatrix}$$

where each partial derivative in the matrix F is evaluated at $q^i = q^i_0$. The equation (2.75) now becomes simply

$$\ddot{\boldsymbol{\eta}} = F\boldsymbol{\eta} \tag{2.76}$$

Our strategy is simple: we search for eigenvectors of F. If F were a symmetric matrix, it would have a complete set of orthogonal eigenvectors with real eigenvalues. Unfortunately, there's no reason to think that F is symmetric. Nonetheless, it is true that for equations of the form (2.76) arising from physical Lagrangian systems, the eigenvalues will be real. We shall postpone a proof of this fact for a couple of paragraphs and continue under the assumption that F has real eigenvalues.

In general, F will have different left and right eigenvectors,

$$F\boldsymbol{\mu}_a = \lambda_a^2 \boldsymbol{\mu}_a \qquad \text{and} \qquad \boldsymbol{\zeta}_a^T F = \lambda_a^2 \boldsymbol{\zeta}_a^T \qquad a = 1, \ldots, N \tag{2.77}$$

where there's no sum over a in these equations. Note that although the eigenvectors differ, the eigenvalues λ_a^2 for $a = 1, \ldots, n$ are the same. Note also that, somewhat unusually, we've denoted the eigenvalues as square numbers, λ_a^2. This is to make the expressions below simpler. However, the fact that the eigenvalue is λ_a^2 does *not* imply that it's a positive number: it could be positive or negative (or, indeed, zero).

By combining the two expressions above, we have

$$\boldsymbol{\zeta}_a^T F \boldsymbol{\mu}_b = \lambda_a^2 \boldsymbol{\zeta}_a \cdot \boldsymbol{\mu}_b = \lambda_b^2 \boldsymbol{\zeta}_a \cdot \boldsymbol{\mu}_b$$

If all eigenvalues are distinct, so $\lambda_a^2 \neq \lambda_b^2$, then the expression above tells us that we necessarily have $\boldsymbol{\zeta}_a \cdot \boldsymbol{\mu}_b = \delta_{ab}$. If F has degenerate eigenvalues, it turns out that it's always possible to pick eigenvectors so that $\boldsymbol{\zeta}_a \cdot \boldsymbol{\mu}_b = \delta_{ab}$.

With the eigenvectors and eigenvalues to hand, the most general solution to (2.77) is

$$\boldsymbol{\eta}(t) = \sum_a \boldsymbol{\mu}_{a=1}^N \left[A_a e^{\lambda_a t} + B_a e^{-\lambda_a t} \right]$$

where A_a and B_a are $2N$ integration constants. Again, we have three possibilities for each eigenvalue

- $\lambda_a^2 < 0$: In this case $\lambda_a = \pm i\omega_a$ for some real number ω_a. The system will be stable in the corresponding direction $\boldsymbol{\eta} = \boldsymbol{\mu}_a$.

- $\lambda_a^2 > 0$: Now $\pm\lambda_a$ are real and the system exhibits a linear instability in the direction $\boldsymbol{\eta} = \boldsymbol{\mu}_a$.

- $\lambda_a^2 = 0$: In this case, the linearised analysis that we've done here is insufficient to tell us the full story.

The eigenvectors $\boldsymbol{\mu}_a$ are called *normal modes*. The equilibrium point is only stable if $\lambda_a^2 < 0$ for every $a = 1, \ldots, N$. If this holds, then the system will oscillate around the equilibrium point as a linear superposition of all the normal modes, each typically vibrating at a different frequency.

To keep things real, we can write the most general solution as

$$\boldsymbol{\eta}(t) = \sum_{a\,:\,\lambda_a^2>0} \boldsymbol{\mu}_a \left[A_a e^{\lambda_a t} + B_a e^{-\lambda_a t}\right] + \sum_{a\,:\,\lambda_a^2<0} \boldsymbol{\mu}_a A_a \cos(\omega_a(t - t_a))$$

where now A_a, B_a and t_a are the $2N$ integration constants.

The Reality of the Eigenvalues

Finally, let's show what we put off above: that the eigenvalues λ_a^2 are real for matrices F derived from a physical Lagrangian system. Consider a general Lagrangian of the form,

$$L = \tfrac{1}{2}T_{ij}(q)\dot{q}^i\dot{q}^j - V(q)$$

We will require that $T_{ij}(q)$ is invertible and positive definite for all q. Here the matrix T_{ij} plays the same role as the metric in Section 2.4.5. Expanding about an equilibrium point as in (2.74), to linear order in η_i the equations read

$$T_{ij}\ddot{\eta}_j = -V_{ij}\eta_j$$

where $T_{ij} = T_{ij}(q^0)$ and $V_{ij} = \partial^2 V/\partial q^i \partial q^j$, also evaluated at $q^i = q_0^i$. Then in the matrix notation of (2.76), we have $F = -T^{-1}V$. Both T_{ij} and V_{ij} are symmetric, but not necessarily simultaneously diagonalisable. This means that F_{ij} is not necessarily symmetric. Nevertheless, F does have real eigenvalues. To see this, look at

$$F\boldsymbol{\mu} = \lambda^2\boldsymbol{\mu} \quad \Rightarrow \qquad V\boldsymbol{\mu} = -\lambda^2\, T\boldsymbol{\mu} \tag{2.78}$$

So far, both $\boldsymbol{\mu}$ and λ^2 could be complex. We will now show that they're not. Take the inner product of this equation with the complex conjugate eigenvector $\boldsymbol{\mu}^\dagger$,

$$\boldsymbol{\mu}^\dagger V \boldsymbol{\mu} = \lambda^2 \, \boldsymbol{\mu}^\dagger T \boldsymbol{\mu}$$

But for any symmetric matrix S, the quantity $\boldsymbol{\mu}^\dagger S \boldsymbol{\mu}$ is real, a fact that follows from expanding $\boldsymbol{\mu}$ in the complete set of real, orthogonal eigenvectors of S, each of which has a real eigenvalue. Therefore both $\boldsymbol{\mu}^\dagger V \boldsymbol{\mu}$ and $\boldsymbol{\mu}^\dagger T \boldsymbol{\mu}$ are both real. Since we have assumed that T is invertible and positive definite, we know that $\boldsymbol{\mu}^\dagger T \boldsymbol{\mu} 0$ so, from (2.78), we conclude that the eigenvalue λ^2 is indeed real.

We now illustrate these ideas with a couple of examples.

2.6.1 The Double Pendulum Revisited

We derived the Lagrangian for the double pendulum in Section 2.5.3. Restricting to the case where the two masses are the same $m_1 = m_2 = m$ and the two lengths are the same $l_1 = l_2 = l$, we derived the Lagrangian (2.68) for arbitrary oscillations

$$L = ml^2\dot{\theta}_1^2 + \tfrac{1}{2}ml^2\dot{\theta}_2^2 + ml^2\cos(\theta_1 - \theta_2)\dot{\theta}_1\dot{\theta}_2 + 2mgl\cos\theta_1 + mgl\cos\theta_2$$

The stable equilibrium point is clearly $\theta_1 = \theta_2 = 0$. (You could check mathematically if you're sceptical). Let's expand for small θ_1 and θ_2. If we want to linearise the equations of motion for θ, then we must expand the Lagrangian to second order (so that after we take derivatives, there's still a θ left standing). We have

$$L \approx ml^2\dot{\theta}_1^2 + \tfrac{1}{2}ml^2\dot{\theta}_2^2 + ml^2\dot{\theta}_1\dot{\theta}_2 - mgl\theta_1^2 - \tfrac{1}{2}mgl\theta_2^2$$

where we've thrown away an irrelevant constant. From this we can use Lagrange's equations to derive the two linearised equations of motion

$$\begin{aligned} 2ml^2\ddot{\theta}_1 + ml^2\ddot{\theta}_2 &= -2mgl\theta_1 \\ ml^2\ddot{\theta}_2 + ml^2\ddot{\theta}_1 &= -mgl\theta_2 \end{aligned}$$

Or, writing $\boldsymbol{\theta} = (\theta_1, \theta_2)^T$, this becomes

$$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \ddot{\boldsymbol{\theta}} = -\frac{g}{l}\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}\boldsymbol{\theta} \qquad \Rightarrow \qquad \ddot{\boldsymbol{\theta}} = -\frac{g}{l}\begin{pmatrix} 2 & -1 \\ -2 & 2 \end{pmatrix}\boldsymbol{\theta}$$

We have two eigenvectors. They are

1. $\boldsymbol{\mu}_1 = (1, \sqrt{2})^T$ which has eigenvalue $\lambda_1^2 = -(g/l)(2 - \sqrt{2})$. This corresponds to the motion shown on left-hand side of Figure 15.

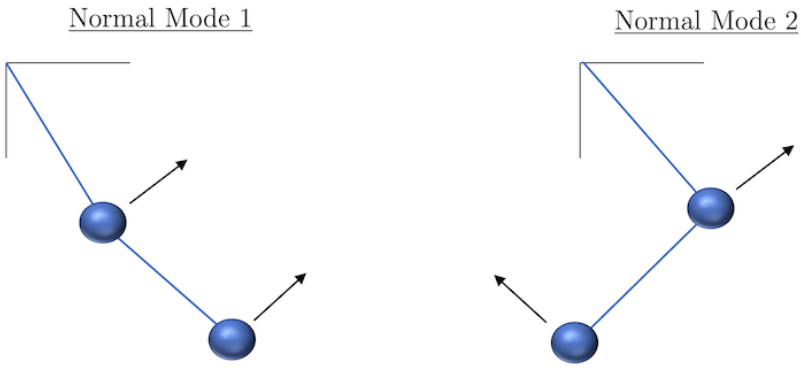


Figure 15: The two normal modes of the double pendulum

2. $\boldsymbol{\mu}_2 = (1, -\sqrt{2})^T$ which has eigenvalue $\lambda_2^2 = -(g/l)(2+\sqrt{2})$. This corresponds to the motion shown on the right-hand side of Figure 15.

We see that the frequency of the mode in which the two rods oscillate in different directions should be higher than that in which they oscillate together.

2.6.2 The Linear Triatomic Molecule

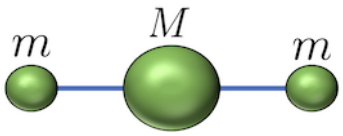


Figure 16:

Consider a linear molecule, comprised of three atoms. The other two have mass m, the middle one has mass M. An example is shown in the figure. This can be viewed as a rough approximation of CO_2. We'll only consider motion in the direction parallel to the molecule for each atom, in which case the Lagrangian for this molecule takes the form,

$$L = \tfrac{1}{2}m\dot{x}_1^2 + \tfrac{1}{2}M\dot{x}_2^2 + \tfrac{1}{2}m\dot{x}_3^2 - V(x_1 - x_2) - V(x_2 - x_3)$$

The function V is some rather complicated interatomic potential. But, the point of this section is that if we're interested in oscillations around equilibrium, this doesn't matter. Assume that $x_i = x_i^0$ in equilibrium. By symmetry, we have $|x_1^0 - x_2^0| = |x_2^0 - x_3^0| = r_0$. We write deviations from equilibrium as

$$x_i(t) = x_i^0 + \eta_i(t)$$

Taylor expanding the potential about the equilibrium point,

$$V(r) = V(r_0) + \left.\frac{\partial V}{\partial r}\right|_{r=r_0}(r - r_0) + \tfrac{1}{2}\left.\frac{\partial^2 V}{\partial r^2}\right|_{r=r_0}(r - r_0)^2 + \ldots$$

Here the first term $V(r_0)$ is a constant and can be ignored, while the second term $\partial V/\partial r$ vanishes since we are in equilibrium. Substituting into the Lagrangian, we have

$$L \approx \tfrac{1}{2}m\dot{\eta}_1^2 + \tfrac{1}{2}M\dot{\eta}_2^2 + \tfrac{1}{2}m\dot{\eta}_3^2 - \frac{k}{2}\left[(\eta_1-\eta_2)^2 + (\eta_2-\eta_3)^2\right]$$

where $k = \partial^2 V/\partial r^2$ evaluated at $r = r_0$. Then the equations of motion are

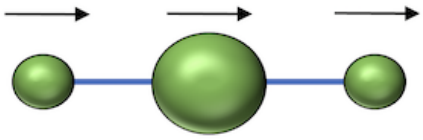
$$\begin{pmatrix} m\ddot{\eta}_1 \\ M\ddot{\eta}_2 \\ m\ddot{\eta}_3 \end{pmatrix} = -k \begin{pmatrix} \eta_1 - \eta_2 \\ (\eta_2-\eta_1) + (\eta_2-\eta_3) \\ \eta_3 - \eta_2 \end{pmatrix}$$

or, putting it in the form $\ddot{\boldsymbol{\eta}} = F\boldsymbol{\eta}$, we have

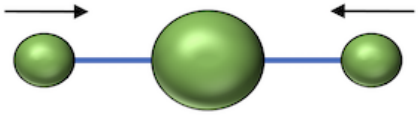
$$F = \begin{pmatrix} -k/m & k/m & 0 \\ k/M & -2k/M & k/M \\ 0 & k/m & -k/m \end{pmatrix}$$

Again, we must look for eigenvectors of F. There are three:

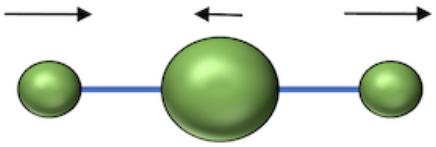
1. $\boldsymbol{\mu} = (1,1,1)^T$ which has eigenvalue $\lambda_1^2 = 0$. But this is just an overall translation of the molecule. It's not an oscillation.



2. $\boldsymbol{\mu}_2 = (1,0,-1)^T$ which has eigenvalue $\lambda_2^2 = -k/m$. In this motion, the outer two atoms oscillate out of phase, while the middle atom remains stationary. The oscillation has frequency $\omega_2 = \sqrt{k/m}$.



3. $\boldsymbol{\mu}_3 = (1,-2m/M,1)^T$ which has eigenvalue $\lambda_3^2 = -(k/m)(1+2m/M)$. This oscillation is a little less obvious. The two outer atoms move in the same direction, while the middle atom moves in the opposite direction. The frequency of this vibration $\omega_3 = \sqrt{-\lambda_3^2}$ is greater than that of the second normal mode.



For small deviations from equilibrium, the most general motion is a superposition of all of these modes.

$$\boldsymbol{\eta}(t) = \boldsymbol{\mu}_1(A + Bt) + \boldsymbol{\mu}_2 C\cos(\omega_2(t - t_2)) + \boldsymbol{\mu}_3 D\cos(\omega_3(t - t_3))$$

with A, B, C and D all arbitrary integration constants, albeit ones that should be taken to be suitably small.

2.6.3 The Stability of Lagrange Points

In Section 2.4.2, we studied the three-body problem, at least in the restricted sense. Recall that we consider a light particle, with mass m_3, moving in the gravitational potential of two much heavier particles, with masses $m_1, m_2 \gg m_3$.

For simplicity, we took the two heavy particles to orbit in a circle of radius r. The frequency of their orbit is, by Kepler's third law,

$$\omega^2 = \frac{G(m_1 + m_2)}{r^3}$$

In the reference frame that rotates at the same frequency ω, we derived the equations of motion for the third particle (2.32),

$$\begin{aligned} m_3\ddot{x} &= 2m_3\omega\dot{y} + m_3\omega^2 x - \frac{\partial V}{\partial x} \\ m_3\ddot{y} &= -2m_3\omega\dot{x} + m_3\omega^2 y - \frac{\partial V}{\partial y} \end{aligned} \tag{2.79}$$

Here the potential experienced by the third particle is given by

$$V = -\frac{Gm_1m_3}{r_{13}} - \frac{Gm_2m_3}{r_{23}}$$

where the separations are given by

$$r_{13}^2 = (x + r\mu/m_1)^2 + y^2 + z^2 \quad \text{and} \quad r_{23}^2 = (x - r\mu/m_2)^2 + y^2 + z^2$$

which contains the reduced mass $\mu = m_1m_2/(m_1 + m_2)$.

In Section 2.4.2, we found the stationary solutions to (2.79), satisfying $\dot{\mathbf{x}} = \ddot{\mathbf{x}} = 0$. There are five such solutions, known as Lagrange points. Three of these points, L_1, L_2, and L_3, lie on the $y = 0$ axis, co-linear with the original orbiting particles. The other two, L_4 and L_5, lie in the orbital plane, but with $y \neq 0$.

Our goal in this section is to understand the stability of these Lagrange points. Suppose that you sit just away from them. Do you get pushed further away, or slowly drift back?

To answer this, first recall that, as part of our derivation in Section 2.4.2, we assumed that the third particle was moving in the same orbital plane as the first two which, in our notation, is the $z = 0$ plane. It's simple to check that the particle is stable to perturbations out of this plane. The additional equation of motion is just

$$\ddot{z} = -\frac{\partial^2 V}{\partial z^2}$$

It's simple to check that $\partial^2 V/\partial z^2\big|_{z=0} > 0$, which means that each of the Lagrange points is at least stable to perturbations out of the $z = 0$ plane. But what about within the plane?

This calculation is a little more involved. We expand

$$x = x_0 + \delta x \quad \text{and} \quad y = y_0 + \delta y$$

Here (x_0, y_0) is one of the Lagrange points: we know that these obey $m_3\omega^2\mathbf{x}_0 = \nabla V\big|_{\mathbf{x}_0}$. Substituting this into the equation of motion (2.79), and expanding to leading order in $\delta\mathbf{x}$, we have

$$\begin{aligned} \delta\ddot{x} - 2\omega\delta\dot{y} &= \omega^2\delta x - \frac{\partial^2 V}{\partial x^2}\delta x - \frac{\partial^2 V}{\partial x\partial y}\delta y \\ \delta\ddot{y} + 2\omega\delta\dot{x} &= \omega^2\delta y - \frac{\partial^2 V}{\partial y^2}\delta y - \frac{\partial^2 V}{\partial x\partial y}\delta x \end{aligned} \tag{2.80}$$

At this stage we make an ansatz for the time dependence,

$$\delta x = \delta x_0 e^{\lambda t} \quad \text{and} \quad \delta y = \delta y_0 e^{\lambda t}$$

Here $\delta\mathbf{x}_0$ is the amplitude of the perturbation which is left arbitrary (but necessarily small). We know from the general analysis of this section that the system will be unstable if λ is real, but will oscillate about the equilibrium point if λ is complex.

Substituting this into the perturbed equation of motion (2.80) gives us the matrix equation

$$\mathcal{M}\begin{pmatrix}\delta x\\ \delta y\end{pmatrix}=0 \quad \text{with} \quad \mathcal{M}=\begin{pmatrix}\lambda^2-\omega^2+A & -2\omega\lambda+B\\ 2\omega\lambda+B & \lambda^2-\omega^2+C\end{pmatrix} \tag{2.81}$$

where we've defined the partial derivatives

$$A=\frac{\partial^2 V}{\partial x^2} \quad , \quad B=\frac{\partial^2 V}{\partial x\partial y} \quad , \quad C=\frac{\partial^2 V}{\partial y^2}$$

each of which should be evaluated at the particular Lagrange point. The expressions for these partial derivatives are easily found, but are a little messy,

$$\begin{aligned}
A &= \frac{Gm_1(y^2-2(x+r\mu/m_1)^2)}{r_{13}^5}+\frac{Gm_2(y^2-2(x-r\mu/m_2)^2)}{r_{23}^5}\\
B &= -\frac{3Gm_1y(x+r\mu/m_1)}{r_{13}^5}-\frac{3Gm_2y(x-r\mu/m_2)}{r_{23}^5}\\
C &= \frac{Gm_1((x+r\mu/m_1)^2-2y^2)}{r_{13}^5}+\frac{Gm_2((x-r\mu/m_2)^2-2y^2)}{r_{23}^5}
\end{aligned}$$

where we've taken the liberty of setting $z=0$ since each Lagrange point sits in the plane. We'll see below how to extract what we need from these expressions.

The matrix equation (2.81) is like an eigenvalue equation: we should solve it at each Lagrange point for λ. We can do this by taking the determinant,

$$\det\mathcal{M}=(\lambda^2-\omega^2+A)(\lambda^2-\omega^2+C)+4\omega^2\lambda^2-B^2=0 \tag{2.82}$$

This is a quadratic equation in λ^2, so that solutions come in $\pm\lambda$ pairs. An equilibrium point will be stable only if λ is purely imaginary, which means that all the roots λ^2 must be real and negative for stability.

The story is slightly different for the three co-linear Lagrange points L_1, L_2 and L_3 and the other two. We'll start with the co-linear points. These sit at $y=0$, which means that the partial derivatives are

$$B=0 \quad \text{and} \quad A=-2C \quad \text{with} \quad C=\frac{Gm_1}{(|x+r\mu/m_1|^3}+\frac{Gm_2}{|x-r\mu/m_2|^3}>0$$

Solving the quadratic (2.82) in this case gives

$$\lambda^2=\frac{1}{2}\left(C-2\omega^2\pm\sqrt{C(9C-8\omega^2)}\right)$$

The roots are clearly real only if $9C \geq 8\omega^2$. But, for stability, we also require that both roots are negative and this gives the extra condition

$$\frac{8}{9}\omega^2 \leq C < \omega^2$$

So does this condition hold? The answer is no. To see this, we need to use the fact, derived in (2.35), that at these co-linear Lagrange points obey

$$\omega^2 x = Gm_1 \frac{x + r\mu/m_1}{|x + r\mu/m_1|^3} + Gm_2 \frac{x - r\mu/m_2}{|x - r\mu/m_2|^3}$$

Dividing through by x, we can use this to write

$$C = \omega^2 + \frac{Gr\mu}{x}\left(\frac{1}{|x - r\mu/m_2|^3} - \frac{1}{|x + r\mu/m_1|^3}\right)$$

But it's not hard to show that the extra term is always positive, so $C > \omega^2$. For example, at the Lagrange point L_2, we have $x > r\mu/m_2 > 0$ and the term in brackets above is positive. Similar arguments hold for L_1 and L_3. This means that the co-linear Lagrange points are all unstable.

We learned in Section 2.4.2 that L_1 and L_2 are particularly useful points for placing satellites. Now we see that they are unstable. Without doing any further work, we can see that the time scale of the instability is set by $\omega = (\text{year})^{-1}$. In fact, a more accurate analysis shows that it's unstable on a time scale of less than a month. This means that satellites must be equipped with the ability to correct their position.

This leaves us with the Lagrange points L_4 and L_5, in which the three particles make an equilateral triangle, so that $r_{13} = r_{23} = r$. Some simple geometry shows that $(x + r\mu/m_1) = -(x - r\mu/m_2) = r/2$ and $y = \pm\sqrt{3}r/2$, where the $\pm$ sign are for L_4 and L_5 respectively. From this, we can explicitly calculate the values of the partial derivatives. We have

$$A = \frac{1}{4}\frac{G(m_1 + m_2)}{r^3} = \frac{1}{4}\omega^2 \quad , \quad C = -\frac{5}{4}\frac{G(m_1 + m_2)}{r^3} = -\frac{5}{4}\omega^2$$

while the mixed partial derivative depends on the ratio of masses,

$$B = -\frac{3\sqrt{3}}{4}\frac{G(m_1 - m_2)}{r^3} = -\frac{3\sqrt{3}}{4}\omega^2\left(\frac{m_1 - m_2}{m_1 + m_2}\right)$$

With these results, it's straightforward to find the roots of the quadratic (2.82). We have

$$\frac{\lambda^2}{\omega^2} = -\frac{1}{2} \pm \frac{1}{2}\sqrt{1 - \frac{27}{4}\left(1 - \left(\frac{m_1 - m_2}{m_1 + m_2}\right)^2\right)}$$

Recall that both roots should be real and negative for stability. This holds provided that one of the original masses is sufficiently smaller than that other, so that

$$\left(\frac{m_1 - m_2}{m_1 + m_2}\right)^2 > \frac{23}{27}$$

If we take $m_2 < m_1$, then the condition for stability becomes

$$\frac{m_2}{m_1 + m_2} \lesssim 0.0385$$

The Earth-Sun system comfortably obeys this, with $m_{\rm Earth}/m_{\rm Sun} \approx 10^{-5}$ and a collection of dust has gathered at L_4 and L_5. It is also obeyed by Jupiter, with $m_{\rm Jupiter}/m_{\rm Earth} \approx 0.01$. in this case, a number of large asteroids have gathered at L_4 and L_5, known as *trojans*.

2.7 A First Look at Perturbation Theory

In the previous section, we looked at small perturbations about an equilibrium point. In this section, we develop perturbation theory in a slightly different direction. Here's the basic idea. Suppose that you have a problem that you can't solve. (This, it turns out, is the normal way of things when doing research.) If you're lucky, it might be close to a problem that you can solve. The hope, then, is that you can build a solution to the problem you care about by starting with a problem that you can make progress on.

There is a lot to say about these kinds of perturbation methods. Here we just sketch the basics and highlight a simple, but important issue that arises. We'll do this by looking at a particular example.

2.7.1 The Anharmonic Oscillator

Our example of choice is the anharmonic, or non-linear oscillator. This has Lagrangian

$$L = \frac{1}{2}\dot{x}^2 - \frac{1}{2}\omega^2 x^2 - \frac{1}{4}\lambda x^4$$

where we've set the mass of the particle to unity. The equation of motion is

$$\ddot{x} = -\omega^2 x - \lambda x^3 \tag{2.83}$$

When $\lambda = 0$, this is just the usual harmonic oscillator. It's clear that the additional term shouldn't qualitatively change the dynamics of the system: for $\lambda > 0$ the extra force pushes us back towards the origin, just like the linear term. Moreover, the energy

$$E = \frac{1}{2}\dot{x}^2 + \frac{1}{2}\omega^2 x^2 + \frac{1}{4}\lambda x^4$$

is conserved and so we see that the particle will oscillate back and forth in a potential $V = \frac{1}{2}\omega^2 x^2 + \frac{1}{4}\lambda x^4$, just like a harmonic oscillator. However, we'd like to do better than this qualitative analysis and actually find a solution to (2.83). Sadly, however, the general solution cannot be written down in closed form.

Our goal here is to make progress when λ is, in some sense, small. This looks promising because we certainly know the solution when $\lambda = 0$. It is (ignoring a possible phase),

$$x_0(t) = A\cos(\omega(t - t_0)) \tag{2.84}$$

The hope is that we can build off this solution to find an approximation to the true solution when λ is small.

Our first task is to clarify what it means for λ to be small. This is because λ has dimensions $[\lambda] = L^{-2}T^{-2}$ and so can only be small relative to something else. In fact there's a unique combination of parameters that have the same dimension and so we're looking for solutions in the situation that

$$\lambda \ll \frac{\omega^2}{A^2}$$

Note that ω is a parameter of our equation of motion (2.83), but A is an integration constant for our starting solution (2.84). This reflects the fact that we're not going to find approximations to the general solution to (2.83) using this method: only to some restricted class of solutions. In fact, it's perhaps better to say that this method works for any λ, but restricted to solutions with $A^2 \ll \omega^2/\lambda$. Either way, we introduce the *dimensionless* small quantity

$$\epsilon = \frac{\lambda A^2}{\omega^2} \ll 1$$

Our task is to set up a perturbative expansion in ϵ.

The obvious way to proceed is to postulate a solution that has the expansion

$$x(t) = x_0(t) + \epsilon x_1(t) + \epsilon^2 x_2(t) + \ldots \tag{2.85}$$

We substitute this into the equation of motion (2.83) and equate terms order by order in ϵ. The first first equation is simply $\ddot{x}_0 = -\omega^2 x_0$ and is solved by (2.84). To find the subsequent equations, it's best to think of $\lambda \sim \epsilon$ and let the dimensions come out in

the wash. We have

$$\begin{aligned} \mathcal{O}(\epsilon): \qquad & \ddot{x}_1 + \omega^2 x_1 = -\frac{\omega^2}{A^2} x_0^3 \\ \mathcal{O}(\epsilon^2): \qquad & \ddot{x}_2 + \omega^2 x_2 = -\frac{3\omega^2}{A^2} x_0^2 x_1 \end{aligned} \tag{2.86}$$

We see that we can now solve these equations successively. The solution x_0 acts as a source for x_1. Both x_0 and x_1 act as sources for x_2. And so on.

Let's start by solving the first equation (2.86). For simplicity, I'll set the integration constant $t_0 = 0$ in (2.84). (As we'll see, this will turn out be not as innocent as it looks!) Plugging in the solution (2.84), we have

$$\ddot{x}_1 + \omega^2 x_1 = -\omega^2 A \cos^3(\omega t) = -\frac{\omega^2 A}{4}(\cos(3\omega t) + 3\cos(\omega t)) \tag{2.87}$$

This is a linear equation in x_1, with two sources on the right-hand side: one proportional to $\cos(3\omega t)$ and one proportional to $\cos(\omega t)$. We can solve each of these in turn. First, we have

$$\ddot{x}_1 + \omega^2 x_1 = -\frac{\omega^2 A}{4}\cos(3\omega t) \quad \Rightarrow \quad x_1 = A_1 \cos(3\omega t)$$

where a quick calculation shows that the overall amplitude is $A_1 = A/32$. This is typical of the kind of behaviour that arises in perturbation theiry: we started with a system oscillating at one frequency ω, and the perturbation induces an oscillation at a different frequency, here 3ω.

Next we turn to the second source term in (2.87). And here we hit a snag. This is because the frequency of the source is ω, which coincides with the natural frequency of x_1. The solution isn't $x_1 \sim \cos(\omega t)$ because that's the homogeneous solution. Instead the solution is

$$\ddot{x}_1 + \omega^2 x_1 = -\frac{3\omega^2 A}{4}\cos(\omega t) \quad \Rightarrow \quad x_1 = A_1' \omega t \sin(\omega t) \tag{2.88}$$

where, again, a short calculation shows that $A_1' = -3A/8$. Rather worryingly, this solution isn't oscillatory, but grows with t. This means that beyond some time, $t \sim 1/\omega\epsilon$, the term $x_1(t)$ will not be a small correction to our original solution $x_0(t)$. Instead it will dominate. And that's bad because our whole strategy was based on finding increasingly small corrections. Term like (2.88) that grow with time are called *secular*.

What to do about this? The answer depends on the problem at hand. For some problems, the secular growth may well be telling us that the original solution was unstable. However, that can't be the case for the anharmonic oscillator as we've already seen, on energy grounds, that the solution can't run away. Instead, the culprit is our original expansion (2.85).

The problem is that, to leading order, the non-linear term has two effects: it induces a second frequency, 3ω, in the oscillation. And it also changes the original frequency ω. This means that a better perturbative expansion, at leading order, is to take

$$x(t) = A\cos((\omega + \epsilon\omega_1)t) + \epsilon x_1(t) + \mathcal{O}(\epsilon^2)$$

which allows for a change in the frequency of the original solution. If we expand out this first term to leading order in ϵ, we have

$$\cos((\omega + \epsilon\omega_1)t) \approx \cos(\omega t) - \epsilon\omega_1 t\sin(\omega t) + \ldots$$

There we see the same $\omega t\sin(\omega t)$ behaviour that appeared in (2.88). Except now we understand the meaning of this: it is just a small change to the original frequency. Comparing to our solution (2.88), this shift in the frequency must be

$$\Delta\omega = \epsilon\omega_1 = \frac{3}{8}\epsilon\omega \tag{2.89}$$

The frequency should be expected to pick up further shifts at $\mathcal{O}(\epsilon^2)$ and higher.

There is another way to think about this. Our original solution (2.84) included an integration constant t_0 that we set to zero in our subsequent analysis. The shift of frequency (2.89) can be viewed as endowing this integration constant with some time dependence, so that $t_0 \to -\frac{3}{8}\epsilon t$. This kind of behaviour is common in many problems. As a general rule of thumb, when secular terms arise one should see if they can be accounted for by endowing integration constants with some small time dependence.

Figure 17: Wolfgang Pauli and Niels Bohr stare in wonder at a spinning top.

3. The Motion of Rigid Bodies

In this section, we turn to a slightly different topic. We will look at the motion of rotating objects. This is a story that we started in the lectures on Dynamics and Relativity but, as we shall see, there is a lot more to say. One of the lessons is that things that spin can be somewhat counterintuitive.

For the most part, we will consider *rigid bodies*. These are extended objects that don't have any internal degrees of freedom. You could have in your mind a bunch of point masses, fixed together by rigid rods as shown on the right. Mathematically, the definition of a rigid body is a collection of N points constrained so that the distance between the points is fixed. i.e.

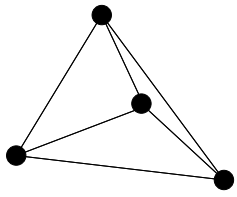


Figure 18:

$$|\mathbf{r}_i - \mathbf{r}_j| = \text{constant}$$

for all $i, j = 1, \ldots, N$. In many situations, we're interested in continuous solids, rather than discrete point masses. As we'll see, it's simple to generalise our ideas to this situation.

3.1 Kinematics

Our first tasks are to understand the ways in which a rigid body can move, and then to develop a mathematical description of this. We already made a start at this in our earlier lectures classical mechanics. For example, there we learned that a rigid body has six degrees of freedom: three translations, and three rotations. Now we will finish the job.

To kick things off, we will ignore translations. We achieve this by considering a rigid body that is fixed at some point, P. This means that the only allowed motion is rotation about P. We want to describe this rotation.

To do this, we introduce two different sets of axes. The first is a set of axes that is fixed in space. We denote these as $\{\tilde{\mathbf{e}}_a : a = 1, 2, 3\}$. The second is a set of axes that is embedding in the rigid body. We denote these as $\{\mathbf{e}_a : a = 1, 2, 3\}$. The idea is that the rotation of the rigid body can be described by saying how the body frame axes $\mathbf{e}_a(t)$ move with respect to the space frame axes $\tilde{\mathbf{e}}_a$.

We take both sets of axes to be orthonormal, meaning that

$$\tilde{\mathbf{e}}_a \cdot \tilde{\mathbf{e}}_b = \delta_{ab} \quad \text{and} \quad \mathbf{e}_a(t) \cdot \mathbf{e}_b(t) = \delta_{ab} \tag{3.1}$$

These two orthonormal bases must are related by time-dependent rotation matrix $R(t)$, so that

$$\mathbf{e}_a(t) = R_{ab}(t)\tilde{\mathbf{e}}_b \tag{3.2}$$

The fact that $R(t)$ is a rotation matrix follows from (3.1): substituting (3.2), we have $\mathbf{e}_a \cdot \mathbf{e}_b = (R_{ac}\tilde{\mathbf{e}}_a) \cdot (R_{bd}\tilde{\mathbf{e}}_d) = R_{ac}R_{bc} = \delta_{ab}$. Losing the indices, this tells us that R is an orthogonal matrix, satisfying $RR^T = 1$. Any such matrix describes a rotation, together with a reflection.

We will assume that the bases $\{\mathbf{e}_a\}$ and $\{\tilde{\mathbf{e}}_a\}$ share the same handedness. (For example, it may be useful to take $\mathbf{e}_a(t = 0) = \tilde{\mathbf{e}}_a$.) This means that $\det R = 1$ and is a rotation matrix. For concreteness, we take both bases to be right-handed.

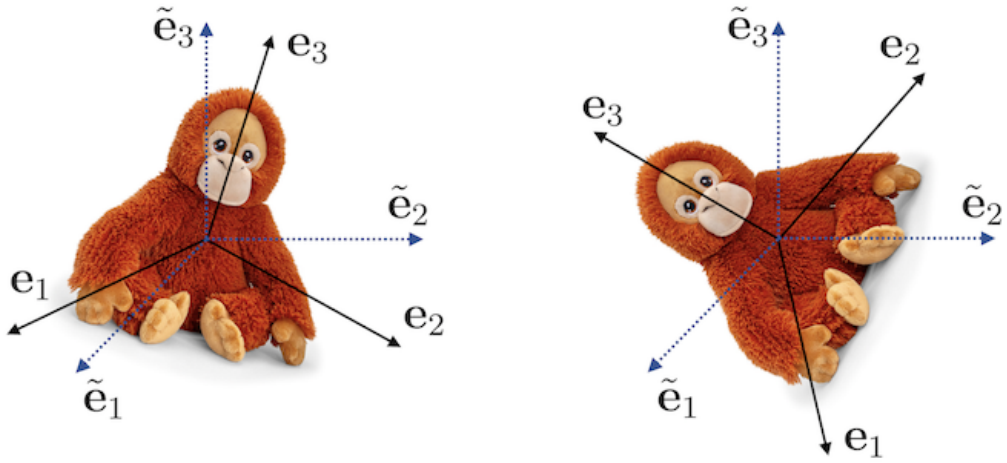


Figure 19: The space frame $\{\tilde{\mathbf{e}}_a\}$ stays fixed, while the body frame $\{\mathbf{e}_a\}$ changes with time, rotating with the rigid body.

As the rigid body rotates it is described by a time dependent matrix $R(t) \in SO(3)$. Conversely, every one-parameter family $R(t)$ describes a possible motion of the rigid body. Translating this into the language of classical dynamics, it means that the configuration space $\mathcal{C}$ of a rotating rigid body is identified with the space of all 3×3 orthogonal matrices with unit determinant. Or, more concisely,

$$\mathcal{C} = SO(3)$$

A 3×3 matrix has 9 components but the condition of orthogonality $R^T R = 1$ imposes 6 relations, so the configuration space $\mathcal{C}$ is 3 dimensional and we need 3 generalised coordinates to parameterise $\mathcal{C}$. We shall describe a useful choice of coordinates, known as Euler angles, in section 3.5.

3.1.1 Angular Velocity

Any point $\mathbf{r}$ in the body can be expanded in either the space frame or the body frame:

$$\begin{aligned} \mathbf{r}(t) &= \tilde{r}_a(t)\, \tilde{\mathbf{e}}_a \quad \text{in the space frame} \\ &= r_a\, \mathbf{e}_a(t) \quad \text{in the body frame} \end{aligned} \tag{3.3}$$

Note the different places where time dependence occurs. From the perspective of the space frame, any point in the body is moving: hence $\tilde{r}_a(t)$. In contrast, from the perspective of the body frame all the points are at some fixed distance r_a, independent of time. Instead, it is the axes $\mathbf{e}_a(t)$ that move.

Substituting the relation (3.2) into (3.3), we get the relationship between coordinates

$$\tilde{r}_b(t) = r_a R_{ab}(t) \tag{3.4}$$

Next, we can take time derivatives to see how things change with time. In fact, there are two different questions that we can ask: how does the body basis $\mathbf{e}_a(t)$ change with time, and how do the coordinates $\tilde{r}_a(t)$ change with time? We'll answer both here, although it turns out that the former is the more useful, so we start with that. From (3.2), we have

$$\frac{d\mathbf{e}_a}{dt} = \frac{dR_{ab}}{dt}\tilde{\mathbf{e}}_b = \frac{R_{ab}}{dt}R^{-1}_{bc}\mathbf{e}_c \tag{3.5}$$

The matrix $\dot{R}R^{-1}$ is important enough to deserve its own name. We call it

$$\Omega = \frac{dR}{dt}R^{-1} = \frac{dR}{dt}R^T \tag{3.6}$$

It's simple to show that Ω is an anti-symmetric matrix. This follows from the fact that $RR^T = 1$ which, upon differentiation, gives

$$\Omega = \frac{dR}{dt}R^T = -R\frac{dR^T}{dt} = -\Omega^T$$

We write the entries of this anti-symmetric matrix as

$$\Omega = \begin{pmatrix} 0 & \omega_3 & -\omega_2 \\ -\omega_3 & 0 & \omega_1 \\ \omega_2 & -\omega_1 & 0 \end{pmatrix} \quad \text{or} \quad \Omega_{ab} = \epsilon_{abc}\omega_c \quad \Rightarrow \quad \omega_a = \frac{1}{2}\epsilon_{abc}\Omega_{bc}$$

We think of the ω_a as the components of a vector *in the body frame*, and define

$$\boldsymbol{\omega} = \omega_a \mathbf{e}_a$$

The vector $\boldsymbol{\omega}$ is called the *instantaneous angular velocity*. As we see above, its components ω_a are measured with respect to the body frame. Our expression (3.5) for the change of the body basis then becomes

$$\frac{d\mathbf{e}_a}{dt} = \Omega_{ac}\mathbf{e}_c = \epsilon_{acb}\omega_b\mathbf{e}_c = \boldsymbol{\omega} \times \mathbf{e}_a \tag{3.7}$$

where, in the final step, we've used the fact that our body frame basis is right-handed and so obeys $\mathbf{e}_a \times \mathbf{e}_b = \epsilon_{abc}\mathbf{e}_c$. Equivalently, expanding $\mathbf{r} = r_a\mathbf{e}_a(t)$ in the body frame, we have

$$\frac{d\mathbf{r}}{dt} = r_a\frac{d\mathbf{e}_a}{dt} = r_a\boldsymbol{\omega} \times \mathbf{e}_a = \boldsymbol{\omega} \times \mathbf{r} \tag{3.8}$$

We've seen the final expression (3.7) in our previous lectures on classical mechanics. It is the usual description of a basis rotating with angular velocity $\boldsymbol{\omega}$. Here the derivation was a little more formal from what we saw previously but we can recapture our previous intuition for $\boldsymbol{\omega}$ by drawing a simple picture, as shown on the right. Consider a displacement of a given point $\mathbf{r}$ in the body by rotating an infinitesimal amount $d\phi$ about an axis $\hat{\mathbf{n}}$. From the figure, we see that $|d\mathbf{r}| = |\mathbf{r}|\, d\phi \sin\theta$. Moreover, this displacement is perpendicular to $\mathbf{r}$ since the distance to P is fixed by the definition of a rigid body. So we have

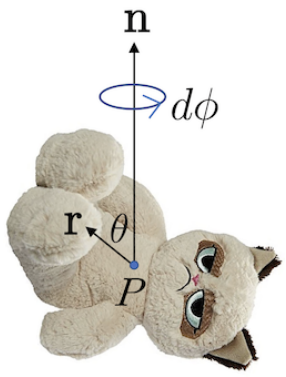


Figure 20:

$$d\mathbf{r} = d\boldsymbol{\phi} \times \mathbf{r} \tag{3.9}$$

with $d\boldsymbol{\phi} = \hat{\mathbf{n}} d\phi$. "Dividing" this equation by dt, we have the result

$$\dot{r} = \boldsymbol{\omega} \times \mathbf{r} \tag{3.10}$$

where $\boldsymbol{\omega} = d\boldsymbol{\phi}/dt$ is the instantaneous angular velocity. In general, both the axis of rotation $\hat{\mathbf{n}}$ and the rate of rotation $d\phi/dt$ will change over time. This, of course, is our result (3.8).

Finally, we mentioned above that, in addition to asking how $\mathbf{e}_a(t)$ changes with time, we could also ask how the space frame coordinates $\tilde{r}_a(t)$ change with time. Differentiating (3.4), we have

$$\frac{d\tilde{r}_a}{dt} = r_a \frac{dR_{ab}}{dt} = \tilde{r}_c R^{-1}_{ac} \frac{dR_{ab}}{dt}$$

Stare closely at the indices and you'll see that it's *not* the angular velocity matrix $\Omega = \dot{R}R^T$ that appears in this expression. Instead, we have $\dot{\tilde{r}}_a = \tilde{r}_c \tilde{\Omega}_{ca}$, with

$$\tilde{\Omega} = R^{-1} \frac{dR}{dt} \tag{3.11}$$

with R^{-1}m and $\dot{R}$ appearing in the opposite order in this new matrix $\tilde{\Omega}$ from the angular velocity matrix Ω. The matrix $\tilde{\Omega}$ is sometimes called the *convective angular velocity*. You can think of the two as like the difference between passive and active transformations. As these things are confusing, it's best to pick one and stick with it. So for the rest of these lectures we'll work only with our original $\Omega = \dot{R}R^T$ defined in (3.6).

3.1.2 Path Ordered Exponentials

Our goal throughout this chapter will be to figure out the angular velocity vector $\boldsymbol{\omega}(t)$ of various objects as they spin and turn. In general, once we've got an expression for $\boldsymbol{\omega}(t)$ we'll simply declare success and move on. However, as we've seen, the configuration space $\mathcal{C}$ is actually parameterised by the $SO(3)$ matrix $R(t)$. So one might wonder how, given $\boldsymbol{\omega}(t)$, we can reconstruct the evolution of $R(t)$. The purpose of this short section is to briefly explain how to do this, at least in principle.

As we've seen above, the angular momentum vector $\boldsymbol{\omega}$ is equivalent to the anti-symmetric matrix $\Omega_{ab} = \epsilon_{abc}\omega_c$ and we can determine the rotation matrix $R(t)$ by solving the differential equation

$$\Omega = \frac{dR}{dt}\, R^{-1} \tag{3.12}$$

So what do solutions to this equation look like? If Ω and R were scalar functions of time, then life would be straightforward: we could simply integrate this equation to get the solution

$$R(t) = \exp\left(\int_0^t \Omega(t')\, dt'\right) \tag{3.13}$$

which satisfies the initial condition $R(0) = 1$. But things are slightly more complicated because both Ω and R are matrices. Let's first describe how we take the exponential of a matrix. And, as we now show, the would-be solution (3.13) no longer works for matrices.

To see what goes wrong with (3.13), first recall that the exponential of any matrix M is defined by the Taylor expansion,

$$\exp(M) \equiv 1 + M + \tfrac{1}{2}M^2 + \ldots \tag{3.14}$$

This means that we should think of the would-be solution (3.13) as a Taylor expansion in $M = \int_0^t \Omega(t')dt'$. The trouble comes when we get to the quadratic term in the expansion. If we differentiate this with respect to time, we have

$$\frac{1}{2}\frac{d}{dt}\left(\int_0^t \Omega(t')\, dt'\right)^2 = \frac{1}{2}\,\Omega(t)\left(\int_0^t \Omega(t')\, dt'\right) + \frac{1}{2}\left(\int_0^t \Omega(t')\, dt'\right)\Omega(t) \tag{3.15}$$

The ordering here is important. In particular, we can't commute the $\Omega(t)$ past the terms in the integral with $\Omega(t')$ with $t' \neq t$ because there is no reason to think that the angular velocity matrices $\Omega(t)$ commute at different times. This is the reason that

(3.13) does *not* solve our original equation (3.12), which can be written as $\dot{R} = \Omega R$. The first term on the right-hand side of (3.15) looks good, since this also appears in the Taylor expansion of ΩR. But the second term isn't right just because the $\Omega(t)$ sits in the wrong place. The upshot is that equation (3.13) is *not* the solution to (3.12) when Ω and R are matrices.

However, seeing how (3.13) fails to solve (3.12) does give us an idea of how to proceed. But it does give us a hint about how we should proceed. Since the problem is in the ordering of the matrices, the correct solution to (3.12) takes a similar form as (3.13), but with a different ordering. This is known as the *time ordered exponential*, and denoted as

$$R(t) = T\,\exp\left(\int_0^t \Omega(t')\,dt'\right) \tag{3.16}$$

All the subtleties are hiding in that capital T that sits in front. This means that when we Taylor expand the exponential, all matrices are ordered so that later times appear on the left. In other words

$$R(t) = 1 + \int_0^t \Omega(t')\,dt' + \int_0^{t''}\int_{t'}^t \Omega(t'')\,\Omega(t')\;dt'dt'' + \ldots \tag{3.17}$$

The double integral is taken over the range $0 < t' < t'' < t$. If we now differentiate this double integral with respect to t, we get just the one term,

$$\frac{d}{dt}\int_0^{t''}\int_{t'}^t \Omega(t'')\,\Omega(t')\;dt'dt'' = \Omega(t)\,\left(\int_0^t \Omega(t')\,dt'\right)$$

instead of the two that appear in (3.15). It can be checked that the higher terms in the Taylor expansion also have the correct property if they are ordered so that matrices evaluated at later times appear to the left in the integrals.

In some ways, the solution (3.17) feels a bit like cheating. To some extent, we solved the problem by introducing a new notation - the time ordering symbol T – to do the job for us. And it's certainly true that for a given $\Omega(t)$ it can be difficult to explicitly construct the corresponding $R(t)$. Nonetheless, this kind of ordered integral shows up frequently in theories involving non-commuting matrices. Examples include Dyson's formula in the lectures on Quantum Field Theory and the Wilson line in the lectures on Gauge Theories.

As an aside, there's some interesting group theory lurking in (3.17). We've already seen that the rotation matrix $R(t)$ lies in the group $SO(3)$. Continuous groups of this kind are known as *Lie groups*. It's a fact that any element of a Lie group can be written as the exponential of a different kind of matrix which, in this case, is the anti-symmetric angular velocity matrix Ω. The matrices that sit in the exponents are said to belong to the *Lie algebra*, usually denoted in lower case as $so(3)$. And the Lie algebra $so(3)$ is defined to be the space of 3×3 anti-symmetric matrices.

3.2 The Inertia Tensor

Having discussed the way to describe a rotating body, we now turn to the dynamics. Our route to the dynamics is through the kinetic energy which, for our pinned, rotating rigid body, is given by

$$\begin{aligned} T &= \tfrac{1}{2}\sum_i m_i \dot{\mathbf{r}}_i^2 \\ &= \tfrac{1}{2}\sum_i m_i \left(\boldsymbol{\omega}\times\mathbf{r}_i\right)\cdot\left(\boldsymbol{\omega}\times\mathbf{r}_i\right) \\ &= \tfrac{1}{2}\sum_i m_i \Big((\boldsymbol{\omega}\cdot\boldsymbol{\omega})(\mathbf{r}_i\cdot\mathbf{r}_i) - (r_i\cdot\boldsymbol{\omega})^2\Big) \end{aligned}$$

We can succinctly write this as

$$T = \frac{1}{2}\omega_a \mathcal{I}_{ab}\omega_b \tag{3.18}$$

where $\mathcal{I}_{ab}$, $a, b = 1, 2, 3$ are the components of the *inertia tensor* measured in the body frame

$$\mathcal{I}_{ab} = \sum_i m_i \left((\mathbf{r}_i\cdot\mathbf{r}_i)\delta_{ab} - (\mathbf{r}_i)_a(\mathbf{r}_i)_b\right) \tag{3.19}$$

Note that the inertia tensor is symmetric: $\mathcal{I}_{ab} = \mathcal{I}_{ba}$. Furthermore, the components are independent of time since they are measured with respect to the body frame. As we mentioned in the introduction to this section, we will often be interested in continuous bodies rather than discrete point masses. In this case, there's an obvious, analogous expression for the inertia tensor

$$\mathcal{I} = \int d^3\mathbf{r}\,\rho(\mathbf{r}) \begin{pmatrix} y^2+z^2 & -xy & -xz \\ -xy & x^2+z^2 & -yz \\ -xz & -yz & x^2+y^2 \end{pmatrix}$$

Since $\mathcal{I}_{ab}$ is a symmetric real matrix, we can diagonalise it. Moreover, it can be diagonalised by an *orthogonal* matrix O, such that

$$O\mathcal{I}O^T = \mathcal{I}' = \text{diag}(I_1, I_2, I_3)$$

Because O is orthogonal, it can be viewed as a rotation of the original body frame basis $\{\mathbf{e}_a\}$ to a new basis $\{O\mathbf{e}_a\}$. This new basis is precisely the eigenvectors of the inertia tensor $\mathcal{I}$.

Physically, there's something rather surprising about this result. It means that every object, no matter how complicated, has a preferred set of orthogonal axes sitting within it. For an object that has a lot of symmetry, this is no surprise as the axes simply align with the symmetry. For example, for a cube rotating around its middle point, the axes will be perpendicular to the faces of the cube. But complicated objects that have no symmetry at all also have a preferred set of axes. These are the eigenvectors of the inertia tensor.

The preferred body frame axes are known as *principal axes*. As we have seen, in this frame the inertia tensor takes the diagonal form

$$\mathcal{I} = \text{diag}(I_1, I_2, I_3)$$

The eigenvalues I_a are called the *principal moments of inertia*. The kinematical properties of a rigid body are fully determined by its mass, principal axes, and moments of inertia.

These principal moments of inertia are real and positive. To see this, consider some arbitrary vector $\mathbf{n}$ and look at

$$\mathcal{I}_{bc}n^b n^c = \sum_i m_i(\mathbf{r}_i^2\mathbf{n}^2 - (\mathbf{r}_i \cdot \mathbf{n})^2) \geq 0$$

The inequality is a strict equality only if all the $\mathbf{r}_i$ lie on a line. If $\mathbf{n}$ is the a^{th} eigenvector of I then this result becomes $\mathcal{I}_{bc}n^b n^c = I_a|\mathbf{n}|^2$ which tells us $I_a \geq 0$.

In our previous lectures on classical mechanics, we introduced the idea of a moment of inertia about a given axis $\mathbf{n}$. The inertia tensor is the grown-up version of this. We can use it to compute the moment of inertia around any axis $\mathbf{n}$ by $I = \hat{\mathbf{n}} \cdot \mathcal{I}\hat{\mathbf{n}}$. We'll now illustrate these ideas with some simple examples, all of which we also covered in our previous course.

Example: The Rod

As a first example, consider a uniform rod of length l and mass M. We will compute the inertia tensor about the centre of the rod. By symmetry, the tensor takes the form $\mathcal{I} = \text{diag}(I_1, I_1, 0)$ where a quick calculation is needed to figure out the moment of inertia I_1,

$$I_1 = \int_{-l/2}^{l/2} \rho\, x^2\, dx = \frac{1}{12} M l^2$$

This is a result that we saw in our previous course.

Another Example: The Disc

Now consider a uniform disc of radius R and mass M, lying in the (x, y)-plane. The density of the disc is $\rho = M/\pi r^2$ We'll again compute $\mathcal{I}$ about the centre of the disc. This time we have $\mathcal{I} = \text{diag}(I_1, I_2, I_3)$, with

$$I_1 = \int \rho y^2 d^2x \quad , \quad I_2 = \int \rho x^2 d^2x$$

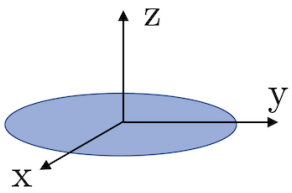


Figure 21:

so $I_1 = I_2$ by symmetry, while

$$I_3 = \int \rho(x^2 + y^2) d^2x$$

Therefore

$$I_3 = I_1 + I_2 = 2\pi\rho \int_0^r r'^3\, dr' = \tfrac{1}{2} M r^2 \tag{3.20}$$

and the moments of inertia are $I_1 = I_2 = \frac{1}{4} M r^2$ and $I_3 = \frac{1}{2} M r^2$.

3.2.1 Parallel Axis Theorem

The inertia tensor depends on what point P in the body is held fixed. In general, if we know $\mathcal{I}$ about a point P then it is messy to compute $\mathcal{I}'$ about some other point P'. But this computation becomes very simple if P happens to coincide with the centre of mass of the object. This follows from the *parallel axis theorem*:

Claim: If $\mathcal{I}$ is the inertia tensor about the centre of mass and P' is displaced by a vector $\mathbf{d}$ from the centre of mass, then the inertia tensor about P' is

$$\mathcal{I}'_{ab} = \mathcal{I}_{ab} + M(\mathbf{d}^2 \delta_{ab} - \mathbf{d}_a \mathbf{d}_b) \tag{3.21}$$

Note that the additional term $M(\mathbf{d}^2\delta_{ab} - \mathbf{d}_a\mathbf{d}_b)$ is the inertia tensor we would find for a point mass M sitting at point $\mathbf{d}$.

Proof: The proof is of the plug-it-in-and-check variety

$$\begin{aligned}\mathcal{I}'_{ab} &= \sum_i m_i \left\{(\mathbf{r}_i - \mathbf{d})^2\delta_{ab} - (\mathbf{r}_i - \mathbf{d})_a(\mathbf{r}_i - \mathbf{d})_b\right\} \\ &= \sum_i m_i\Big(r_i^2\delta_{ab} - (\mathbf{r}_i)_a(\mathbf{r}_i)_b + \mathbf{d}^2\delta_{ab} - \mathbf{d}_a\mathbf{d}_b - [2\mathbf{r}_i\cdot\mathbf{d}\delta_{ab} - (\mathbf{r}_i)_a\mathbf{d}_b - (\mathbf{r}_i)_b\mathbf{d}_a]\Big)\end{aligned}$$

But the term in square brackets is linear in $\mathbf{r}$, and so proportional to $\sum m_i\mathbf{r}_i$ which vanishes if $\mathbf{r}_i$ is measured from the centre of mass. The remaining terms give the promised result (3.21). □

Example: The Rod Again

The inertia tensor of the rod about one of its ends is $I_1 = \frac{1}{12}Ml^2 + M(l/2)^2 = \frac{1}{3}Ml^2$.

Another Example: The Disc Again

We can consider our previous example of the disc, now displaced by $\mathbf{d} = (d, 0, 0)$ from the centre. We have

$$\mathcal{I}'_{\mathbf{c}} = \frac{M}{4}\begin{pmatrix} r^2 & & \\ & r^2 & \\ & & 2r^2 \end{pmatrix} + M\begin{pmatrix} 0 & & \\ & d^2 & \\ & & d^2 \end{pmatrix} = \frac{M}{4}\begin{pmatrix} r^2 & & \\ & r^2+4d^2 & \\ & & 2r^2+4d^2 \end{pmatrix}$$

Note that $\mathcal{I}'$ remains diagonal because we have shifted the point P along a principal axis.

3.2.2 Angular Momentum

Next we turn to the angular momentum of a rigid body. Again, we consider the case where the body is pinned at some point P, about which it rotates. We have

$$\mathbf{L} = \sum_i m_i\mathbf{r}_i \times \dot{\mathbf{r}}_i = \sum_i m_i\mathbf{r}_i \times (\boldsymbol{\omega}\times\mathbf{r}_i) = \sum_i m_i(r_i^2\boldsymbol{\omega} - (\boldsymbol{\omega}\cdot\mathbf{r}_i)\mathbf{r}_i)$$

We recognise the two terms in the sum as the inertia tensor (3.19), now multiplying the angular velocity $\boldsymbol{\omega}$,

$$\mathbf{L} = \mathcal{I}\boldsymbol{\omega}$$

Note that, perhaps counterintuitively, the angular momentum $\mathbf{L}$ is not necessarily aligned with the spin $\boldsymbol{\omega}$. If the spin happens to point along a principal axis of the rigid body then $\boldsymbol{\omega}$ and $\mathbf{L}$ do align. Otherwise $\boldsymbol{\omega}$ and $\mathbf{L}$ point in different directions. Many of the more peculiar properties of spinning objects follow from this simple fact.

The components of the inertia tensor $\mathcal{I}_{ab}$ are found in the body frame. Here we can write

$$L_a = I_{ab}\omega_b \tag{3.22}$$

where $\mathbf{L} = L_a\mathbf{e}_a$ and $\boldsymbol{\omega} = \omega_a\mathbf{e}_a$.

3.3 Free Tops

> "To those who study the progress of exact science, the common spinning-top is a symbol of the labours and the perplexities of men."
>
> *James Clerk Maxwell, no less*

We now have the tools at our disposal to understand the the motion of objects as they twist and turn.

Until now, we've only considered rigid bodies pinned at some point P. Now we free things up. The body will be suspended in space, free to move or rotate in any direction. We won't, for now, consider the effects of gravity, an omission that we'll remedy in Section 3.5. That means that the analysis of this section will describe planets or satellites, or simply objects in free fall where we ignore the fact that it's accelerating downwards.

The most general motion of a rigid body is an overall translation, superposed with a rotation. In principle, we could think about this rotation as being about any point inside or outside the body. But, as we now show, nice things happen if we consider the point of rotation to be about the centre of mass. We will write the position of a general point $\mathbf{r}_i$ in the rigid body as

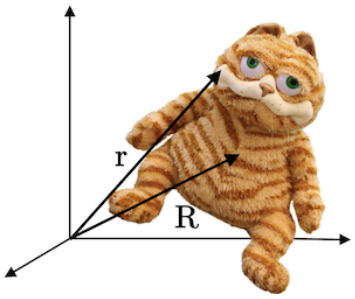


Figure 22:

$$\mathbf{r}_i(t) = \mathbf{R}(t) + \mathbf{y}_i(t)$$

where $\mathbf{R}$ is the centre of mass and, by definition, $\mathbf{y}_i$ is the position measured from the centre of mass. Considering the body to be rotating about the centre of mass means that

$$\dot{\mathbf{y}}_i = \boldsymbol{\omega} \times \mathbf{y}_i$$

The kinetic energy for the rigid body is then

$$\begin{aligned} T &= \tfrac{1}{2}\sum_i m_i \dot{\mathbf{r}}_i^2 \\ &= \sum_i m_i \left[\tfrac{1}{2}\dot{\mathbf{R}}^2 + \dot{\mathbf{R}}\cdot(\boldsymbol{\omega}\times\mathbf{y}_i) + \tfrac{1}{2}(\boldsymbol{\omega}\times\mathbf{y}_i)^2\right] \\ &= \tfrac{1}{2}M\dot{\mathbf{R}}^2 + \tfrac{1}{2}\boldsymbol{\omega}_a\mathcal{I}_{ab}\boldsymbol{\omega}_b \end{aligned}$$

where, to go to the final line, we've used the fact that $\sum_i m_i\mathbf{y}_i = 0$, which ensures that the term linear in $\boldsymbol{\omega}$ vanishes. We've also used our original definition of the inertia tensor (3.18) as it appears in the kinetic energy of a rotating body. This is a rather nice result: it tells us that the dynamics separates into the translational motion of the centre of mass $\mathbf{R}$, together with rotation about the centre of mass. Happily, this means that the work we did in the previous sections describing an object that rotates about a fixed point is also valid for a free object.

3.3.1 Euler's Equations

We know how the centre of mass of an object moves. For this reason, we will focus only on the rotational motion of the rigid body.

The basic idea underlying the rotation of a free, rigid body is very simple: angular momentum is conserved. This gives us the vector equation

$$\frac{d\mathbf{L}}{dt} = 0 \tag{3.23}$$

We expand this in the body frame, with $\mathbf{L} = L_a\mathbf{e}_a$. Both the components L_a and the basis vectors $\mathbf{e}_a$ can change with time, so we have

$$0 = \frac{d\mathbf{L}}{dt} = \frac{dL_a}{dt}\,\mathbf{e}_a + L_a\,\frac{d\mathbf{e}_a}{dt} = \frac{dL_a}{dt}\,\mathbf{e}_a + L_a\,\boldsymbol{\omega}\times\mathbf{e}_a \tag{3.24}$$

where we've used the expression $\dot{\mathbf{e}}_a = \boldsymbol{\omega}\times\mathbf{e}_a$, derived in (3.7), to express the change of the body frame basis in terms of the angular velocity.

The next step is to combine (3.24) with the relation $\mathbf{L} = \mathcal{I}\boldsymbol{\omega}$ from (3.22). Things are significantly simpler if we pick the body frame axes $\{\mathbf{e}_a\}$ to coincide with the principal axes. Then the inertia tensor is diagonal and the expression $L_a = \mathcal{I}_{ab}\omega_b$ becomes $L_1 = I_1\omega_1$ and so on. The conservation of angular momentum (3.23) now becomes three non-linear coupled first order differential equations,

$$\begin{aligned} I_1\dot{\omega}_1 + \omega_2\omega_3(I_3 - I_2) &= 0 \\ I_2\dot{\omega}_2 + \omega_3\omega_1(I_1 - I_3) &= 0 \\ I_3\dot{\omega}_3 + \omega_1\omega_2(I_2 - I_1) &= 0 \end{aligned} \tag{3.25}$$

These are *Euler's Equations*. (As an aside, Euler has way too many things named after him! Euler's equations are not to be confused with the more general Euler-Lagrange equations, nor with the equation for a fluid which is usually called *the* Euler equation.)

It is straightforward to extend the analysis above to include a torque $\boldsymbol{\tau}$ acting on the body. The equation of motion becomes $\dot{\mathbf{L}} = \boldsymbol{\tau}$ and repeating the steps above results in Euler's equations (3.25), now with the components of the torque $\boldsymbol{\tau} = \tau_a \mathbf{e}_a$, expanded in the body frame, arising on the right-hand side.

3.3.2 The Symmetric Top

The simplest example of a rigid body is the sphere. It is, sadly, too simple. The moments of inertia all coincide, $I_1 = I_2 = I_3$, which means that the inertia tensor $\mathcal{I}$ is diagonal and the angular momentum $\boldsymbol{\omega}$ is parallel to the angular momentum $\mathbf{L}$. Euler's equations (3.25) then tell us that $\boldsymbol{\omega}$ is constant. If you spin a sphere about some axis, then it continues to spin about the same axis.

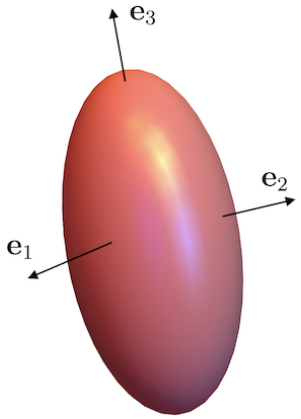


Figure 23:

To find something more interesting, we need to go to the next simplest case. This an object with $I_1 = I_2 \neq I_3$, known as the *symmetric top*. An example is shown in the figure to the right. Euler's equations (3.25) for the symmetric top are,

$$\begin{aligned} I_1\dot{\omega}_1 &= \omega_2\omega_3(I_1 - I_3) \\ I_2\dot{\omega}_2 &= -\omega_1\omega_3(I_1 - I_3) \\ I_3\dot{\omega}_3 &= 0 \end{aligned} \tag{3.26}$$

We see that ω_3, which is the spin about the axis of symmetry, is a constant of motion. In contrast, the spins about the other two axes vary in time, and obey the equations

$$\dot{\omega}_1 = \Omega\omega_2 \quad \text{and} \quad \dot{\omega}_2 = -\Omega\omega_1 \quad \text{with} \quad \Omega = \frac{I_1 - I_3}{I_1}\omega_3 \tag{3.27}$$

These are familiar equations, with (ω_1, ω_2) tracing out a circle

$$\omega_1 = \omega_0 \sin\Omega t \quad \text{and} \quad \omega_2 = \omega_0 \cos\Omega t \tag{3.28}$$

for some integration constant ω_0. Recall that the components ω_a in Euler's equations describe the angular velocity in the body frame. The solution (3.28) is then telling us that, in the body frame, the direction of the spin precesses about the axis of symmetry $\mathbf{e}_3$, with frequency Ω. Note from (3.27) that the sign of Ω depends on whether $I_1 > I_3$

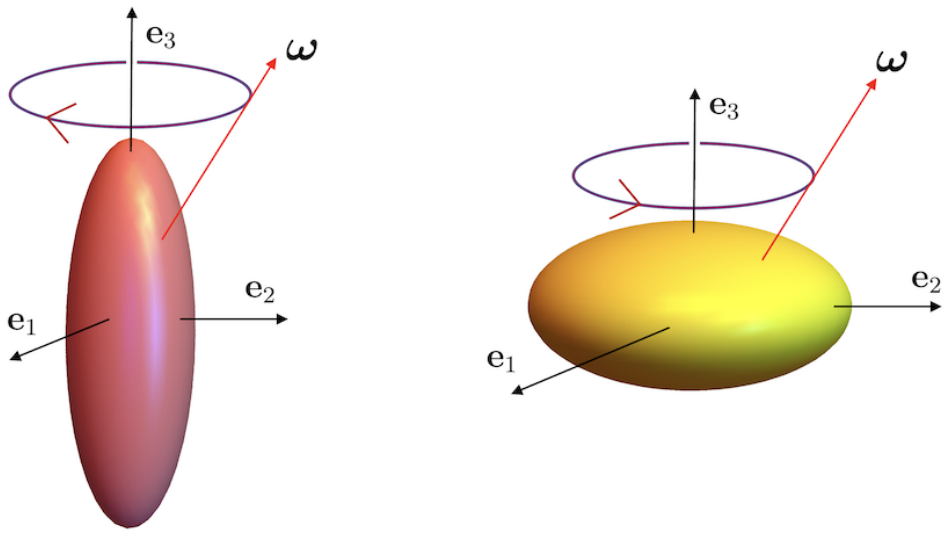


Figure 24: The precession of the spin: the direction of precession depends on whether the object is tall and skinny with $I_3 < I_1$ (known as a *prolate spheroid* or is short and fat with $I_3 > I_1$ (known as an *oblate spheroid*.

or $I_1 < I_3$. For a tall, skinny object, like the one shown in the figure above, $I_3 < I_1$ and $\Omega > 0$. Meanwhile, a short, fat object has $I_1 < I_3$ and $\Omega < 0$. Examples of the spin precession are shown in Figure 24.

The story above holds in the body frame, meaning it's relevant for someone unlucky enough to be clinging to the spinning object (presumably trying to keep their food down). But what does the situation look like for someone in the inertial, space frame? The angular momentum $\mathbf{L}$ is a fixed vector. But ω_3, and hence L_3, are also fixed which ensures that the angle between $\mathbf{e}_3$ and $\mathbf{L}$ doesn't change in time. Instead, $\mathbf{e}_3$ precesses around $\mathbf{L}$, while the body simultaneously spins such that $\boldsymbol{\omega}$ remains between $\mathbf{e}_3$ and $\mathbf{L}$. A picture of this motion, albeit one that is necessarily rather static, is shown to the right[1]. The fact that $\boldsymbol{\omega}$ precesses around the body frame axis $\mathbf{e}_3$ is sometimes referred to, reasonably, as a *wobble*. We'll return to the motion of the symmetric top in the space frame in Section 3.4.3 where we'll parameterise the motion in terms of Euler angles.

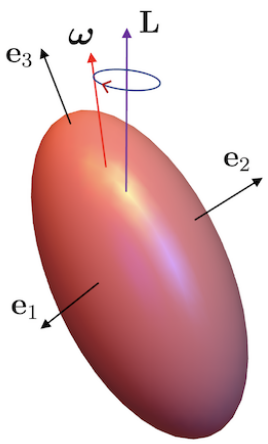


Figure 25:

[1]You can find more enlightening videos of this motion on YouTube.

An Example: The Earth's Wobble

When we were children, our teachers told us that the Earth is round. They presumably meant this as a topological statement rather than a statement about the metric. The spin of the Earth causes it to bulge at the equator so it is no longer a sphere but is closer to an oblate ellipsoid. In other words, the Earth is a symmetric top. The moments of inertia turn out to be

$$\frac{I_1 - I_3}{I_1} \approx -\frac{1}{300}$$

Our childhood teachers also told us that $\omega_3 = (1 \text{ day})^{-1}$. This information is enough to calculate the frequency of the earth's wobble; from (3.27), it should be

$$\Omega_{earth} = \frac{1}{300} \text{ day}^{-1}$$

This calculation was first performed by Euler in 1749 who predicted that the Earth completes a wobble every 300 days. Note, however, that the size of the wobble isn't predicted by the analysis above: it is given by the integration constant ω_0.

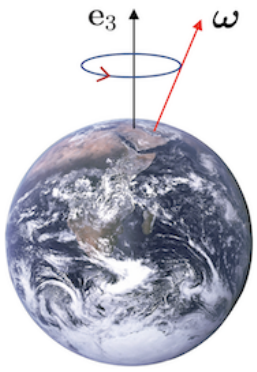


Figure 26:

Despite many searches, this effect wasn't detected until 1891 when Chandler re-analysed the data and saw a wobble with a period of 427 days. It is now known as the *Chandler wobble*. It is very small! The angular velocity $\boldsymbol{\omega}$ intercepts the surface of the earth approximately 10 metres from the North pole and precesses around it. A greatly exaggerated picture of the wobble is shown in the figure.

More recent measurements place the frequency at 435 days. The discrepancy with the predicted 300 days can be traced to the fact that the Earth is not, in fact, a rigid body. Instead it is flexible because of tidal effects. Less well understood is why these same tidal effects haven't caused the wobble to dampen and disappear completely. There are various theories about what keeps the wobble alive, from earthquakes to fluctuating pressure at the bottom of the ocean.

3.3.3 The Asymmetric Top: Stability

The most general rigid body has no symmetries and distinct moments of inertia, $I_a \neq I_b$ for $a \neq b$. Euler's equations are now non-linear differential equations and the general solution is complicated; we will give some insight into it in Section 3.3.4. Here, we look at a special case of the motion and present a simple but striking result.

The special situation that we're interested in is when the spin is aligned with one of the principal axes, say $\mathbf{e}_1$, so that

$$\omega_1 = \Omega \quad \text{and} \quad \omega_2 = \omega_3 = 0$$

This is a steady state solution to Euler's equations (3.25), with $\dot{\omega}_a = 0$. The question we want to ask is: what happens if we perturb the system, so that the spin lies close to the $\mathbf{e}_1$ axis, but isn't quite aligned. To answer this, we write the spin as

$$\omega_1 = \Omega + \eta_1 \quad , \quad \omega_2 = \eta_2 \quad , \quad \omega_3 = \eta_3 \tag{3.29}$$

where η_a, $a = 1, 2, 3$ are all taken to be small which, in this case, means $\eta_a \ll \Omega$. Substituting this into Euler's equations and ignoring terms of order η^2 and higher, we get the system of linear differential equations

$$\begin{aligned} I_1\dot{\eta}_1 &= 0 \\ I_2\dot{\eta}_2 &= \Omega\eta_3(I_3 - I_1) \\ I_3\dot{\eta}_3 &= \Omega\eta_2(I_1 - I_2) \end{aligned}$$

We substitute the third equation into the second to find an equation for just one of the perturbations, say η_2,

$$I_2\ddot{\eta}_2 = \Omega^2\frac{(I_3 - I_1)(I_1 - I_2)}{I_3}\eta_2 \equiv A\eta_2$$

The fate of the small perturbation depends on the sign of the quantity A. We have two possibilities

- $A < 0$: In this case, the disturbance will oscillate around the original solution.
- $A > 0$: In this case, the disturbance will grow exponentially.

Examining the definition of A, we find that if the body was originally spinning around the axis with the middle moment of inertia, meaning

$$I_2 < I_1 < I_3 \quad \text{or} \quad I_3 < I_1 < I_2,$$

then the motion is unstable. In contrast, if the body was originally spinning about an axis with either the smallest moment of inertia $I_1 < I_2, I_3$ or the largest moment of inertia $I_1 > I_2, I_3$, then the motion will be stable.

You can get a vivid (and, if you're not careful, painful) demonstration of this result by picking up some suitable object – say a book or, if you're brave, a tennis racket – and trying it for yourself. Throw the object upwards, spinning around either the axis with the largest or smallest moment of inertia, and it will gracefully return to you. Throw it spinning around the intermediate axis then good luck catching it as it tumbles and turns in unexpected directions due to the instability.

3.3.4 The Poinsot Construction

As we mentioned above, for unequal moments of inertia, Euler's equations are non-linear. But they're not terribly non-linear and there's no chaotic motion. Indeed, it turns out that there is an analytic solution for the general solution of an asymmetric top, given in terms of elliptic functions. We won't describe this here, but instead present a nice, geometrical perspective on the motion, due to Poinsot.

We start by working in the body frame. As always, we make progress by focussing on conserved quantities. There are two: the kinetic energy T and the magnitude of the angular momentum $\mathbf{L}^2$. In terms of the angular velocity, they are

$$2T = I_1\omega_1^2 + I_2\omega_2^2 + I_3\omega_3^2 \tag{3.30}$$

$$\mathbf{L}^2 = I_1^2\omega_1^2 + I_2^2\omega_2^2 + I_3^2\omega_3^2 \tag{3.31}$$

Each of these equations defines an ellipsoid in $\boldsymbol{\omega}$ space. The motion of the vector $\boldsymbol{\omega}$ is constrained to lie on the intersection of these two ellipsoids. The first of these ellipsoids, defined by

$$\frac{I_1}{2T}\,\omega_1^2 + \frac{I_2}{2T}\,\omega_2^2 + \frac{I_3}{2T}\,\omega_3^2 = 1 \tag{3.32}$$

is known as the *inertia ellipsoid* (or, sometimes, the inertia quadric). If we fix the kinetic energy, we can think of this abstract ellipsoid as embedded within the object, rotating with it.

As in the previous section, we'll take $I_1 > I_2 > I_3$. This means that ω_3 is the major axis of the inertial ellipsoid, while ω_1 is the minor axis. The second ellipsoid, defined by (3.31), has the same major and minor axes because $I_1^2 > I_2^2 > I_3^2$. To understand the motion in the body frame, we can consider the intersection of the two ellipsoids for different values of $\mathbf{L}^2$. These lines of intersection are drawn on the inertia ellipsoid in the figure.

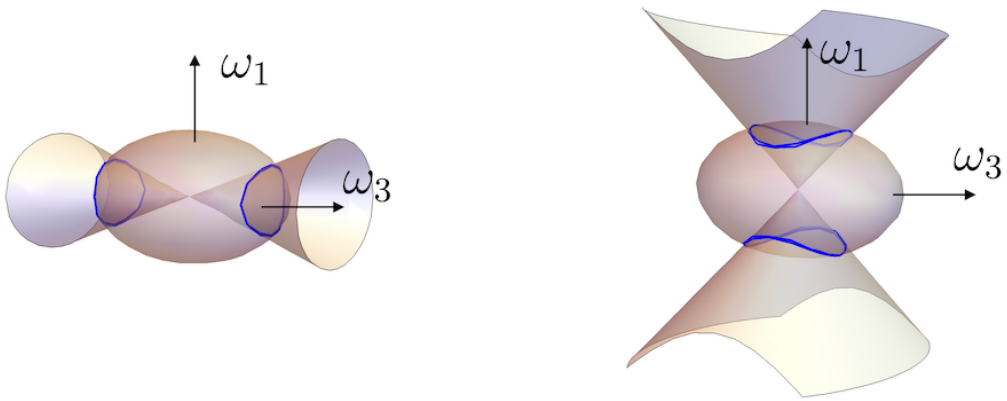


Figure 27: The intersection of the inertia ellipsoid with the polhode cone. On the left, $\mathbf{L}^2 > 2I_2T$ and the cone is oriented along the ω_3 axis. On the right, $2I_2T < \mathbf{L}^2 < 2I_3T$, and the cone is oriented along the ω_1 axis.

On the major and minor axes, the two ellipsoids intersect in small closed orbits, as shown. Meanwhile, the behaviour is different along the intermediate axis ω_2, where the intersection appears as two crossed lines. This gives a graphical rendering of the result that saw in the previous section. If the body is spinning along either the ω_1 or ω_3 axis, then a perturbation is stable, and the spin will precess around the original axis. Meanwhile, if it was originally spinning around ω_2, then a general perturbation will be unstable and the orientation of the spin will rapidly change when perturbed.

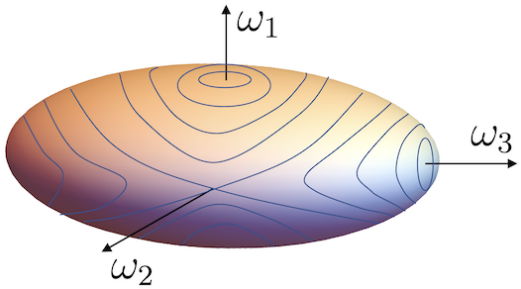


Figure 28:

The path that $\boldsymbol{\omega}$ traces on the inertia ellipsoid is known as the *polhode* curve. One of the things we learn from this pictorial representation is that the polhode curves are always closed and therefore motion in the body frame is periodic.

There is a slightly different way to view this. Equations (3.30) and (3.31) can be rearranged to give the condition

$$(2I_1T - \mathbf{L}^2)\, I_1\omega_1^2 + (2I_2T - \mathbf{L}^2)\, I_2\omega_2^2 + (2I_3T - \mathbf{L}^2)\, I_3\omega_3^2 = 0 \tag{3.33}$$

This is the equation for a cone. To see this, note that if $\boldsymbol{\omega}$ satisfies (3.33), then so does any rescaled vector $\lambda\omega$ for all real numbers λ. It is known as the *polhode cone*.

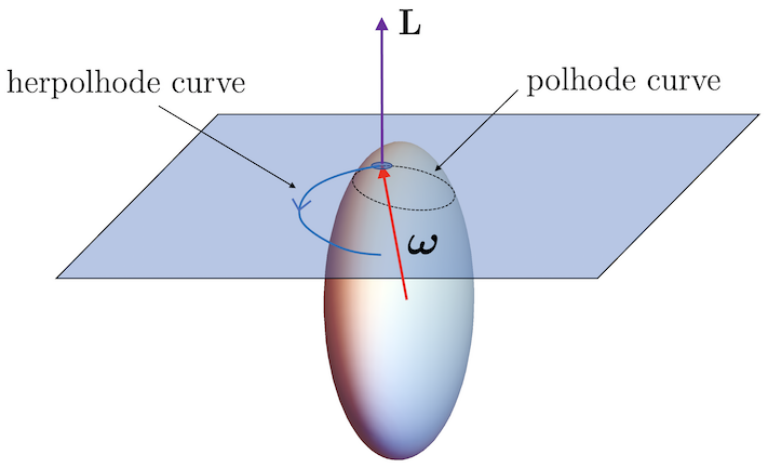


Figure 29: The inertia ellipsoid rolling around on the invariable plane, with the polhode and herpolhode curves drawn for a fixed time period.

For fixed T and $\mathbf{L}^2$, the motion of $\boldsymbol{\omega}$ in the body frame runs along the intersection of the inertia ellipsoid and the polhode cone. Two examples are shown in Figure 27, the first with $\mathbf{L}^2 < 2I_2T$, the second with $2I_2T < \mathbf{L}^2 < 2I_3T$, so that the orientation of the cone is different in the two cases.

The discussion above holds in the body frame. What does all this look like in the space frame? The vector $\mathbf{L}$ is a constant of motion. Since the kinetic energy $2T = \mathbf{L}\cdot\boldsymbol{\omega}$ is also constant, we learn that $\boldsymbol{\omega}$ must lie in a fixed plane perpendicular to $\mathbf{L}$. This is known as the *invariable plane*. The inertia ellipsoid touches the invariable plane at the point defined by the angular velocity vector $\boldsymbol{\omega}$. Moreover, the invariable plane is always tangent to the inertial ellipsoid at the point $\boldsymbol{\omega}$. To see this, note that the angular momentum can be written as

$$\mathbf{L} = \nabla_\omega T \tag{3.34}$$

where the gradient operator is in $\boldsymbol{\omega}$ space, i.e. $\nabla_\omega = (\partial/\partial\omega_1, \partial/\partial\omega_2, \partial/\partial\omega_3)$. But recall that the inertia ellipsoid is defined as a level surface of T, so equation (3.34) tells us that the angular momentum $\mathbf{L}$ is always perpendicular to the ellipsoid. This, in turn, ensures that the invariable plane is always tangent to the ellipsoid.

In summary, the angular velocity traces out two curves: one on the inertia ellipsoid, known as the polhode curve, and another on the invariable plane, known as the *herpolhode* curve. The body moves as if it is embedded within the inertia ellipsoid, which rolls around the invariable plane without slipping, with the centre of the ellipsoid a constant

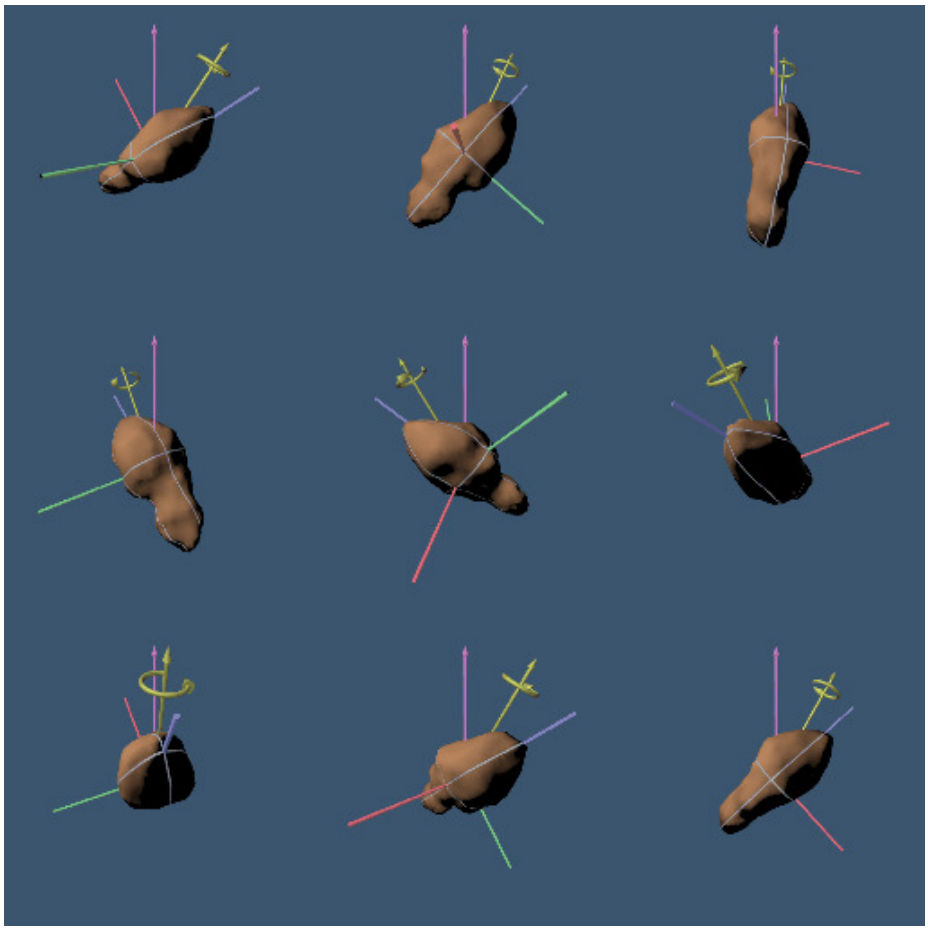


Figure 30: By Toutatis! The three principal axes are shown in red, green and blue (without arrows). The angular momentum $\mathbf{L}$ is the vertical, purple arrow. The angular velocity $\boldsymbol{\omega}$ is the circled, yellow arrow.

distance from the plane. The motion is shown in Figure 29. Unlike the polhode curve, the herpolhode curve does not necessarily close.

An Example: The Asteroid Toutatis

Astronomical objects are usually symmetric, but there's an important exception wandering around our solar system, shown[2] in Figure 30 This is the asteroid Toutatis.

In September 2004, Toutatis passed the earth at a distance of about four times that to the moon. This is (hopefully!) the closest any asteroid will come for the next 60

[2]This picture was created by Scott Hudson of Washington State University. You can find many interesting facts about the asteroid on his webpage.

years. The orbit of Toutatis is thought to be chaotic, which could potentially be bad news for Earth a few centuries from now. As you can see from the picture, its tumbling motion is complicated. It is aperiodic. The pictures show the asteroid at intervals of a day. The angular momentum vector $\mathbf{L}$ remains fixed and vertical throughout the motion. The angular velocity $\boldsymbol{\omega}$ traces out the herpolhode curve in the horizontal plane, perpendicular to $\mathbf{L}$. The angular momentum vector $\boldsymbol{\omega}$ also traces out a curve over the asteroid's surface: this is the polhode curve. It has a period of 5.4 days which you can observe by noting that $\boldsymbol{\omega}$ has roughly the same orientation relative to the principal axes every five to six days.

However, there are further effects at play in a real object like Toutatis which is not spinning around a principal axis. Various stresses and strains lead to dissipation. This means that the angular velocity $\boldsymbol{\omega}$ does not quite follow the polhode curve. Instead it begins close to the major axis ω_3 and slowly spirals towards the minor axis ω_1. This is why we see so few wobbling asteroids.

3.4 Euler Angles

So far, we've made a fair bit of progress simply by using the relation $\mathbf{L} = I\boldsymbol{\omega}$ between angular momentum and angular velocity, and interpreting what this means for the tumbling motion of rigid bodies. Notice, in particular, that we haven't needed to say anything about what this motion looks like in the configuration space $\mathcal{C}$ of the rigid body.

However, to make further progress we will need to get a better handle on the way to describe the orientation of the rigid body. Recall that the configuration space is given by all possible rotations of a set of axes. An example of such a rotation is shown in Figure 31. Our first task is to find a way to explicitly parameterise such a rotation. Happily, there is a way do do this, due to Euler, that often leads to simple solutions.

As an aside, we mentioned previously that the configuration space is actually the group $\mathcal{C} = SO(3)$. This means that we will be looking for a way to parameterise the group manifold $SO(3)$. We will use the following fact:

Euler's Theorem: An arbitrary rotation may be expressed as the product of 3 successive rotations about 3 (in general) different axes.

Proof: Let $\{\tilde{\mathbf{e}}_a\}$ be space frame axes. Let $\{\mathbf{e}_a\}$ be body frame axes. We want to find the rotation R so that $\mathbf{e}_a = R_{ab}\tilde{\mathbf{e}}_b$. We can accomplish this in three steps:

$$\{\tilde{\mathbf{e}}_a\} \xrightarrow{R_3(\phi)} \{\mathbf{e}'_a\} \xrightarrow{R_1(\theta)} \{\mathbf{e}''_a\} \xrightarrow{R_3(\psi)} \{\mathbf{e}_a\} \tag{3.35}$$

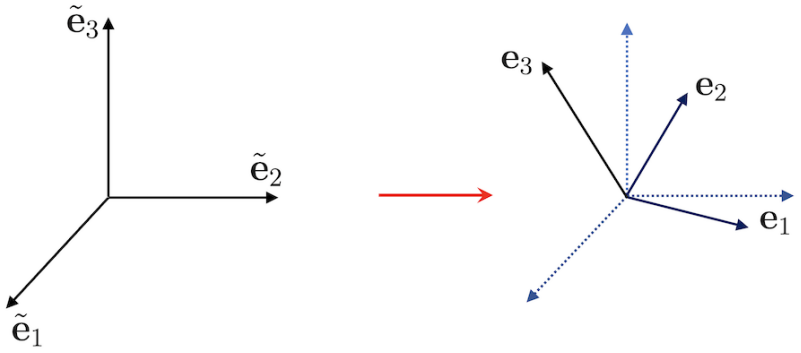


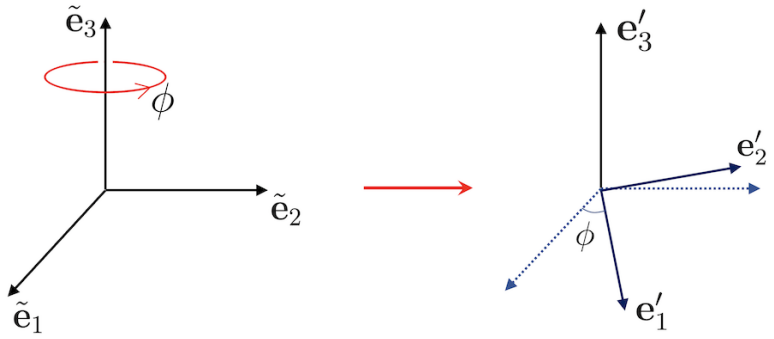
Figure 31: The rotation from space frame $\{\tilde{\mathbf{e}}_a\}$ to body frame $\{\mathbf{e}_a\}$.

Let's look at these step in turn.

Step 1: Rotate by ϕ about the $\tilde{\mathbf{e}}_3$ axis. So $\mathbf{e}'_a = R_3(\phi)_{ab}\tilde{\mathbf{e}}_b$ with

$$R_3(\phi) = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{3.36}$$

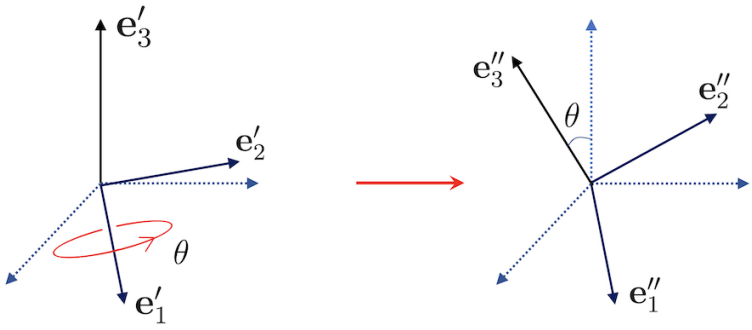
This rotates the axes like this:



Step 2: The next step is to totate by θ about the *new* axis $\mathbf{e}'_1$. This axis $\mathbf{e}'_1$ is sometimes called the "line of nodes". We write $\mathbf{e}''_a = R_1(\theta)_{ab}\mathbf{e}'_b$ with

$$R_1(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix} \tag{3.37}$$

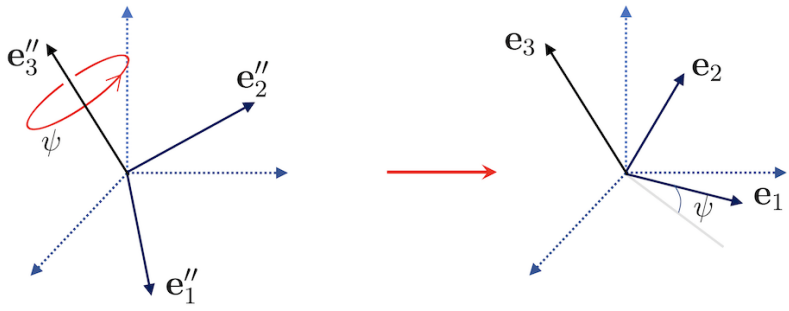
This time, it acts on the axes like this:



Step 3: Finally, we rotate by ψ about the *new new* axis $\mathbf{e}''_3$ so $\mathbf{e}_a = R_3(\psi)_{ab}\mathbf{e}''_b$ with

$$R_3(\psi) = \begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{3.38}$$

This final rotation looks like this:



This brings us to the promised rotation shown in Figure 31. Putting it all together, we have

$$R_{ab}(\phi,\theta,\psi) = [R_3(\psi)R_1(\theta)R_3(\phi)]_{ab} \tag{3.39}$$

This is our claimed result. The angles ϕ, θ and ψ are known as *Euler angles*. □

If we write out the matrix $R(\phi,\theta,\psi)$ in longhand, it reads

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

Before we proceed, a warning. Here we've thought about the rotations as acting on the basis vectors $\{\mathbf{e}_a\}$. Some textbooks instead choose to think of the rotations as acting on the components of a vector $\mathbf{r} = r_a\mathbf{e}_a = \tilde{r}_a\tilde{\mathbf{e}}_a$, writing $\tilde{r}_b = r_a R_{ab}$. When expanded in terms of Euler angles, this can result in an apparent reversal of the ordering of the three rotation matrices in (3.39).

3.4.1 Leonhard Euler (1707-1783)

As is clear from the section headings, the main man for this chapter is Euler, by far the most prolific mathematician of all time. As well as developing the dynamics of rotations, he made huge contributions to the fields of number theory, geometry, topology, analysis and fluid dynamics. For example, the lovely equation $e^{i\theta} = \cos\theta + i\sin\theta$ is due to Euler. In 1744 he was the first to correctly present a limited example of the calculus of variations (which we saw in section 2.1) although he generously gives credit to a rather botched attempt by his friend Maupertuis in the same year. Euler also invented much of the modern notation of mathematics: $f(x)$ for a function; e for exponential; π for, well, π and so on.

Euler was born in Basel, Switzerland and held positions in St Petersburg, Berlin and, after falling out with Frederick the Great, St Petersburg again. He was pretty much absorbed with mathematics day and night. Upon losing the sight in his right eye in his twenties he responded with: "Now I will have less distraction". Even when he went completely blind later in life, it didn't slow him down much as he went on to produce over half of his total work. The St Petersburg Academy of Science continued to publish his work for a full 50 years after his death.

3.4.2 Angular Velocity

There is a simple expression for the instantaneous angular velocity $\boldsymbol{\omega}$ in terms of Euler angles. One way straightforward, but slightly tedious, way to derive this is to plug in the the expression (3.39) into the definition (3.6) of the angular momentum matrix $\Omega = \dot{R}R^{-1}$. However, a little thought about what this means physically will get us there quicker. Consider the motion of a rigid body in an infinitesimal time dt, during which each of the Euler angles will change a little

$$(\psi, \theta, \phi) \to (\psi + d\psi, \theta + d\theta, \phi + d\phi) \tag{3.40}$$

From the definition of the Euler angles, the angular velocity must be of the form

$$\boldsymbol{\omega} = \dot{\phi}\,\tilde{\mathbf{e}}_3 + \dot{\theta}\,\mathbf{e}'_1 + \dot{\psi}\,\mathbf{e}_3 \tag{3.41}$$

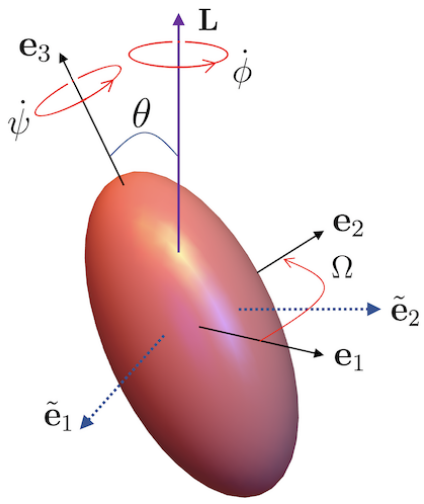


Figure 32: Euler angles for the free symmetric top when $\mathbf{L}$ coincides with $\tilde{\mathbf{e}}_3$

We just need to write each of these vectors in the same basis. We'll choose to write everything in the body frame basis $\{\mathbf{e}_a\}$. We can write the first two vectors above as

$$\begin{aligned}\tilde{\mathbf{e}}_3 &= \sin\theta\sin\psi\,\mathbf{e}_1 + \sin\theta\cos\psi\,\mathbf{e}_2 + \cos\theta\,\mathbf{e}_3\\ \mathbf{e}_1' &= \cos\psi\,\mathbf{e}_1 - \sin\psi\,\mathbf{e}_2\end{aligned}$$

from which we find the expression for the angular velocity $\boldsymbol{\omega}$ in terms of Euler angles in the body frame axis

$$\boldsymbol{\omega} = [\dot{\phi}\sin\theta\sin\psi + \dot{\theta}\cos\psi]\mathbf{e}_1 + [\dot{\phi}\sin\theta\cos\psi - \dot{\theta}\sin\psi]\mathbf{e}_2 + [\dot{\psi} + \dot{\phi}\cos\theta]\mathbf{e}_3 \qquad (3.42)$$

By playing a similar game, we can also express $\boldsymbol{\omega}$ in the space frame axis.

3.4.3 The Free Symmetric Top Revisited

To get some intuition for Euler angles, we can return to the free symmetric top that we looked at in Section 3.3.2. There we studied the angular velocity components in the body frame and found that ω_3 is constant while, as shown in (3.28), ω_1 and ω_2 precess as

$$(\omega_1, \omega_2) = \omega_0(\sin\Omega t, \cos\Omega t) \quad \text{with} \quad \Omega = \omega_3\frac{(I_1 - I_3)}{I_1} \qquad (3.43)$$

But what does this look like in the space frame? Now that we have parametrised motion in the space frame in terms of Euler angles, we can answer this question. This is simplest if we choose the angular momentum $\mathbf{L}$ to lie along the $\tilde{\mathbf{e}}_3$ space-axis. The resulting rigid body, garlanded with its many angles, is shown in Figure 32.

Because both $\mathbf{L}$ and its body-frame component $L_3 = I_3\omega_3$ are conserved, so too is the angle between them. But this angle is identified with the Euler angle θ. This means that $\dot{\theta} = 0$.

Next, we want to compute $\dot{\psi}$. The simplest way to do this is to state at Figure 32 and convince yourself that

$$\dot{\psi} = \Omega$$

Alternatively, you can reach the same result by comparing the expression for the angular velocity (3.42) with our result (3.43) for the precession. Next, we can get an expression for $\dot{\phi}$ by using the component of the body frame angular velocity

$$\omega_3 = \dot{\psi} + \dot{\phi}\cos\theta$$

Substituting our expressino $\Omega = \dot{\psi}$, we find the precession frequency

$$\dot{\phi} = \frac{I_3\omega_3}{I_1\cos\theta} \tag{3.44}$$

This is the rate of the wobble as seen from the space frame.

An Example: The Wobbling Plate

The physicist Richard Feynman tells the following story:

> *"I was in the cafeteria and some guy, fooling around, throws a plate in the air. As the plate went up in the air I saw it wobble, and I noticed the red medallion of Cornell on the plate going around. It was pretty obvious to me that the medallion went around faster than the wobbling.*
>
> *I had nothing to do, so I start figuring out the motion of the rotating plate. I discover that when the angle is very slight, the medallion rotates twice as fast as the wobble rate – two to one. It came out of a complicated equation!*
>
> *I went on to work out equations for wobbles. Then I thought about how the electron orbits start to move in relativity. Then there's the Dirac equation in electrodynamics. And then quantum electrodynamics. And before I knew it....the whole business that I got the Nobel prize for came from that piddling around with the wobbling plate."*

Feynman was right about quantum electrodynamics. But what about the plate?

We can look at this easily using what we've learnt. The spin of the plate is ω_3, while the precession, or wobble, rate $\dot{\phi}$ is given in (3.44). The relevant moments of inertia for a plate were calculated in Section 3.2 where we found that $I_3 = 2I_1$. From (3.43), we then have $\Omega = -\omega_3$. We can use this to see that $\dot{\psi} = -\omega_3$ for this example and so, for slight angles θ, with $\cos\theta \approx 1$, we have

$$\dot{\phi} \approx -2\dot{\psi} \tag{3.45}$$

This means that the wobble rate of the plate is twice as fast as the spin. It's the opposite to how Feynman remembers!

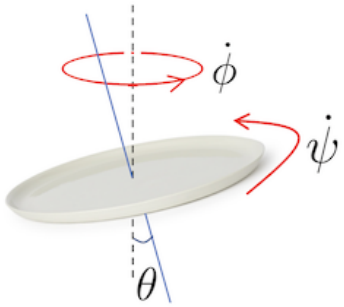


Figure 33:

There is another elegant and simple method you can use to see that Feynman was wrong: you can pick up a plate and throw it. It's hard to see that the wobble to spin ratio is exactly two. But it's easy to see that it wobbles faster than it spins.

3.5 The Heavy Symmetric Top

Until now, all our rigid bodies have been free. They tumble and turn, only because of conservation of angular momentum. Now we start adding forces to the story. And the most obvious force to add is gravity.

We'll work with a symmetric top, pinned at a point P which is a distance l from the centre of mass as shown in Figure 34. The principal axes are $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$ and we have $I_1 = I_2$. It's straightforward to write down the Lagrangian in terms of Euler angles

$$\begin{aligned} L &= \tfrac{1}{2}I_1(\omega_1^2 + \omega_2^2) + \tfrac{1}{2}I_3\omega_3^2 - Mgl\cos\theta \\ &= \tfrac{1}{2}I_1(\dot{\theta}^2 + \sin^2\theta\dot{\phi}^2) + \tfrac{1}{2}I_3(\dot{\psi} + \cos\theta\,\dot{\phi})^2 - Mgl\cos\theta \end{aligned}$$

Clearly both $\psi$and ϕ are ignorable coordinates, giving us two constants of motion. The first of these is

$$p_\psi = I_3(\dot{\psi} + \cos\theta\,\dot{\phi}) = I_3\omega_3 \tag{3.46}$$

This is the angular momentum about the symmetry axis $\mathbf{e}_3$ of the top. The angular velocity ω_3 about this axis is called the *spin* of the top and, as for the free symmetric top, it is a constant. The other constant of motion is

$$p_\phi = I_1\sin^2\theta\,\dot{\phi} + I_3\cos\theta\,(\dot{\psi} + \dot{\phi}\cos\theta) \tag{3.47}$$

In addition to these two conjugate momenta, the total energy E is also conserved

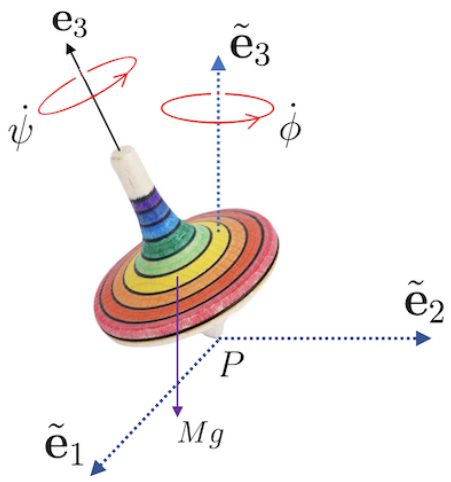


Figure 34: The heavy top with its Euler angles

$$E = T + V = \tfrac{1}{2} I_1 (\dot\theta^2 + \dot\phi^2 \sin^2\theta) + \tfrac{1}{2} I_3 \omega_3^2 + Mgl\cos\theta$$

To simplify these equations, we define the two constants

$$a = \frac{I_3\omega_3}{I_1} \quad \text{and} \quad b = \frac{p_\phi}{I_1} \tag{3.48}$$

Then we can write

$$\dot\phi = \frac{b - a\cos\theta}{\sin^2\theta} \tag{3.49}$$

and

$$\dot\psi = \frac{I_1 a}{I_3} - \frac{(b - a\cos\theta)\cos\theta}{\sin^2\theta} \tag{3.50}$$

So if we can solve $\theta = \theta(t)$ somehow, then we can always integrate these two equations to get $\phi(t)$ and $\psi(t)$. But first we have to figure out what θ is doing. To do this, we define the "reduced energy" $E' = E - \frac{1}{2} I_3 \omega_3^2$. Because both E and ω_3 are constant, so too is E'. We have

$$E' = \tfrac{1}{2} I_1 \dot\theta^2 + V_{\text{eff}}(\theta) \tag{3.51}$$

where the effective potential is

$$V_{\text{eff}}(\theta) = \frac{I_1 (b - a\cos\theta)^2}{2\sin^2\theta} + Mgl\cos\theta \tag{3.52}$$

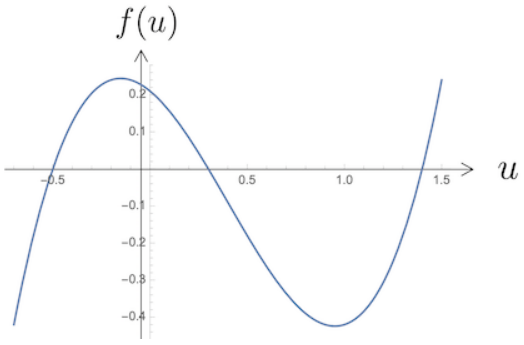


Figure 35: The cubic $f(u)$, defined in (3.54)

This now looks like the kind of things that we can understand, simply by plotting the effective potential $V_{\text{eff}}(\theta)$. It turns out to be simpler if we define the new coordinate

$$u = \cos\theta$$

Clearly $-1 \leq u \leq 1$. We'll also define two further constants to help put the equations in the most concise form

$$\alpha = \frac{2E'}{I_1} \quad \text{and} \quad \beta = \frac{2Mgl}{I_1} \tag{3.53}$$

With all these redefinitions, the equations of motion (3.49), (3.50) and (3.51) can be written as

$$\dot{u}^2 = (1-u^2)(\alpha - \beta u) - (b - au)^2 \equiv f(u) \tag{3.54}$$

$$\dot{\phi} = \frac{b - au}{1-u^2} \tag{3.55}$$

$$\dot{\psi} = \frac{I_1 a}{I_3} - \frac{u(b-au)}{1-u^2}$$

We could take the square root of equation (3.54) and integrate to reduce the problem to quadrature. The resulting integral is known as an *elliptic integral*. But, rather than doing this, there's a better way to understand the physics qualitatively.

Note that the function $f(u)$ defined in (3.54) is a cubic polynomial that behaves as

$$f(u) \to \begin{cases} +\infty \text{ as } u \to +\infty \\ -\infty \text{ as } u \to -\infty \end{cases}$$

and $f(\pm 1) = -(b \mp a)^2 \leq 0$. This means that if we plot the function $f(u)$ then it looks like figure 35

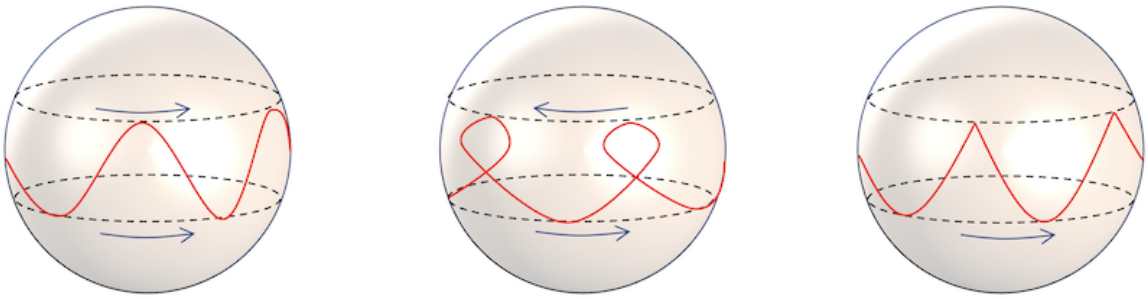


Figure 36: The three different types of motion depend on the direction of precession $\dot{\phi}$ when θ reaches its maximum and minimum values..

The physical range is $\dot{u}^2 = f(u) > 0$ and $-1 \leq u \leq 1$ so we find that, like in the spherical pendulum and central force problem, the system is confined to lie between the two roots of $f(u)$.

There are three possibilities for the motion depending on the sign of $\dot{\phi}$ at the two roots $u = u_1$ and $u = u_2$ as determined by (3.55). These are

- $\dot{\phi} > 0$ at both $u = u_1$ and $u = u_2$
- $\dot{\phi} > 0$ at $u = u_1$, but $\dot{\phi} < 0$ at $u = u_2$
- $\dot{\phi} > 0$ at $u = u_1$ and $\dot{\phi} = 0$ at $u = u_2$

The different paths of the top corresponding to these three possibilities are shown in figure 36. Motion in ϕ is called *precession* while motion in θ is known as *nutation*.

Letting the Top Go

The last of these three motions is not as unlikely as it may first appear. Suppose we spin the top and let it go at some angle θ. What happens? We have the initial conditions

$$\begin{aligned} & \dot{\theta}_{t=0} = 0 \quad \Rightarrow \quad f(u_{t=0}) = 0 \quad \Rightarrow \quad u_{t=0} = u_2 \\ \text{and} \quad & \dot{\phi}_{t=0} = 0 \quad \Rightarrow \quad b - a u_{t=0} = 0 \quad \Rightarrow \quad u_{t=0} = \frac{b}{a} \end{aligned}$$

Remember also that the quantity

$$p_\phi = I_1 \dot{\phi} \sin^2\theta + I_3 \omega_3 \cos\theta = I_3 \omega_3 \cos\theta_{t=0} \tag{3.56}$$

is a constant of motion. We now have enough information to figure out the qualitative motion of the top. Firstly, it starts to fall under the influence of gravity, so θ increases.

But as the top falls, $\dot{\phi}$ must turn and increase in order to keep p_ϕ constant. Moreover, we also see that the direction of the precession $\dot{\phi}$ must be in the same direction as the spin ω_3 itself. What we get is motion of the third kind.

3.5.1 Uniform Precession

Can we make the top precess without bobbing up and down? This would require $\dot{\theta} = 0$ and $\dot{\phi}$ constant. For this to happen, we would need the function $f(u)$ to have a single root u_0 lying in the physical range $-1 \leq u_0 \leq +1$, as shown in the figure. This root must satisfy,

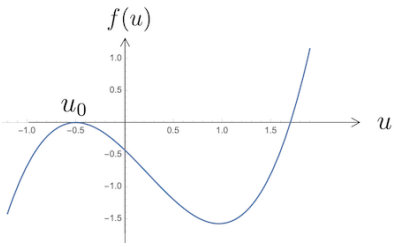


Figure 37:

$$f(u_0) = (1-u_0^2)(\alpha - \beta u_0) - (b - a u_0)^2 = 0$$
$$\text{and} \quad f'(u_0) = -2u_0(\alpha - \beta u_0) - \beta(1-u_0^2) + 2a(b - a u_0) = 0$$

Combining these, we find $\frac{1}{2}\beta = a\dot{\phi} - \dot{\phi}^2 u_0$. Substituting the definitions $I_1 a = I_3\omega_3$ and $\beta = 2Mgl/I_1$ into this expression, we find

$$Mgl = \dot{\phi}(I_3\omega_3 - I_1\dot{\phi}\cos\theta_0) \tag{3.57}$$

The interpretation of this equation is as follows: for a fixed value of ω_3 (the spin of the top) and θ_0 (the angle at which you let it go), we need to give exactly the right push $\dot{\phi}$ to make the top spin without bobbing. In fact, since equation (3.57) is quadratic in $\dot{\phi}$, there are two frequencies with which the top can precess without bobbing.

Of course, these "slow" and "fast" precessions only exist if equation (3.57) has any solutions at all. Since it is quadratic, this is not guaranteed, but requires

$$\omega_3 > \frac{2}{I_3}\sqrt{MglI_1\cos\theta_0} \tag{3.58}$$

So we see that, for a given θ_0, the top has to be spinning fast enough in order to have uniform solutions. What happens if it's not spinning fast enough? Well, the top falls over!

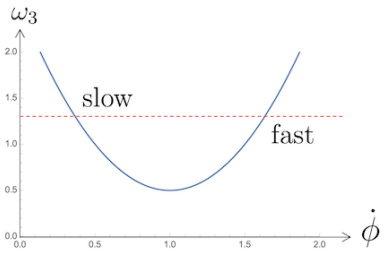


Figure 38:

3.5.2 The Sleeping Top

Suppose we start the top spinning in an upright position, with

$$\theta = \dot{\theta} = 0 \tag{3.59}$$

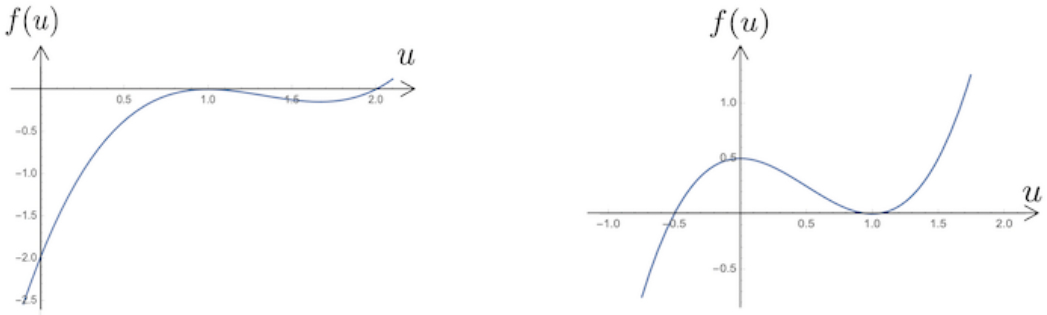


Figure 39: The function $f(u)$ for the stable sleeping top (on the left) and the unstable sleeping top (on the right).

When it spins upright, it is called a *sleeping top*. The question we want to answer is: will it stay there? Or will it fall over? From (3.54), we see that the function $f(u)$ must have a root at $\theta = 0$, or $u = +1$: $f(1) = 0$. From the definitions (3.48) and (3.53), we can check that $a = b$ and $\alpha = \beta$ in this situation and $f(u)$ actually has a double zero at $u = +1$,

$$f(u) = (1-u)^2(\alpha(1+u) - a^2) \tag{3.60}$$

The second root of $f(u)$ is at $u_2 = a^2/\alpha - 1$. There are two possibilities

- The first possibility is when $u_2 > 1$ or, equivalently, $\omega_3^2 > 4I_1Mgl/I_3^2$. In this case, the graph of $f(u)$ is drawn in first in figure 39. This motion is stable: if we perturb the initial conditions slightly, we will perturb the function $f(u)$ slightly, but the physical condition that we must restrict to the regime $f(u) > 0$ means that the motion will continue to be trapped near $u = 1$
- The second possibility is when $u_2 < 1$, or $\omega_3^2 < 4I_1Mgl/I_3^2$. In this case, the function $f(u)$ looks like the second figure of 39. Now the top is unstable; slight changes in the initial condition allow a large excursion.

In practice, the top spins upright until it is slowed by friction to $I_3\omega_3 = 2\sqrt{I_1Mgl}$, at which point it starts to fall and precess.

3.5.3 The Precession of the Equinox

The Earth spins at an angle of $\theta = 23.5^o$ to the plane of its orbit around the Sun, known as the *plane of the elliptic*. The spin of the earth is familiar: it is $\dot{\psi} = (\text{day})^{-1}$. This spin causes the Earth to bulge at the equator so it is no longer a sphere, but rather a

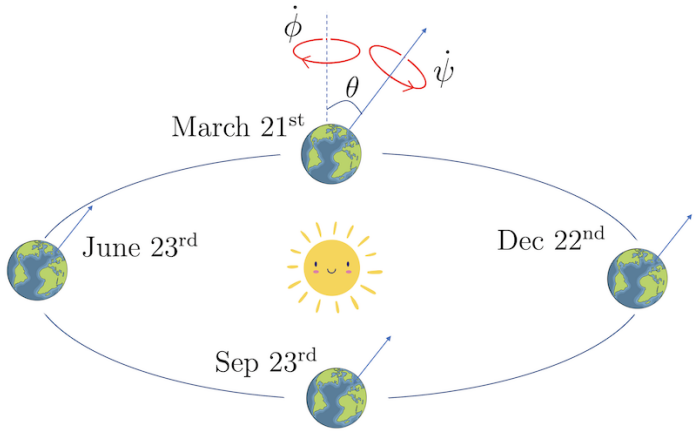


Figure 40: The precession of the Earth. (Not to scale.)

symmetric top. This, in turn, allows both the Sun and Moon to exert a torque on the Earth which produces a precession $\dot{\phi}$. Physically this means that the direction in which the North Pole points traces a circle in the sky and what we currently call the "pole star" will no longer be in several thousand years time. It turns out that this precession is "retrograde" i.e. opposite to the direction of the spin. The relevant angles are shown in Figure 40.

One can calculate the precession $\dot{\phi}$ of the earth due to the moon and sun using the techniques described in the chapter. We won't. Instead, we will use a more novel technique to calculate the precession of the earth: astrology. After all, it's got to be good for something[3].

To compute the precession of the earth, the first fact we need to know is that Jesus was born in the age of Pisces. This doesn't mean that Jesus looked up Pisces in his daily horoscope (while scholars are divided over the exact date of his birth, he seems to exhibit many traits of a typical Capricorn) but rather refers to the patch of the sky in which the sun appears during the first day of spring. Known in astronomical terms as the "vernal equinox", this day of the year is defined by the property that the sun sits directly above the equator at midday. As the earth precesses, this event takes place at a slightly different point in its orbit each year, with a slightly different backdrop of stars as a result. The astrological age is defined to be the background constellation in which the sun rises during vernal equinox.

[3]I learned about this fact from John Baez' website.

It is easy to remember that Jesus was born in the age of Pisces since the fish was used as an early symbol for Christianity. The next fact that we need to know is that we're currently entering the age of Aquarius (which anyone who has seen the musical Hair will know). So we managed to travel backwards one house of the zodiac in 2,000 years. We've got to make it around 12 in total, giving us a precession time of $2,000 \times 12 = 24,000$ years. The actual value of the precession is $25,700$ years. Our calculation is pretty close considering the method!

The earth also undergoes other motion. The value of θ varies from 22.1^o to 24.5^o over a period of $41,000$ years, mostly due to the effects of the other planets. These also affect the eccentricity of the orbit over a period of 105,000 years.

3.6 The Motion of Deformable Bodies

Take a lively cat. (Not one that's half dead like Schrödinger's). Hold it upside down and drop it. The cat will twist its body and land sprightly on its feet[4]. Yet it doesn't do this by pushing against anything and its angular momentum is zero throughout. If the cat were rigid, such motion would be impossible since a change in orientation for a rigid body necessarily requires non-vanishing angular momentum. But the cat isn't rigid (indeed, it can be checked that dead cats are unable to perform this feat) and bodies that can deform are able to reorient themselves without violating the conservation of angular momentum. In this section we'll describe some of the beautiful mathematics that lies behind this. I should warn you that this material is somewhat more advanced than the motion of rigid bodies. The theory described below was first developed in the late 1980s in order to understand how micro-organisms swim[5].

3.6.1 Kinematics

We first need to describe the configuration space $\mathcal{C}$ of a deformable body. We factor out translations by insisting that all bodies have the same center of mass. Then the configuration space $\mathcal{C}$ is the space of all shapes with some orientation.

Rotations act naturally on the space $\mathcal{C}$: they simply rotate each shape. This allows us to define the smaller *shape space* $\tilde{\mathcal{C}}$ so that any two configurations in $\mathcal{C}$ which are related by a rotation are identified in $\tilde{\mathcal{C}}$. In other words, any two objects that have the same shape, but different orientation, are described by different points in $\mathcal{C}$, but the same point in $\tilde{\mathcal{C}}$. Mathematically, we say $\tilde{\mathcal{C}} \cong \mathcal{C}/SO(3)$.

[4]Here is Étienne-Jules Marey's famous video of a falling cat from 1894.

[5]See A. Shapere and F. Wilczek, "*Geometry of Self-Propulsion at Low Reynolds Number*", J. Fluid Mech. **198** 557 (1989) . This is the same Frank Wilczek who won the 2004 Nobel prize for his work on quark interactions.

Figure 41: Three possible shapes of a deformable object.

We can describe this in more detail for a body consisting of N point masses, each with position $\mathbf{r}_i$. Unlike in section 3.1, we do not require that the distances between particles are fixed, i.e. $|\mathbf{r}_i - \mathbf{r}_j| \neq$ constant. (However, there may still be some restrictions on the $\mathbf{r}_i$). The configuration space $\mathcal{C}$ is the space of all possible configurations $\mathbf{r}_i$. For each different shape in $\mathcal{C}$, we pick a representative $\tilde{\mathbf{r}}_i$ with some, fixed orientation. It doesn't matter what representative we choose — just as long as we pick one. These variables $\tilde{\mathbf{r}}_i$ are coordinates on the space shape $\tilde{\mathcal{C}}$. For each $\mathbf{r}_i \in \mathcal{C}$, we can always find a rotation matrix $R \in SO(3)$ such that

$$\mathbf{r}_i = R\,\tilde{\mathbf{r}}_i \tag{3.61}$$

As in section 3.1, we can always do this to continuous bodies. In this case, the configuration space $\mathcal{C}$ and the shape space $\tilde{\mathcal{C}}$ may be infinite dimensional. Examples of different shapes for a continuously deformable body are shown in figure 41.

We want to understand how an object rotates as it changes shape keeping its angular momentum fixed (for example, keeping $\mathbf{L} = 0$ throughout). The first thing to note is that we can't really talk about the rotation between objects of different shapes. (How would you say that the the third object in figure 41 is rotated with respect to the first or the second?). Instead, we should think of an object moving through a sequence of shapes before returning to its initial shape. We can then ask if there's been a net rotation. As the object moves through its sequence of shapes, the motion is described by a time dependent $\tilde{\mathbf{r}}_i(t)$, while the corresponding change through the configuration space is

$$\mathbf{r}_i(t) = R(t)\,\tilde{\mathbf{r}}(t) \tag{3.62}$$

where the 3×3 rotation matrix $R(t)$ describes the necessary rotation to go from our fixed orientation of the shape $\tilde{\mathbf{r}}$ to the true orientation. As in section 3.1.1, we can define the 3×3 anti-symmetric matrix that describes the instantaneous angular velocity of

the object. In fact, it will for once prove more useful to work with the "convective angular velocity" matrix defined around equation (3.11)

$$\tilde{\Omega} = R^{-1}\,\frac{dR}{dt} \tag{3.63}$$

This angular velocity is non-zero due to the changing shape of the object, rather than the rigid rotation that we saw before. Let's do a quick change of notation and write coordinates on the shape space $\tilde{\mathcal{C}}$ as x^A, with $A = 1, \ldots, 3N$ instead of in vector notation $\tilde{\mathbf{r}}_i$, with $i = 1, \ldots, N$. Then, since Ω is linear in time derivatives, we can write

$$\tilde{\Omega} = \Omega_A(x)\,\dot{x}^A \tag{3.64}$$

The component $\Omega_A(x)$ is the 3×3 angular velocity matrix induced if the shape changes from x^A to $x^A + \delta x^A$. It is independent of time: all the time dependence sits in the $\dot{x}^A$ factor which tells us how the shape is changing. The upshot is that for each shape $x \in \tilde{C}$, we have a 3×3 anti-symmetric matrix Ω_A associated to each of the $A = 1, \ldots, 3N$ directions in which the shape can change.

However, there is an ambiguity in defining the angular velocity Ω. This comes about because of our arbitrary choice of reference orientation when we picked a representative $\tilde{\mathbf{r}}_i \in \tilde{\mathcal{C}}$ for each shape. We could quite easily have picked a different orientation,

$$\tilde{\mathbf{r}}_i \to S(x^A)\,\tilde{\mathbf{r}}_i \tag{3.65}$$

where $S(x^A)$ is a rotation that, as the notation suggests, can vary for each shape x^A. If we pick this new set of representative orientations, then the rotation matrix R defined in (3.62) changes: $R(t) \to R(t)\,S^{-1}(x^A)$. Equation (3.63) then tells us that the angular velocity also change as

$$\Omega_A \to S\,\Omega_A\,S^{-1} + S\,\frac{\partial S^{-1}}{\partial x^A} \tag{3.66}$$

This ambiguity is related to the fact that we can't define the meaning of rotation between two different shapes. Nonetheless, we will see shortly that when we come to compute the net rotation of the same shape, this ambiguity will disappear, as it must. Objects such as Ω_A which suffer an ambiguity of form (3.66) are extremely important in modern physics and geometry. They are known as non-abelian *gauge potentials* to physicists, or as *connections* to mathematicians.

3.6.2 Dynamics

So far we've learnt how to describe the angular velocity $\tilde{\Omega}$ of a deformable object. The next step is to see how to calculate $\tilde{\Omega}$. We'll now show that, up to the ambiguity described in (3.66), the angular velocity $\tilde{\Omega}$ is specified by the requirement that the angular momentum $\mathbf{L}$ of the object is zero.

$$\begin{aligned}\mathbf{L} &= \sum_i m_i\, \mathbf{r}_i \times \dot{\mathbf{r}}_i \\ &= \sum_i m_i \left[(R\tilde{\mathbf{r}}_i) \times (R\dot{\tilde{\mathbf{r}}}_\mathbf{i}) + (\mathbf{R}\tilde{\mathbf{r}}_\mathbf{i}) \times (\dot{\mathbf{R}}\tilde{\mathbf{r}}_\mathbf{i}) \right] = 0 \end{aligned} \tag{3.67}$$

In components this reads

$$L_a = \epsilon_{abc} \sum_i m_i \left[R_{bd} R_{ce} (\tilde{\mathbf{r}}_i)_d (\dot{\tilde{\mathbf{r}}}_i)_e + R_{bd} \dot{R}_{ce} (\tilde{\mathbf{r}}_i)_d (\tilde{\mathbf{r}}_i)_e \right] = 0 \tag{3.68}$$

The vanishing $\mathbf{L} = 0$ is enough information to determine the following result:

Claim: The 3×3 angular velocity matrix $\tilde{\Omega}_{ab} = R^{-1}_{ac} \dot{R}_{cb}$ is given by

$$\tilde{\Omega}_{ab} = \epsilon_{abc}\, \tilde{I}^{-1}_{cd}\, \tilde{L}_d \tag{3.69}$$

where $\tilde{I}$ is the instantaneous inertia tensor of the shape described by $\tilde{\mathbf{r}}_i$,

$$\tilde{I}_{ab} = \sum_i m_i ((\tilde{\mathbf{r}}_i \cdot \tilde{\mathbf{r}}_i)\delta_{ab} - (\tilde{\mathbf{r}}_i)_a (\tilde{\mathbf{r}}_i)_b) \tag{3.70}$$

and $\tilde{L}_a$ is the apparent angular momentum

$$\tilde{L}_a = \epsilon_{abc} \sum_i m_i\, (\tilde{\mathbf{r}}_i)_b (\dot{\tilde{\mathbf{r}}}_i)_c \tag{3.71}$$

Proof: We start by multiplying L_a by ϵ_{afg}. We need to use the fact that if we multiply two ϵ-symbols, we have $\epsilon_{abc}\epsilon_{afg} = (\delta_{bf}\delta_{cg} - \delta_{bg}\delta_{cf})$. Then

$$\begin{aligned}\epsilon_{afg} L_a = \sum_i m_i \big[& R_{fd} R_{ge} (\tilde{\mathbf{r}}_i)_d (\dot{\tilde{\mathbf{r}}}_i)_e - R_{gd} R_{fe} (\tilde{\mathbf{r}}_i)_d (\dot{\tilde{\mathbf{r}}}_i)_e \\ & - R_{gd} \dot{R}_{fe} (\tilde{\mathbf{r}}_i)_d (\tilde{\mathbf{r}}_i)_e + R_{fd} \dot{R}_{ge} (\tilde{\mathbf{r}}_i)_d (\tilde{\mathbf{r}}_i)_e \big] = 0 \end{aligned} \tag{3.72}$$

Now multiply by $R_{fb}R_{gc}$. Since R is orthogonal, we known that $R_{fb}R_{fd} = \delta_{bd}$ which, after contracting a bunch of indices, gives us

$$R_{fb}R_{gc}\epsilon_{afg}L_a = \sum_i m_i \left[(\tilde{\mathbf{r}}_i)_b (\dot{\tilde{\mathbf{r}}}_i)_c - (\tilde{\mathbf{r}}_i)_c (\dot{\tilde{\mathbf{r}}}_i)_b - \tilde{\Omega}_{bd} (\tilde{\mathbf{r}}_i)_c (\tilde{\mathbf{r}}_i)_d + \tilde{\Omega}_{cd} (\tilde{\mathbf{r}}_i)_b (\tilde{\mathbf{r}}_i)_d \right] = 0$$

This is almost in the form that we want, but the indices aren't quite contracted in the right manner to reproduce (3.69). One can try to play around to get the indices working right, but at this stage it's just as easy to expand out the components explicitly. For example, we can look at

$$\begin{aligned}\tilde{L}_1 &= \sum_i m_i \left[(\tilde{\mathbf{r}}_i)_2(\dot{\tilde{\mathbf{r}}}_i)_3 - ((\tilde{\mathbf{r}}_i)_3(\dot{\tilde{\mathbf{r}}}_i)_2\right] \\ &= \sum_i m_i \left[\tilde{\Omega}_{21}(\tilde{\mathbf{r}}_i)_3(\tilde{\mathbf{r}}_i)_1 + \tilde{\Omega}_{23}(\tilde{\mathbf{r}}_i)_3(\tilde{\mathbf{r}}_i)_3 - \tilde{\Omega}_{31}(\tilde{\mathbf{r}}_i)_2(\tilde{\mathbf{r}}_i)_1 - \tilde{\Omega}_{32}(\tilde{\mathbf{r}}_i)_2(\tilde{\mathbf{r}}_i)_2\right] \\ &= \tilde{I}_{11}\tilde{\Omega}_{23} + \tilde{I}_{12}\tilde{\Omega}_{31} + \tilde{I}_{13}\tilde{\Omega}_{12} = \tfrac{1}{2}\,\epsilon_{abc}\tilde{I}_{1a}\tilde{\Omega}_{bc} \end{aligned} \tag{3.73}$$

where the first equality is the definition of $\tilde{L}_1$, while the second equality uses our result above, and the third equality uses the definition of $\tilde{I}$ given in (3.70). There are two similar equations, which are summarised in the formula

$$\tilde{L}_a = \tfrac{1}{2}\,\epsilon_{bcd}\tilde{I}_{ab}\tilde{\Omega}_{cd} \tag{3.74}$$

Multiplying both sides by $\tilde{I}^{-1}$ gives us precisely the claimed result (3.69). This concludes the proof. □.

To summarise: a system with no angular momentum that can twist and turn and change its shape has an angular velocity (3.69) where $\tilde{\mathbf{r}}_i(t)$ is the path it chooses to take through the space of shapes. This is a nice formula. But what do we do with it? We want to compute the net rotation R as the body moves through a sequence of shapes and returns to its starting point at a time T later. This is given by solving (3.63) for R. The way to do this was described in section 3.1.2. We use *path ordered exponentials*

$$R = \tilde{P}\exp\left(\int_0^T \tilde{\Omega}(t)\,dt\right) = \tilde{P}\exp\left(\oint \Omega_A\,dx^A\right) \tag{3.75}$$

The path ordering symbol $\tilde{P}$ puts all matrices evaluated at later times to the right. (This differs from the ordering in section 3.1.2 where we put later matrices to the left. The difference arises because we're working with the convective angular velocity $\tilde{\Omega} = R^{-1}\dot{R}$ instead of the angular velocity $\Omega = \dot{R}R^{-1}$). In the second equality above, we've written the exponent as an integral around a closed path in shape space. Here time has dropped out. This tells us an important fact: it doesn't matter how quickly we perform the change of shapes — the net rotation of the object will be the same.

In particle physics language, the integral in (3.75) is called a "Wilson loop". We can see how the rotation fares under the ambiguity (3.64). After some algebra, you can find that the net rotation R of an object with shape x^A is changed by

$$R \to S(x^A)\,R\,S(x^A)^{-1} \tag{3.76}$$

This is as it should be: the S^{-1} takes the shape back to our initial choice of standard orientation; the matrix R is the rotation due to the change in shape; finally S puts us back to the new, standard orientation. So we see that even though the definition of the angular velocity is plagued with ambiguity, when we come to ask physically meaningful questions — such as how much has a shape rotated — the ambiguity disappears. However, if we ask nonsensical questions — such as the rotation between two different shapes — then the ambiguity looms large. In this manner, the theory contains a rather astonishing new ingredient: it lets us know what are the sensible questions to ask! Quantities for which the ambiguity (3.64) vanishes are called *gauge invariant*.

In general, it's quite hard to explicitly compute the integral (3.75). One case where it is possible is for infinitesimal changes of shape. Suppose we start with a particular shape x^0_A, and move infinitesimally in a loop in shape space:

$$x_A(t) = x^0_A + \alpha_A(t) \tag{3.77}$$

Then we can Taylor expand our angular velocity components,

$$\Omega_A(x(t)) = \Omega_A(x^0) + \left.\frac{\partial \Omega_A}{\partial x^B}\right|_{x^0} \alpha_B \tag{3.78}$$

Expanding out the rotation matrix (3.75) and taking care with the ordering, one can show that

$$\begin{aligned} R &= 1 + \frac{1}{2} F_{AB} \oint \alpha_A \dot{\alpha}_B \, dt + \mathcal{O}(\alpha^3) \\ &= 1 + \frac{1}{2} \int F_{AB} \, dA_{AB} + \mathcal{O}(\alpha^3) \end{aligned} \tag{3.79}$$

where F_{AB} is anti-symmetric in the shape space indices A and B, and is a 3×3 matrix (the $a, b = 1, 2, 3$ indices have been suppressed) given by

$$F_{AB} = \frac{\partial \Omega_A}{\partial x^B} - \frac{\partial \Omega_B}{\partial x^A} + [\Omega_A, \Omega_B] \tag{3.80}$$

It is known as the *field strength* to physicists (or the *curvature* to mathematicians). It is evaluated on the initial shape x^0_A. The second equality in (3.79) gives the infinitesimal rotation as the integral of the field strength over the area traversed in shape space. This field strength contains all the information one needs to know about the infinitesimal rotations of objects induced by changing their shape.

One of the nicest things about the formalism described above is that it mirrors very closely the mathematics needed to describe the fundamental laws of nature, such as the strong and weak nuclear forces and gravity. They are all described by "non-abelian gauge theories", with an object known as the gauge potential (analogous to Ω_A) and an associated field strength.

4. The Hamiltonian Formalism

In this section, we describe a different approach to classical mechanics, initially developed by Hamilton around 1830. Our goal here is not to solve increasingly complicated problems. Instead, we will use the Hamiltonian formalism to the structure underlying classical dynamics. If you like, it will help us understands what questions we should ask. In addition, we will start to see striking connections to quantum mechanics.

4.1 Hamilton's Equations

The Lagrangian formulation of classical mechanics starts with a function, the eponymous the Lagrangian, $L(q^i, \dot{q}^i, t)$. As the notation shows, this function depends on n generalised coordinates, q^i, $i = 1, \ldots, n$, their time derivatives $\dot{q}^i$ and, possibly, time. The Euler-Lagrange equations are

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} = 0 \tag{4.1}$$

These are n 2^{nd} order differential equations. Crucially, if we wish to find a solution then we have to give $2n$ initial conditions, say q^i and $\dot{q}^i$ at time $t = 0$.

Roughly speaking, the basic idea of Hamilton's approach is to try and place positions and velocities on a more symmetric footing. The "rough speaking" here is because it turns out that it's not quite the velocities that we should consider: instead it is the generalised momenta that we introduced in Section 2.2,

$$p_i = \frac{\partial L}{\partial \dot{q}^i} \tag{4.2}$$

In terms of the generalised momenta, the Euler-Lagrange equations (4.1) take the particularly simple form

$$\dot{p}_i = \frac{\partial L}{\partial q^i} \tag{4.3}$$

We should think of these generalised momenta as functions $p_i(q^j, \dot{q}^j, t)$. In many familiar situations the Lagrangian is quadratic in velocities (because the kinetic energy is $\frac{1}{2}m\mathbf{v}^2$) in which case the momenta will be linearly related to the velocities. But this need not be the case, and we'll see examples in this section where p and $\dot{q}$ have a different relationship.

Now we can state more precisely the key idea of the Hamiltonian formalism: it is to put coordinates q^i and momenta p^i on a more symmetric footing.

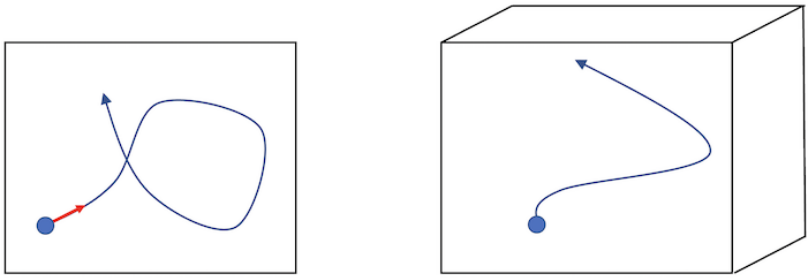


Figure 42: Motion in configuration space on the left, and in phase space on the right.

Phase Space

First, a definition. Recall that the generalised coordinates q^i parameterise the configuration space $\mathcal{C}$ of the system. Time evolution can be viewed as a path in in $\mathcal{C}$. However, if we know the that the system sits at a point in $\mathcal{C}$ at some particular time t_0, that's not enough information to determine where the system is at some later time. This is because we also need the initial conditions: we need $\dot{q}^i$ at time t_0 as well. Only then is the full evolution of the system uniquely determined.

As we mentioned above, throughout this section we will replace the velocities $\dot{q}^i$ in favour of the momenta p_i. So, alternatively, we could say that the future evolution of the system is uniquely determined by specifying the pair (q^i, p_i) at some moment in time. This pair (q^i, p_i) defines the *state* of the system. The space of all states comprise the $2n$-dimensional *phase space*, $\mathcal{M}$, parameterised by (q^i, p_i).

In contrast to configuration space, if you sit at a point in phase space at some time then your future (and, indeed past) evolution is uniquely determined. This means, among other things, that paths in phase space can never cross. We say that evolution is governed by a *flow* in phase space.

An Example: The Pendulum

Here's a simple example. The pendulum has a configuration space $\mathcal{C} = \mathbf{S}^1$, parameterised by an angle $\theta \in [-\pi, \pi)$. The phase space is then the cylinder $\mathcal{M} = \mathbf{R} \times \mathbf{S}^1$, with the $\mathbf{R}$ factor corresponding to the momentum p_θ. We've drawn this in Figure 43 by flattening out the cylinder, so that the left-hand edge at $\theta = -\pi$ and the right-hand edge at $\theta = \pi$ are identified. The phase space flows are also shown.

Two different types of motion are clearly visible in the flows. For small θ and small momentum, the pendulum oscillates back and forth, motion which appears as an ellipse

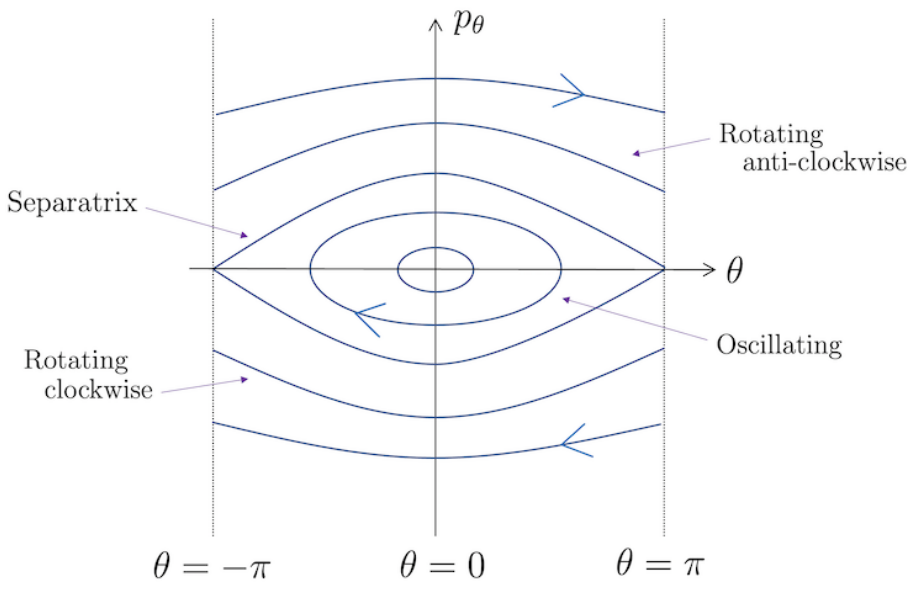


Figure 43: Flows in the phase space of a pendulum.

in phase space. But for large momentum, the pendulum swings all the way around, which appears as lines wrapping around the $\mathbf{S}^1$ of phase space. Separating these two different motions is the special case where the pendulum starts upright, falls, and just makes it back to the upright position. This curve in phase space is called the *separatrix*.

4.1.1 The Hamiltonian

We've seen that dynamical evolution is given by a flow on phase space. Our next goal is to write down equations that describe this flow. We do this by constructing the *Hamiltonian*. This is a function $H(q, p)$, on phase space. (Note that when we write $H(q, p)$, with no indices on q and p, we mean that it depends on all coordinates q^i and p_i with $i = 1, \ldots, n$.) The Hamiltonian is defined in terms of the Lagrangian by

$$H(q, p, t) = p_i \dot{q}^i - L(q, \dot{q}, t) \tag{4.4}$$

We've met something very similar to the Hamiltonian function before: we saw in (2.19) that, if $\partial L/\partial t = 0$, then the function $H(q, p)$ is identified as the conserved energy. This, of course, is also familiar from quantum mechanics where the Hamiltonian is identified as the energy operator.

However, there's more to the Hamiltonian than just the energy. In what follows, it's crucial that we think of the Hamiltonian as a function on phase space. This means that

H is a function of q^i and p_i. This is in contrast to the Lagrangian which is a function of q^i and $\dot{q}^i$.

In practice, this means that, given a Lagrangian, you first compute the momentum

$$p_i = \frac{\partial L}{\partial \dot{q}^i}$$

and then compute the right-hand side of (4.4). Often the resulting expression will be given as a function of q^i and $\dot{q}^i$. The last step is to then invert $\dot{q}^i = \dot{q}^i(q, p, t)$, so that the Hamiltonian takes the desired for $H(q, p, t)$ as a function over phase space.

The difference between thinking of $H(q, p, t)$ as a function of momentum vs velocities may seem like a pedantic one. But it's a difference on which the whole Hamiltonian formalism hangs. We'll see this in some examples as we move forwards. But, mathematically, the reason is that the Hamiltonian is a *Legendre transform* of the Lagrangian. We now take a small detour to explain what this means.

An Aside: The Legendre Transform

The Legendre transform arises in a number of places in physics. As we'e seen above, in classical mechanics it relates the Lagrangian and Hamiltonian. But we will also see the same transformation in thermodynamics, where it relates different functions of state, like energy and free energy. Here we give a brief account of the mathematics behind the Legendre transform.

First, consider an arbitrary function $f(x, y)$. Here the variable y will just go along for the ride. (It will be like the coordinates q in the Lagrangian.) Meanwhile, we're going to do the transform with respect to the variable x (which is like the $\dot{q}$ in the Lagrangian).

The total derivative of f is

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy$$

Next, define the related function

$$g(x, y, u) = ux - f(x, y) \tag{4.5}$$

The function $g(u, x, y)$ depends on the original two variable x and y, as well as a third variable u that, for now, just appears linearly in the definition. The total derivative of g is

$$dg = d(ux) - df = u\,dx + x\,du - \frac{\partial f}{\partial x}dx - \frac{\partial f}{\partial y}dy \tag{4.6}$$

So far, we've viewed u as a new independent variable. At this point, we choose it to be a specific function of x and y, defined by

$$u(x,y) = \frac{\partial f}{\partial x} \tag{4.7}$$

With this choice, the term proportional to dx in (4.6) vanishes. We have

$$dg = x\,du - \frac{\partial f}{\partial y}dy \tag{4.8}$$

This is telling us that g can be thought of as a function of just u and y. If we want an explicit expression for $g(u,y)$, we must first invert (4.7) to get $x = x(u,y)$ and then insert this into the definition (4.5), so that

$$g(u,y) = u\,x(u,y) - f(x(u,y),y) \tag{4.9}$$

This is the Legendre transform. It takes us from one function $f(x,y)$ to a different function $g(u,y)$ where $u = \partial f/\partial x$.

A crucial fact about the Legendre transformation is that (with one caveat to be explained below) it doesn't lose any information about the original function. To see this, we just need to note that we can recover the original $f(x,y)$ from $g(u,y)$ by doing a second Legendre transformation. We have

$$\left.\frac{\partial g}{\partial u}\right|_y = x(u,y) \qquad \text{and} \qquad \left.\frac{\partial g}{\partial y}\right|_u = -\frac{\partial f}{\partial y} \tag{4.10}$$

This assures us that the inverse Legendre transform $f = (\partial g/\partial u)u - g$ does indeed take us back to the original function.

We can get more insight into the Legendre transform by thinking geometrically. The figure shows the two curves $f(x,y)$ and ux for fixed y and for fixed u. The function $g(u;x)$, as defined in (4.5), is just the difference between these two curves. But the all-important condition (4.7) tells us that, for fixed u, the function $g(u)$ is the *maximum* distance between these curves. You can see this geometrically (because the maximum distance occurs when the tangents of the two curves are equal). Or you can see this algebraically by extremising the distance

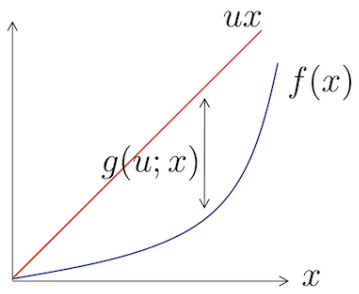


Figure 44:

$$\frac{d}{dx}\left(ux - f(x)\right) = 0 \quad \Rightarrow \quad u = \frac{\partial f}{\partial x}$$

This, then, is the geometrical interpretation of the Legendre transform. As you change u, you change the slope of the line ux. The function $g(u)$ is the maximum distance between that line and the original curve $f(x)$.

The graphical approach also hints at the caveat that we mentioned above. This is because there is only a unique, local maximum distance between the curves when the original function $f(x)$ is *convex*.

Hamilton's Equations

With a better understanding of the Legendre transform, we can now return to physics. We have the Hamiltonian $H(q, p, t)$, defined in (4.4), and we know from the discussion above that this contains the same information as our original Lagrangian $L(q, \dot{q}, t)$. Let's follow the discussion of the Legendre transform and consider the total derivative of the Hamiltonian

$$\begin{aligned} dH &= (\dot{q}^i dp_i + p_i \, d\dot{q}^i) - \left(\frac{\partial L}{\partial q^i} dq^i + \frac{\partial L}{\partial \dot{q}^i} d\dot{q}^i + \frac{\partial L}{\partial t} dt \right) \\ &= \dot{q}^i dp_i - \frac{\partial L}{\partial q^i} dq^i - \frac{\partial L}{\partial t} dt \end{aligned} \tag{4.11}$$

We see that the $d\dot{q}^i$ terms have cancelled upon using the definition of generalised momenta $p_i = \partial L/\partial \dot{q}^i$. This mimics the fact that dx term vanished in (4.8) and is telling us that we can think of $H = H(q, p, t)$. In other words, we must have

$$dH = \frac{\partial H}{\partial q^i} \, dq^i + \frac{\partial H}{\partial p_i} \, dp_i + \frac{\partial H}{\partial t} \, dt \tag{4.12}$$

Now we simply equate terms between (4.11) and (4.12). So far this is repeating the steps of the Legendre transformation. We now add a physics ingredient into the mix in the guise of the Euler-Lagrange equation (4.3) which reads $\dot{p}_i = \partial L/\partial q^i$. The result is a collection of first order differential equations,

$$\dot{p}_i = -\frac{\partial H}{\partial q^i} \quad \text{and} \quad \dot{q}^i = \frac{\partial H}{\partial p_i} \tag{4.13}$$

These are *Hamilton's equations*. In addition, if the Lagrangian L has explicit time dependence, then so too does the Hamiltonian with

$$-\frac{\partial L}{\partial t} = \frac{\partial H}{\partial t} \tag{4.14}$$

In Hamilton's equations (4.13), we have replaced n second order equations of motion with $2n$ first order equations of motion. This may not seem like a great leap forward.

In particular, when we come to solve these equations we will often revert back to the second order form. Nevertheless, as we promised above, Hamilton's equations are the starting point for peeling apart the structure of classical mechanics. We'll see more of this as we progress. For now, we just note that we can already see a glimpse of the symmetry between q and p in Hamilton's equations: both sit on a similar footing, with just a minus sign in the first equation of (4.13) to distinguish them.

Briefly, Conservation Laws

We saw in Section 2.3 how conservation laws arise in the Lagrangian formalism. While Noether's theorem lives most naturally in the Lagrangian formalism, some of the simpler results carry over straightforwardly to the Hamiltonian formalism.

For example, if $\partial H/\partial t = 0$ then H itself is a constant of motion. This follows by looking at the total time derivative

$$\frac{dH}{dt} = \frac{\partial H}{\partial q^i}\dot{q}^i + \frac{\partial H}{\partial p_i}\dot{p}_i + \frac{\partial H}{\partial t} = -\dot{p}_i\dot{q}^i + \dot{q}^i\dot{p}_i + \frac{\partial H}{\partial t} = \frac{\partial H}{\partial t}$$

where we've imposed the equations of motion in the second equality. In most examples, H will be identified with the energy.

Similarly, if an coordinate q doesn't appear in the Lagrangian then, by construction, it also doesn't appear in the Hamiltonian. The conjugate momentum p_q is then conserved since

$$\dot{p}_i = \frac{\partial H}{\partial \partial q^i} = 0 \tag{4.15}$$

We'll have more to say about conservation laws in the Hamiltonian formalism in Section 4.4.2.

4.1.2 The Principle of Least Action

The principle of least action is a central piece of the Lagrangian formalism of classical mechanics. We define the action as a functional of all possible paths $q^i(t)$ in configuration space,

$$S[q(t)] = \int_{t_1}^{t_2} L(q, \dot{q}, t)\, dt \tag{4.16}$$

The equations of motion then follow by insisting that $\delta S = 0$ for all paths with fixed end points so that $\delta q^i(t_1) = \delta q^i(t_2) = 0$.

There is a version of this story in the Hamiltonian formalism that goes through with only minor tweaks. This time, we define the action

$$S[q(t), p(t)] = \int_{t_1}^{t_2} (p_i \dot{q}^i - H) dt \tag{4.17}$$

where $\dot{q}^i = \partial H/\partial p_i$. This looks the same as our original action (4.16) but the devil is in the details: the action now is a function of all possible paths $q^i(t)$ and $p_i(t)$ in phase space. We now vary $q^i(t)$ and $p_i(t)$ independently. This differs from the Lagrangian formalism where a variation of $q^i(t)$ automatically leads to a variation of $\dot{q}^i$. But remember that the whole point of the Hamiltonian formalism is that we treat q^i and p_i on equal footing. So we vary both. We have

$$\begin{aligned} \delta S &= \int_{t_1}^{t_2} \left\{ \dot{q}^i \delta p_i + p_i \delta \dot{q}^i - \frac{\partial H}{\partial p_i} \delta p_i - \frac{\partial H}{\partial q^i} \delta q^i \right\} dt \\ &= \int_{t_1}^{t_2} \left\{ \left[\dot{q}^i - \frac{\partial H}{\partial p_i} \right] \delta p_i + \left[-\dot{p}_i - \frac{\partial H}{\partial q^i} \right] \delta q^i \right\} dt + \Big[p_i \delta q^i \Big]_{t_1}^{t_2} \end{aligned} \tag{4.18}$$

and there are Hamilton's equations waiting for us in the square brackets. If we look for extrema $\delta S = 0$ for all δp_i and δq^i we get Hamilton's equations

$$\dot{q}^i = \frac{\partial H}{\partial p_i} \quad \text{and} \quad \dot{p}_i = -\frac{\partial H}{\partial q^i} \tag{4.19}$$

Except there's a very slight subtlety with the boundary conditions. We need the last term in (4.18) to vanish, and so require only that

$$\delta q^i(t_1) = \delta q^i(t_2) = 0$$

while δp_i can be free at the end points $t = t_1$ and $t = t_2$. So, despite our best efforts, q^i and p_i are not quite symmetric in this formalism.

Note that we could simply impose $\delta p_i(t_1) = \delta p_i(t_2) = 0$ if we really wanted to and the above derivation still holds. It would mean we were being more restrictive on the types of paths we considered. But it does have the advantage that it keeps q^i and p_i on a symmetric footing. It also means that we have the freedom to add a function to consider actions of the form

$$S = \int_{t_1}^{t_2} \left(p_i \dot{q}^i - H(q,p) + \frac{dF(q,p)}{dt} \right)$$

so that what sits in the integrand differs from the Lagrangian. For some situations this may be useful.

4.1.3 A Particle in a Potential

We can illustrate the Hamiltonian formalism by looking at a particle moving in $\mathbf{R}^3$ in the presence of a potential $V(\mathbf{x})$. In this simple example, there are no surprises. (This will not continue to be the case as we look at more complicated examples.)

The Lagrangian is

$$L = \frac{1}{2}m\dot{\mathbf{x}}^2 - V(\mathbf{r})$$

The corresponding momentum takes the familiar form

$$\mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{x}}} = m\dot{\mathbf{x}}$$

The Hamiltonian is then given by

$$H = \mathbf{p}\cdot\dot{\mathbf{x}} - L = \frac{1}{2m}\mathbf{p}^2 + V(\mathbf{x})$$

In the final step, we've eliminated $\dot{\mathbf{r}}$ in favour of $\mathbf{p}$. Of course, this is trivial to do in the present example, but it is important nonetheless. It means that the kinetic energy is written as $\mathbf{p}^2/2m$ rather than the $m\dot{\mathbf{x}}^2/2$.

You may recall that that Hamiltonian plays a staring role in quantum mechanics. There too, it should be thought of as a function of momentum, rather than velocity, albeit with the momentum variable replaced by the operators $\mathbf{p} \to -i\hbar\nabla$.

Back in the classical world, Hamilton's equations (4.13) are

$$\dot{\mathbf{x}} = \frac{\partial H}{\partial \mathbf{p}} = \frac{1}{m}\mathbf{p} \quad \text{and} \quad \dot{\mathbf{p}} = -\frac{\partial H}{\partial \mathbf{x}} = -\nabla V$$

Both of these are familiar. The first just recapitulates the definition of velocity in terms of momentum, while the second is Newton's equation of motion for this system.

4.1.4 A Particle in an Electromagnetic Field

We described the Lagrangian for a particle with charge q moving in electric and magnetic fields in Section 2.4.3. One of the surprises was that the Lagrangian necessarily depends on the vector potential $\mathbf{A}$

$$L = \tfrac{1}{2}m\dot{\mathbf{x}}^2 - q\left(\phi - \dot{\mathbf{x}}\cdot\mathbf{A}\right)$$

The conjugate momentum is now given by

$$\mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{x}}} = m\dot{\mathbf{x}} + q\mathbf{A}$$

Crucially, the momentum $\mathbf{p}$ is not just "mass times velocity". It now gets an extra contribution from the vector potential. Moreover, as we stressed in Section 2.4.3, the momentum $\mathbf{p}$ is not gauge invariant, so it's not possible to ascribe physical meaning to the value of $\mathbf{p}$! These issues notwithstanding, it's trivial to invert the relationship above and write the velocity as

$$\dot{\mathbf{x}} = \frac{1}{m}\left(\mathbf{p} - q\mathbf{A}\right) \tag{4.20}$$

We now calculate the Hamiltonian

$$\begin{aligned} H(\mathbf{p},\mathbf{x}) &= \mathbf{p}\cdot\dot{\mathbf{x}} - L \\ &= \frac{1}{m}\mathbf{p}\cdot(\mathbf{p}-q\mathbf{A}) - \left[\frac{1}{2m}\left(\mathbf{p}-q\mathbf{A}\right)^2 - q\phi + \frac{q}{m}\left(\mathbf{p}-q\mathbf{A}\right)\cdot\mathbf{A}\right] \\ &= \frac{1}{2m}\left(\mathbf{p}-q\mathbf{A}\right)^2 + q\phi \end{aligned} \tag{4.21}$$

There's something curious about this expression. If we were to write the Hamiltonian in terms of the velocity $\dot{x}$, then the first term is just the usual kinetic energy $m\dot{\mathbf{x}}^2/2$ and the Hamiltonian coincides with the energy $E = m\dot{\mathbf{x}}^2/2 + q\phi$. This energy has no dependence on the vector potential $\mathbf{A}$ which is just the statement that magnetic fields do no work, and so don't affect the energy. But, as we've stressed above, the Hamiltonian must be thought of as a function of $\mathbf{p}$. And, when viewed as a function of $\mathbf{p}$, the gauge potential $\mathbf{A}$ sneaks back in. This means that the Hamiltonian depends on $\mathbf{A}$, and so magnetic fields can affect the equations of motion, even though the energy is independent of $\mathbf{A}$.

The first of Hamilton's equations reads

$$\dot{\mathbf{x}} = \frac{\partial H}{\partial \mathbf{p}} = \frac{1}{m}\left(\mathbf{p} - q\mathbf{A}\right)$$

which, as usual, is a recapitulation of our earlier expression (4.20). The second of Hamilton's equations, $\dot{\mathbf{p}} = -\partial H/\partial \mathbf{r}$, is best expressed in terms of components

$$\dot{p}_a = -\frac{\partial H}{\partial x^a} = -e\frac{\partial \phi}{\partial x^a} + \frac{e}{m}\left(p_b - eA_b\right)\frac{\partial A_b}{\partial x^a} \tag{4.22}$$

A little rearranging of indices shows that this coincides with the Lorentz force law (2.4.3).

4.1.5 The Relativistic Particle

We met two Lagrangians for a free, relativistic particle in Section 2.4.4. The first of these, (2.46), holds in a particular reference frame

$$S[\mathbf{x}(t)] = \int dt\ L[\dot{\mathbf{x}}] \quad \text{with} \quad L[\dot{\mathbf{x}}] = -mc^2\sqrt{1-\frac{\dot{\mathbf{x}}^2}{c^2}}$$

This has the advantage that it only depends on the physical degrees $\mathbf{x}(t)$, but has the disadvantage that Lorentz invariance is not manifest. We have the conjugate momentum

$$\mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{x}}} = m\gamma\dot{\mathbf{x}} \quad \text{with} \quad \gamma = \sqrt{\frac{1}{1-\dot{\mathbf{x}}^2/c^2}} \tag{4.23}$$

which we can invert to find

$$\dot{\mathbf{x}} = \frac{\mathbf{p}}{\sqrt{m^2+\mathbf{p}^2/c^2}} \tag{4.24}$$

The Hamiltonian is then

$$H[\mathbf{x},\mathbf{p}] = \mathbf{p}\cdot\dot{\mathbf{x}} - L = mc^2\gamma \tag{4.25}$$

This we recognise as the energy of a relativistic particle. The only catch is that, as usual, we should view the Hamiltonian as a function of the momentum $\mathbf{p}$ rather than the velocity. Substituting (4.24) into (4.23), we have

$$\gamma = \sqrt{1+\frac{\mathbf{p}^2}{m^2c^2}}$$

As a sanity check, if you Taylor expand the Hamiltonian (4.25) for small p, you find the usual kinetic energy $H = \mathbf{p}^2/2m + \ldots$.

The second Lagrangian that we met in Section 2.4.4 is more subtle. The action (2.47) is now a functional of the four-vector $X^\mu = (ct, \mathbf{x})$, and is given by the proper time experienced by the particle

$$S[X^\mu(\sigma)] = \int d\sigma\ L[\dot{X}] \quad \text{with} \quad L[\dot{X}] = -mc\sqrt{\dot{X}^\mu\dot{X}_\mu} \tag{4.26}$$

where indices are raised and lowered by the Minkowski metric $\eta_{\mu\nu} = \text{diag}(+1,-1,-1,-1)$, and $\dot{X}^\mu = dX^\mu/d\sigma$, with σ is an arbitrary parameter that labels the worldline. This action has manifest Lorentz invariance. But it also has reparameterisation invariance, meaning that we can choose to redefine the way we label the worldline, so $\sigma \to \tilde{\sigma}(\sigma)$, and the action remains unchanged. As we explained in Section 2.4.4, this reparameterisation invariance allows us to remove one of the four degrees of freedom in X^μ, bringing us back down to the three physical degrees of freedom of a particle moving in $\mathbf{R}^3$.

Let's see what becomes of (4.26) in the Hamiltonian formalism. The conjugate momentum is

$$P_\mu = \frac{\partial L}{\partial \dot{X}^\mu} = -mc\frac{1}{\sqrt{\dot{X}^\nu \dot{X}_\nu}}\dot{X}_\mu$$

The corresponding Hamiltonian is

$$H = P_\mu \dot{X}^\mu - L = 0$$

Wait. What? That's certainly unexpected: the Hamiltonian for this action with reparameterisation invariance vanishes!

What's going on here is that the Hamiltonian governs evolution with respect to the appropriate "time" coordinate which, in the present case, is the label σ on the worldline. But the label σ is a coordinate that has no independent meaning. And the way the formalism deal with this is by handing us a Hamiltonian $H = 0$ which is telling us that there is no evolution with respect to σ.

Instead, you have to look elsewhere for the dynamics in this Hamiltonian. And the right place is in the constraints. Recall from (2.50) that the momenta P_μ are not all independent: instead they obey the equation

$$P^\mu P_\mu = m^2 c^2 \tag{4.27}$$

which is the energy-momentum relation in special relativity.

It turns out that these kind of issues raise their heads whenever we have a theory with reparameterisation invariance and, to a lesser extent, when we have a theory with gauge symmetry. In particular, the Hamiltonian for General Relativity turns out to be $H = 0$. Again, the physics is hiding in the constraints. For Maxwell theory, the Hamiltonian is non-vanishing but there are also constraints that arise from gauge symmetry that complicate the issue.

We will not address these issues further in these lectures, but this is example is there simply as a warning shot. First, you shouldn't blindly equate the Hamiltonian with what you think is the energy of the system. In many examples this will be the case but not always, as illustrated clearly by the covariant Lagrangian (4.26) which has $H = 0$. Second, the Hamiltonian framework may come with certain constraints, like (4.27), that tell you that the momenta are not all independent. These can affect the dynamics in interesting ways.

4.1.6 What's Your Name, Man? William Rowan Hamilton (1805-1865)

The formalism described above arose out of Hamilton's interest in the theory of optics. The ideas were published in a series of books entitled "Theory of Systems of Rays", the first of which appeared while Hamilton was still an undergraduate at Trinity College, Dublin. They also contain the first application of the Hamilton-Jacobi formulation (which we shall see in Section 4.7) and the first general statement of the principal of least action, which sometimes goes by the name of "Hamilton's Principle".

Hamilton's genius was recognised early. His capacity to soak up classical languages and to find errors in famous works of mathematics impressed many. In an unprecedented move, he was offered a full professorship in Dublin while still an undergraduate. He also held the position of "Royal Astronomer of Ireland", allowing him to live at Dunsink Observatory even though he rarely did any observing. Unfortunately, the later years of Hamilton's life were not happy ones. The woman he loved married another and he spent much time depressed, mired in drink, bad poetry and quaternions[6].

4.2 Liouville's Theorem

Now it's time to use the Hamiltonian formalism to uncover some of the features of classical dynamics. In section we'll describe two such features, each of which follows from the description of the dynamics in terms of a flow in phase space, captured by Hamilton's equations (4.13).

Usually when considering classical mechanics, we think of our system at sitting in some point in phase space and then that point moving in time. But for our next result, we want to extend this viewpoint a little: we instead think of some region of phase space evolving in time. There are various reasons to do this, but one obvious one is because we want to admit to some uncertainty. Maybe we're not sure of the exact position or momentum of the system, but can confidently say that it sits in some small region of phase space. We then have the following important result:

Liouville's Theorem: As a region of phase space evolves in time, its shape will typically change, but its volume will not.

Proof: We start by considering the evolution of an infinitesimal volume moving for an infinitesimal time. We'll take this region of phase space to be centred around the point

[6]The "What's your name, man?" joke is not mine. I stole it from the outrageously talented Tim Blais, whose scholarly account of the life of Hamilton can be seen on YouTube.

(q^i, p_i) in phase space, with volume

$$V = dq_1 \dots dq_n dp_1 \dots dp_n \tag{4.28}$$

Then in time dt, we know that

$$q^i \to q^i + \dot{q}^i dt = q^i + \frac{\partial H}{\partial p_i} dt \equiv \tilde{q}^i \tag{4.29}$$

and

$$p_i \to p_i + \dot{p}_i dt = p_i - \frac{\partial H}{\partial q^i} dt \equiv \tilde{p}_i \tag{4.30}$$

So the new volume in phase space is

$$\tilde{V} = d\tilde{q}_1 \dots d\tilde{q}_n d\tilde{p}_1 \dots d\tilde{p}_n = (\det\{\mathcal{J}\})\, V \tag{4.31}$$

where $\det\{\mathcal{J}\}$ is the Jacobian of the transformation defined by the determinant of the $2n \times 2n$ matrix

$$\mathcal{J} = \begin{pmatrix} \partial\tilde{q}^i/\partial q^j & \partial\tilde{q}^i/\partial p_j \\ \partial\tilde{p}_i/\partial q^j & \partial\tilde{p}_i/\partial p_j \end{pmatrix} \tag{4.32}$$

To prove the theorem, we need to show that $\det\{\mathcal{J}\} = 1$. First consider a single degree of freedom (i.e. $n = 1$). Then we have

$$\det\{\mathcal{J}\} = \det \begin{pmatrix} 1 + (\partial^2 H/\partial p \partial q) dt & (\partial^2 H/\partial p^2)\, dt \\ -(\partial^2 H/\partial q^2)\, dt & 1 - (\partial^2 H/\partial q \partial p)\, dt \end{pmatrix} = 1 + \mathcal{O}(dt^2) \tag{4.33}$$

which means that

$$\frac{d(\det\{\mathcal{J}\})}{dt} = 0 \tag{4.34}$$

so that the volume remains constant for all time. Now to generalise this to arbitrary n, we have

$$\det\{\mathcal{J}\} = \det \begin{pmatrix} \delta_{ij} + (\partial^2 H/\partial p_i \partial q^j) dt & (\partial^2 H/\partial p_i \partial p_j)\, dt \\ -(\partial^2 H/\partial q^i \partial q^j)\, dt & \delta_{ij} - (\partial^2 H/\partial q^i \partial p_j)\, dt \end{pmatrix} \tag{4.35}$$

To compute the determinant, we need the result that $\det(1 + \epsilon M) = 1 + \epsilon\, \mathrm{Tr} M + \mathcal{O}(\epsilon^2)$ for any matrix M and small ϵ. Then we have

$$\det\{\mathcal{J}\} = 1 + \sum_i \left(\frac{\partial^2 H}{\partial p_i \partial q^i} - \frac{\partial^2 H}{\partial q^i \partial p_i} \right) dt + \mathcal{O}(dt^2) = 1 + \mathcal{O}(dt^2) \tag{4.36}$$

and we're done. □

Liouville's theorem is telling us that classical systems act, in many ways, like an incompressible fluid. But the fluid is incompressible in phase space, rather than in configuration space.

Notice that Liouville's theorem holds whether or not the system conserves energy. (i.e. whether or not $\partial H/\partial t = 0$). But the system must be described by a Hamiltonian. For example, systems with friction that exhibit dissipation do *not* obey Liouville's theorem. Instead, they tend to grind to a halt, heading to ever smaller regions of phase space with $\dot{q}_i = 0$.

Liouville's theorem states that the volume of a region of phase space stays constant under Hamiltonian evolution. But the theorem says nothing about the shape of that region. In some simple examples, the region might retain its rough shape. But there's nothing that says that this must be the case. Indeed, in systems that exhibit chaos, a given starting region will stretch and twist, keeping its original volume until but ultimately spreading tendril-like over a large region of phase space, rather like a droplet of ink spreading in water. This means that although Liouville's theorem tells us that the volume in phase space is strictly unchanged, in any practical terms our ignorance nonetheless increases.

The central idea of Liouville's theorem – that volume of phase space is constant – is somewhat reminiscent of quantum mechanics. Indeed, this is the first of several occasions where we shall see ideas of quantum physics creeping into the classical world. Suppose we have a system of particles distributed randomly within a square $\Delta q \Delta p$ in phase space. Liouville's theorem implies that if we evolve the system in any Hamiltonian manner, we can cut down the spread of positions of the particles only at the cost of increasing the spread of momentum. We're reminded strongly of Heisenberg's uncertainty relation, which is also written $\Delta q \Delta p =$ constant. While Liouville and Heisenberg seem to be talking the same language, there are very profound differences between them. The distribution in the classical picture reflects our ignorance of the system rather than any intrinsic uncertainty.

4.2.1 Liouville's Equation

To describe the evolution of a region of phase space more quantitatively, we introduce a density function $\rho(q, p, t)$. This could have one of two different interpretations

- If we have a single system, but don't know the exact state very well, then ρ can be interpreted as a probability distribution parameterising our ignorance. In this

case, the distribution should be normalised to

$$\int \prod_i dp_i dq^i \ \rho(q,p,t) = 1$$

- We may have a large number N of identical, non-interacting systems, for example $N = 10^{23}$ gas molecules in a jar. In this case, we really only care about the averaged behaviour. The distribution ρ now satisfies

$$\int \prod_i dq^i dp_i \ \rho(q,p,t) = N$$

In either case, we would like to know how the distribution $\rho(p,q,t)$ evolves in time. In both cases, the normalisations should hold for all time t. We know from Liouville's theorem that co-moving volume elements $dp\,dq$ are preserved, so the normalisations can only hold for all time if

$$\frac{d\rho}{dt} = \frac{\partial\rho}{\partial t} + \frac{\partial\rho}{\partial q^i}\,\dot{q}^i + \frac{\partial\rho}{\partial p_i}\,\dot{p}_i = 0$$

If we replace the expressions for $\dot{p}_i$ and $\dot{q}^i$ by Hamilton's equations (4.13) then, rearranging terms, we get

$$\frac{\partial\rho}{\partial t} = \frac{\partial\rho}{\partial p_i}\frac{\partial H}{\partial q^i} - \frac{\partial\rho}{\partial q^i}\frac{\partial H}{\partial p_i} \tag{4.37}$$

This is *Liouville's equation.*

Time Independent Distributions

Liouville's equation rests on the fact that the total time derivative of the distribution vanishes: $d\rho/dt = 0$. But often in physics we're interested in situations in which there is no explicit time dependence in the probability distribution, so $\partial\rho/\partial t = 0$. In this case, Liouville's equation (4.37) becomes

$$\frac{\partial\rho}{\partial p_i}\frac{\partial H}{\partial q^i} - \frac{\partial\rho}{\partial q^i}\frac{\partial H}{\partial p_i} = 0$$

An important class of solutions are probability distributions that are functions of the Hamiltonian,

$$\rho = \rho(H(q,p))$$

To see that these are indeed time independent, so that $\partial\rho/\partial t = 0$, we just need to check

$$\frac{\partial\rho}{\partial p_i}\frac{\partial H}{\partial q^i} - \frac{\partial\rho}{\partial q^i}\frac{\partial H}{\partial p_i} = \frac{\partial\rho}{\partial H}\frac{\partial H}{\partial p_i}\frac{\partial H}{\partial q^i} - \frac{\partial\rho}{\partial H}\frac{\partial H}{\partial q^i}\frac{\partial H}{\partial p_i} = 0 \tag{4.38}$$

A very famous example of this type is the *Boltzmann distribution*

$$\rho = \exp\left(-\frac{H(q,p)}{k_B T}\right) \tag{4.39}$$

which describes the probability distribution for a system at a temperature T. Here k_B is the Boltzmann constant. You can read (a lot) more about this in the lectures on Statistical Physics.

For example, for a free particle with $H = \mathbf{p}^2/2m$, the Boltzmann distribution is $\rho = \exp(-m\dot{\mathbf{x}}^2/2k_BT)$ which is a Gaussian distribution in velocities. Things get more interesting when you add some interaction terms to describe collisions between atoms in a gas. In fact, one can start from the Boltzmann distribution for 10^{23} atoms and derive the Navier-Stokes equation describing fluids. This, somewhat long story, is described in the lectures on Kinetic Theory.

There's an interesting historical story here. We can look back to the Hamiltonian of a free particle in a magnetic field (4.21), $H = (\mathbf{p} - e\mathbf{A})^2/2m$. But, as we noted back then, written in terms of the velocity, the Boltzmann distribution for a bunch of particles takes the form

$$\rho = \exp\left(-\frac{H(q,p)}{k_B T}\right) = \exp\left(-\frac{m\dot{\mathbf{x}}^2}{2k_B T}\right) \tag{4.40}$$

This is again a Gaussian distribution of velocities, independent of the magnetic field. This is a concern: the magnetism of solids is all about how the motion of electrons is affected by magnetic fields. Yet the magnetic field doesn't affect the velocity distribution of electrons. This is known as the Bohr-van Leeuwen paradox: there can be no magnetism in classical physics. This was one of the motivations for the development of quantum theory.

As an aside: there are actually two mechanisms for magnetism in solids. The first, known as Landau diamagnetism, is a quantum effect and is due to the motion of particles moving in magnetic fields. Roughly speaking, it's a rerun of the above argument, but with a different conclusion because things in the quantum world are different from the classical. The second, known as Pauli paramagnetism, arises because electrons carry a little internal degree of freedom known as spin. This too is, ultimately, a quantum effect. But there's a classical version of spin that we describe in Section 4.3.4.

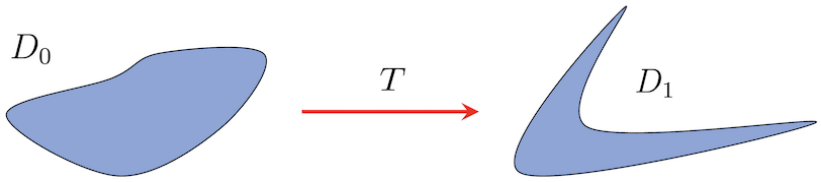


Figure 45: The Hamiltonian map in a time step T.

4.2.2 Poincaré Recurrence Theorem

Next, we turn to another aspect of motion in phase space. The following theorem applies to systems with *bounded* phase space, meaning that it has finite volume. At first glance, this seems unlikely all the phase spaces we've met so far have unbounded phase spaces. But all is not lost. If we have a conserved energy, $E = T + V$ with $T > 0$ and $V > 0$, then the accessible phase space is bounded by the spatial region $V(\mathbf{r}) \leq E$.

With this in mind, we have the following theorem, due to Poincaré around 1890. Roughly speaking, it says that wherever you start in phase space, if you wait long enough you will ultimately return to somewhere arbitrarily close to this point. More precisely:

Theorem: Consider an initial point P in phase space. Then for any neighbourhood D_0 of P, there exists a point $P' \in D_0$ that will return to D_0 in a finite time.

Proof: Consider the evolution of the region D_0 over a finite time interval T. Hamilton's equations provide a map $D_0 \mapsto D_1$ shown in figure 45. By Liouville's theorem, we know that $\text{Vol}(D_0) = \text{Vol}(D_1)$, although the shapes of these two regions will in general be different. Let D_k be the region after time kT, with k a positive integer. Then we first note that there there must exist integers k and k' such that the intersection of D_k and $D_{k'}$ is not empty:

$$D_k \cap D_{k'} \neq \phi$$

To see this, suppose that it's not true. Then the total volume $\bigcup_{k=0}^{\infty} D_k \to \infty$ but, by assumption, the accessible phase space volume is finite.

Take $k' > k$ such that the intersection region $\omega_{k,k'} = D_k \cap D_{k'}$ in not empty. But since the Hamiltonian mapping $D_k \to D_{k+1}$ is invertible, we can track backwards to find $\omega_{0,k'-k} = D_0 \cap D_{k'-k} \neq \phi$, as shown in Figure 46. This tells us that some point $P' \in D_0$ has returned to D_0 in $k' - k$ time steps T. □

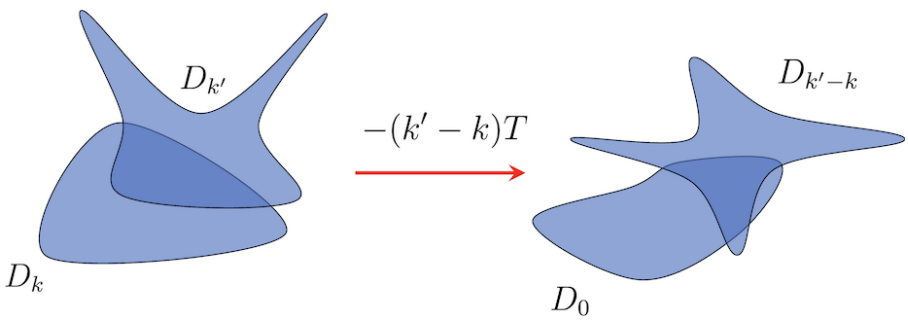


Figure 46: The inverse Hamiltonian map for $k'-k$ time steps proves the Poincaré recurrence theorem.

What does the Poincaré recurrence theorem mean? Consider gas molecules all in one corner of the room. If we let them go, they fill the room. But this theorem tells us that if we wait long enough, they will all return once more to the corner of the room.

That's a little disconcerting. One of the most powerful laws of physics is the second law of thermodynamics which says that entropy always increases. (This is described in detail in the lectures on Statistical Physics.) Yet this appears to be in stark contradiction to the Poincaré recurrence theorem which says that, if you wait long enough, The trick is that the Poincaré recurrence time for this to happen you'll get arbitrarily close to your starting point. If entropy did indeed increase, then it must later decrease.

We reconcile these contradictory statements only by considering the timescales involved. The Poincaré recurrence theorem says that there exists a time $(k'-k)T$ after which we return to our starting point, but it says nothing about how long we would expect to wait. The precise recurrence time depends on how "close" we want the system to return. (In fact, this questions isn't fully well defined in classical mechanics, because phase space doesn't come with a natural metric to allow us to measure distances in it.) But we can get some idea through a simple, crude argument.

To illustrate this, let's go back to a gas, with all the molecules initially confined to one half of a room. We let the gas go and it rapidly fills the room. At some fixed time later, there is a probability $1/2$ that a given gas molecule will be in the original half of the room. If we assume that the motion of the gas molecules are independent, this means that that the probability that all $N = 10^{23}$ gas molecules return to the original half of the room is

$$P = \left(\frac{1}{2}\right)^{10^{23}}$$

That's a very small number! Suppose, however, that a gas molecule takes some takes some characteristic time T to cross the room. (Say, $T = L/v$ with L the size of the room, and v the velocity of the molecule.) Then, one might expect that all molecules return to the original half in some recurrence time of order

$$T_{\text{recurrence}} \sim 2^{10^{23}}\, T$$

That's now a very very long time! Rather amusingly, it doesn't matter what the characteristic time scale T actually is. A reasonable estimate might be $T = 1$ second, but if you replace this with a Planck time $T \sim 10^{-44}$ seconds or a Hubble time, $T = 10$ billion years you still get the same ball park answer for the recurrence time because everything is swamped by that double exponential. (For example, $10^{23} - 43 \approx 10^{23}$.)

More generally, the recurrence time of a complicated system is of order

$$T_{\text{recurrence}} \sim e^S \times \text{some unit of time}$$

with S the entropy of the system. Again, the unit of time is completely irrelevant when S is exponentially large.

There is something of an irony to this. The second law of thermodynamics is one law of physics that we can be confident will never be overthrown. But this is precisely because we know that it's not an exact statement! The second law of thermodynamics is probabilistic in nature. There is a small chance that it's violated, but we know that the probability of this happening is so vanishingly small that it's a good operational definition of the word "never". Relatedly, we know that if we wait long enough then the second law will certainly be violated. But, again, the time that you need to wait it so unfathomably vast that, from a physics perspective, it's again as good as "never".

4.3 Poisson Brackets

Let's now continue with our study of the structure underlying classical mechanics. We start with a definition that looks slightly weird when you first see it.

Let $f(q,p)$ and $g(q,p)$ be two functions on phase space. Then the *Poisson bracket* is defined to be

$$\{f,g\} = \frac{\partial f}{\partial q^i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q^i} \tag{4.41}$$

The Poisson bracket obeys the following properties, all of which are straightforward to prove:

- $\{f, g\} = -\{g, f\}$.
- linearity: $\{\alpha f + \beta g, h\} = \alpha\{f, h\} + \beta\{g, h\}$ for all $\alpha, \beta \in \mathbf{R}$.
- Leibniz rule: $\{fg, h\} = f\{g, h\} + \{f, h\}g$. This follows from the chain rule in differentiation.
- Jacobi identity: $\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0$. To prove this you need a large piece of paper and a hot cup of coffee. Expand out all 24 terms and watch them cancel one by one.

We can also compute the Poisson bracket among the coordinates on phase space,

$$\{q^i, q^j\} = \{p_i, p_j\} = 0 \quad \text{and} \quad \{q^i, p_j\} = \delta^i_j \tag{4.42}$$

There might be something a little familiar about these results. The same four algebraic properties are obeyed by matrix commutators [,]. Moreover, the basic Poisson bracket relations (4.42) on phase space look very much like the commutation relations between position and momentum operators in quantum mechanics, with the Poisson bracket replaced by commutators. This, it turns out, is no coincidence, and the Poisson bracket provides one of the most direct connections to between quantum mechanics and classical mechanics. We"ll explore this further in Section 4.8.

(As an aside: the four algebraic properties are also obeyed by a derivative operator, known as the *Lie derivative*, that naturally arises in differential geometry.)

Another, more relevant aside: the Poisson bracket is an example of what mathematicians call a *symplectic structure* on phase space. Roughly speaking, given an even-dimensional manifold, a symplectic structure is an anti-symmetric pairing coordinates on the space. In physics language, this means that directions on the space are paired into position variables q and their conjugate momenta p, as captured by the relations (4.42).

There is, it turns out, a natural way to describe time evolution in terms of the Poisson bracket.

Claim: For any function $f(q, p, t)$ on phase space, evolution under Hamilton's equations is given by the following lovely result:

$$\frac{df}{dt} = \{f, H\} + \frac{\partial f}{\partial t} \tag{4.43}$$

Proof: We simply need to look at

$$\frac{df}{dt} = \frac{\partial f}{\partial p_i}\dot{p}_i + \frac{\partial f}{\partial q^i}\dot{q}^i + \frac{\partial f}{\partial t} = -\frac{\partial f}{\partial p_i}\frac{\partial H}{\partial q^i} + \frac{\partial f}{\partial q^i}\frac{\partial H}{\partial p_i} + \frac{\partial f}{\partial t} = \{f, H\} + \frac{\partial f}{\partial t}$$

where we've invoked Hamilton's equations in the second equality. □

If we restrict to the function $f = q^i$ or $f = p_i$, them the expression (4.43) simply gives us back Hamilton's equations

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

But (4.43) generalises this to any function $f(q, p)$. This general expression also has a counterpart in quantum mechanics, where an analogous equation describes the time evolution of operators in the so-called Heisenberg picture.

A simple consequence of (4.43) is that any function $Q(p, q)$ on phase space that obeys

$$\{Q, H\} = 0$$

then is a constant of motion. We say that Q and H *Poisson commute*. As an example of this, suppose that some coordinate q^i, so that it doesn't appear in the Hamiltonian. Then, from the definition of the Poisson bracket, we have

$$\dot{p}_i = \{p_i, H\} = 0 \tag{4.44}$$

which tells us that the conjugate momentum p_i is a conserved quantity. This, of course, simply confirms the result that we saw earlier in both the Lagrangian formalism (2.18) and the Hamiltonian formalism (4.15), now in the language of Poisson brackets.

Note that if Q_1 and Q_2 are constants of motion, so that $\{Q_1, H\} = \{Q_2, H\} = 0$, then the Jacobi identity tells us that $\{\{Q_1, Q_2\}, H\} = \{Q_1, \{Q_2, H\}\} + \{\{Q_1 H\}, Q_2\} = 0$ which means that $\{Q_1, Q_2\}$ is also a constant of motion. We say that the constants of motion form a closed algebra under the Poisson bracket.

We now illustrate the Poisson bracket structure with a number of examples.

4.3.1 Angular Momentum

The angular momentum of a particle is given by $\mathbf{L} = \mathbf{x} \times \mathbf{p}$. In components, this reads

$$L_1 = x^2 p_3 - x^3 p_2 \quad , \quad L_2 = x^3 p_1 - x^1 p_3 \quad , \quad L_3 = x^1 p_2 - x^2 p_1 \tag{4.45}$$

The Poisson bracket structure is

$$\begin{aligned}\{L_1, L_2\} &= \{x^2p_3 - x^3p_2, x^3p_1 - x^1p_3\} \\ &= \{x^2p_3, x^3p_1\} + \{x^3p_2, x^1p_3\} \\ &= -x^2p_1 + p_2x^1 \;=\; L_3\end{aligned}$$

So if L_1 and L_2 are conserved, we see that L_3 must also be conserved. By symmetry, this means that the whole vector $\mathbf{L}$ is conserved if any two components are. A similar calculation for other components gives

$$\{L_i, \L_j\} = \epsilon_{ijk}L_k \tag{4.46}$$

Similarly, one can show that

$$\{\mathbf{L}^2, L_i\} = 0 \quad i = 1, 2, 3$$

where $\mathbf{L}^2 = \sum_i L_i^2$. All of this is very reminiscent of the angular momentum commutation relations in quantum mechanics. We again see that what appears to be a novel quantum structure was actually lurking all along in classical mechanics.

In fact, there is some group theory that sits behind this structure. The Poisson bracket structure (4.46) is the structure that underlies the Lie group $SO(3)$ of rotations. Or, more precisely, it is the structure of the related Lie algebra $so(3)$. This is ultimately the reason why this same kind of structure arises in both classical mechanics and quantum mechanics.

4.3.2 A Particle in a Magnetic Field, Revisited

We saw the Hamiltonian for a particle of charge q moving in a magnetic field $\mathbf{B}$ in Section 4.1.4. It is described by

$$H = \frac{1}{2m}\left(\mathbf{p} - q\mathbf{A}\right)^2 = \frac{1}{2}m\dot{\mathbf{x}}^2$$

with the vector potential $\mathbf{A}$ related to the magnetic field by $\mathbf{B} = \nabla \times \mathbf{A}$. As emphasised in the expression above, when written in terms of the velocity $\dot{\mathbf{x}}$, the Hamiltonian takes the same form with and without a magnetic field. All the physics in this example is buried in the relation between the momentum $\mathbf{p}$ and the velocity $\dot{\mathbf{x}}$,

$$\mathbf{p} = m\dot{\mathbf{x}} + q\mathbf{A}$$

This means that, when written int terms of the velocity, the Hamiltonian is $H = \frac{1}{2}m\dot{\mathbf{x}}^2$ and the magnetic field disappears. But it's really hiding in the Poisson bracket structure. The canonical momentum has the Poisson brakcet

$$\{x^i, p_j\} = \delta^i_j \quad \text{and} \quad \{p_i, p_j\} = 0$$

But, as we've mentioned before, the canonical momentum $\mathbf{p}$ is not gauge invariant: it changes under a gauge transformation $\mathbf{A} \to \mathbf{A} + \nabla\chi$. This means that there's no sense in which you can assign an unambiguous number to the momentum because someone else, working with a different but physically equivalent choice of $\mathbf{A}$, will assign a different number. In contrast, the velocity $m\dot{\mathbf{x}}$ is gauge invariant and physical: everyone agrees on the value of the velocity. But its Poisson bracket structure is more complicated

$$\{x^i, \dot{x}^j\} = \frac{1}{m}\delta^{ij} \quad \text{and} \quad \{\dot{x}^i, \dot{x}^j\} = \frac{q}{m^2}\epsilon^{ijk}B_k \tag{4.47}$$

It's this Poisson bracket structure that is responsible for the interesting physics of a particle moving in a magnetic field.

A Curious Example: The Magnetic Monopole

It is an experimental fact that magnets have both a north and south pole. Cut a magnet in two, and each piece also has a north and a south pole.

Nonetheless, we can postulate the existence of a *magnetic monopole* which would be, say, just a north pole on its own. Such a *magnetic monopole* would emit a radial magnetic field,

$$\mathbf{B} = g\,\frac{\mathbf{r}}{r^3} \tag{4.48}$$

with g the "magnetic charge".

Magnetic monopoles have never been observed. Moreover, at first glance there is a law of physics that says that they can't exist. This law of physics follows immediately from the formulation of electromagnetism in terms of the vector potential $\mathbf{A}$. Since we define $\mathbf{B} = \nabla \times \mathbf{A}$, we immediately have one of Maxwell's equations,

$$\nabla \cdot \mathbf{B} = 0 \tag{4.49}$$

This states that any flux that enters a region must also leave. In particular, it prohibits field configurations of the form (4.48) which give rise to a delta-function source on the right-hand side of $\nabla \cdot \mathbf{B}$.

So if magnetic monopoles have never been observed, and, moreover, are forbidden by the laws of physics, then why are we interested in them?! The reason is that every theory that goes beyond Maxwell's equations and tries to unify electromagnetism with the other forces of Nature predicts magnetic monopoles. In fact, the existence of magnetic monopoles are one of the few robust predictions of every attempt to go beyond the

current laws of physics. Which means that there's reason to suspect that, somewhere in the universe, there may be particles with a radial magnetic field given by (4.48). Indeed, magnetic monopoles will be a constant companion as we learn theoretical physics, and will make an appearance in the lectures on Electromagnetism, Solid State Physics and Gauge Theory.

Here we ask a very simple question: what motion does an electron make when it moves in the presence of a magnetic monopole? It's tricky to set up the Lagrangian as we don't have a gauge potential $\mathbf{A}$. (Actually, one can work with certain singular gauge potentials but we won't go there). However, there is a more straightforward way to determine the motion using the Poisson brackets (4.47).

The trick is to look at the angular momentum. Or, more precisely, to look at the modified angular momentum

$$\mathbf{J} = m\mathbf{x} \times \dot{\mathbf{x}} - qg\hat{\mathbf{r}}$$

where $\hat{\mathbf{r}} = \mathbf{x}/|\mathbf{x}|$ is the unit, radial vector. When $g = 0$, this reduces to the usual angular momentum. But the usual angular momentum isn't conserved in the presence of a magnetic monopole. Instead, $\mathbf{J}$ is the conserved quantity. Indeed, it's simple to show, using the Poisson brackets (4.47), together with the Hamiltonian $H = \frac{1}{2}m\dot{\mathbf{x}}^2$, that

$$\{H, \mathbf{J}\} = 0$$

This ensures that $\mathbf{J}$ is a constant of motion. Note that we managed to do this calculation without ever introducing a gauge potential $\mathbf{A}$: the Poisson bracket structure for the velocity depends only on the magnetic field $\mathbf{B}$.

So what do we learn from this? Since $\mathbf{J}$ is conserved, we can look at

$$\hat{\mathbf{r}} \cdot \mathbf{J} = -qg$$

This tells us that the radial direction $\hat{\mathbf{r}}$ of the electron always sits at a constant angle relative to $\mathbf{J}$. In other words, the electron lies on a cone of angle $\cos\theta = qg/J$ pointing away from the vector $\mathbf{J}$. This is rather surprising behaviour: if we throw an electron at a magnetic monopole, it will spiral towards the monopole on the cone, reach some minimal distance, and then spiral back out again.

4.3.3 First Order Vortex Dynamics in the Plane

Sometimes, we might find ourselves given a set of equations of motion and wish to cast them in the framework of Hamiltonian dynamics. Here's an example.

Consider N vortices moving in the (x, y)-plane. The vortices have position $\mathbf{x}_i = (x_i, y_i)$ and each is assigned a *vorticity* γ_i, which you can think of as something akin to electric charge. In certain situations, the dynamics of these vortices is described by the equations of motion

$$\begin{aligned} \dot{x}_i &= -\sum_{j\neq i} \gamma_j \frac{y_i - y_j}{|\mathbf{x}_i - \mathbf{x}_j|^2} \\ \dot{y}_i &= +\sum_{j\neq i} \gamma_j \frac{x_i - x_j}{|\mathbf{x}_i - \mathbf{x}_j|^2} \end{aligned} \tag{4.50}$$

The interaction terms between vortices drops off as $1/r$, which is like the electrostatic force confined to a plane.

The novelty with these equations is that they are first order, rather than second order. It's like the particles live in the world envisaged by Aristotle, in which forces directly change velocities, rather than the world Galileo appreciated in which forces cause only acceleration. It means, in particular, that you only need to specify the initial positions $\mathbf{x}_i$ of each vortex to uniquely solve the equations of motion. In other words, the positions $\mathbf{x}_i$ specify the *state* of the system. This is in contrast to more familiar second-order equations of motion where you need the positions and momenta to specify the state of the system.

But, recall from earlier in this section that the state of the system parameterises what we called phase space. This means that the phase space of a first-order system is parameterised only by the positions which, for N particles, means $\mathcal{M} = \mathbf{R}^{2N}$.

This has consequence for the Hamiltonian formulation of these equations. In particular, the Poisson bracket is a structure on phase space. This means that the bracket $\{f, g\}$ acts on functions $f(\mathbf{x}_i)$ and $g(\mathbf{x}_i)$. For our vortex dynamics, it turns out that the relevant Poisson bracket structure is

$$\{f, g\} = \sum_{i=1}^{n} \frac{1}{\gamma_i} \left(\frac{\partial f}{\partial x_i} \frac{\partial g}{\partial y_i} - \frac{\partial f}{\partial y_i} \frac{\partial g}{\partial x_i} \right) \tag{4.51}$$

In particular, we have

$$\{x_i, y_j\} = \frac{\delta_{ij}}{\gamma_i} \tag{4.52}$$

This is a slightly strange structure: it means that one the positions, say y, is the canonical momentum for the other position, x. It arises whenever we are dealing with first-order, rather than second-order, equations of motion. that the particular Poisson bracket structure (4.52) is nice in the sense that it preserves both translational symmetry and rotational symmetry for each particle.

To specify the dynamics, we also need to write down the Hamiltonian on phase space. We take

$$H = -\sum_{i<j} \gamma_i \gamma_j \log |\mathbf{r}_i - \mathbf{r}_j| \tag{4.53}$$

Note that there is no kinetic energy term because there is no obvious momentum variable $\mathbf{p}$ in phase space. However, the Hamiltonian is not unfamiliar: it takes the same form as a potential between electric charges confined to a plane. We can now check that the Poisson bracket (4.51), together with the Hamiltonian (4.53), does indeed reproduce our equations of motion. We have

$$\dot{x}_i = \{x_i, H\} = \frac{1}{\gamma_i}\frac{\partial H}{\partial y_i} = -\sum_{j\neq i} \gamma_j \frac{y_i - y_j}{|\mathbf{x}_i - \mathbf{x}_j|^2}$$
$$\dot{y}_i = \{y_i, H\} = -\frac{1}{\gamma_i}\frac{\partial H}{\partial x_i} = \sum_{j\neq i} \gamma_j \frac{x_i - x_j}{|\mathbf{x}_i - \mathbf{x}_j|^2}$$

We can also briefly say a few words about the solutions to these equations. First, there are a few conserved quantities. There is one that follows from overall translational symmetry that we wish to call the total momentum although, in this context, it is more like a kind of centre of mass

$$P_x = \sum_i \gamma_i y_i \quad \text{and} \quad P_y = -\sum_i \gamma_i x_i$$

These satisfy $\{P_x, H\} = \{P_y, H\} = 0$, ensuring that they are conserved quantities. We also have $\{P_x, P_y\} = \sum_i \gamma_i$ and the right hand side, being constant, is trivially conserved.

Another conserved quantity follows from rotational invariance on the plane, so we would usually identify it as the total angular momentum. In this case, it takes the form

$$J = -\frac{1}{2}\sum_{i=1}^{n} \gamma_i (x_i^2 + y_i^2) \tag{4.54}$$

which again satisfies $\{J, H\} = 0$, ensuring it is conserved. The full algebra of the conserved quantities includes $\{P_x, J\} = -P_y$ and $\{P_y, J\} = P_x$, so the system closes (meaning we get back something we know on the right hand side). In fact, one can show that H, J and $(P_x^2 + P_y^2)$ provide three mutually Poisson commuting conserved quantities.

So what is the resulting motion of a bunch of vortices? A single vortex doesn't move: it's stuck in position. However, two vortices are induced to move by their mutual interaction. It's simple to solve their equations of motion to find that they move in a circle

$$\begin{aligned} x_1 - x_2 &= R\sin\left(\omega(t-t_0)\right) \\ y_1 - y_2 &= R\cos\left(\omega(t-t_0)\right) \end{aligned} \tag{4.55}$$

where R is the separation between the vortices and $\omega = (\gamma_1 + \gamma_2)/R^2$. So we learn that two vortices orbit each other with frequency inversely proportional to the square of their separation.

For three vortices, it turns out that there is a known solution which is possible because of the three mutually Poisson commuting conserved quantities we saw above. For four or more vortices, the motion turns out to be is chaotic.

The Poisson bracket structure (4.52) is rather odd. However, it also arises naturally in another familiar system: a particle moving in a uniform magnetic field $\mathbf{B} = (0, 0, B)$. We discussed this example previously in Section 2.4.3 and Section 4.1.4 where we saw that the canonical momentum gets modified to

$$\mathbf{p} = m\dot{\mathbf{x}} + q\mathbf{A}$$

If we work in the gauge $\mathbf{A} = (-By, 0, 0)$ then $p_x = m\dot{x} - qBy/m$. We might consider situations in which the magnetic field is strong, so the second term dominates. In this case

$$p_x \approx -\frac{qBy}{m} \quad \text{and} \quad \{x, p_x\} = 1 \quad \Rightarrow \quad \{x, y\} \approx -\frac{m}{qB} \tag{4.56}$$

This is the same Poisson bracket structure (4.52) that we saw for vortices. I should confess that the $\approx$ sign is doing quite a lot of heavy lifting in (4.56). It's not really obvious that we can ignore the $\dot{x}$ term in the momentum in the classical world. However, the same approximation is certainly valid in quantum mechanics where this can be thought of as the projection to the lowest Landau level and underlies, among other things, the quantum Hall effect.

4.3.4 Spin and a Compact Phase Space

Our next example is again a story about angular momentum. This, however, won't become apparent for a while. Instead, we're going to motivate this example with a rather different question.

So far, in all our examples of phase space it's been obvious what directions are "position" and what directions are "momentum". Indeed, we always started with some configuration space $\mathcal{C}$, describing the positions of the system, and then doubled the dimension to get the phase space. (For what it's worth, in fancy mathematical language the phase space in these familiar cases is called the cotangent bundle of the configuration space and denoted $\mathcal{M} = T^\star\mathcal{C}$.)

However, now that we've got a more abstract formulation of classical mechanics, we can spread our wings a little. We could take any even-dimensional space $\mathcal{M}$ (strictly a manifold) and think of it as the phase space of some system. To do this, we just need to equip it with a Poisson bracket structure. Then, adding a function over $\mathcal{M}$ that we identify as the Hamiltonian H will tell us how anything evolves in time, using (4.43).

Here we give the simplest example of this kind of approach. We will consider a phase space that is a two-dimensional sphere,

$$\mathcal{M} = \mathbf{S}^2$$

This certainly doesn't look like our previous phase spaces, not least because it's a compact space meaning that it doesn't extend to infinity in any direction. In contrast, for all our other phase spaces, the momentum direction was always non-compact. For example, for the pendulum example the phase space was a cylinder $\mathcal{M} = \mathbf{R} \times \mathbf{S}^1$, with the momentum parameterising the $\mathbf{R}$ direction.

We'll parameterise phase space with the usual spherical polar coordinates,

$$\theta \in [0, \pi] \quad \text{and} \quad \phi \in [0, 2\pi)$$

and consider the somewhat unusual action

$$S = \int dt \; J\dot{\phi}\cos\theta \tag{4.57}$$

with J a constant that, for reasons that we will see below, should be thought of as the radius of the sphere. This is different from our previous actions because it is only linear in velocities, rather than quadratic. This means that the momentum conjugate to ϕ is

$$p_\phi = \frac{\partial L}{\partial \dot{\phi}} = J\cos\theta$$

which is the other coordinate on the $\mathbf{S}^2$. This is why $\mathbf{S}^2$ is the phase space, rather than the configuration space. Meanwhile, there is no momentum conjugate to θ, since

$p_\theta = \partial L/\partial\dot{\theta} = 0$. The Poisson bracket structure is defined on functions $f(\theta,\phi)$ and $g(\theta,\phi)$ by

$$\{f,g\} = -\frac{1}{J\sin\theta}\left(\frac{\partial f}{\partial\phi}\frac{\partial g}{\partial\theta} - \frac{\partial f}{\partial\theta}\frac{\partial g}{\partial\phi}\right)$$

The peculiar $1/\sin\theta$ factor out front ensures that this gives

$$\{\phi, p_\phi\} = \{\phi, J\cos\theta\} = 1$$

which is the expected Poisson bracket between a coordinate and its conjugate momentum.

So far, it's not obvious that the action (4.57) preserves the $SO(3)$ rotational symmetry of the sphere. It turns out that it does. One way to see this is to construct the 3-vector from the origin to some point on the sphere. We will call this 3-vector $\mathbf{J}$ (preempting the fact advertised above that J is the radius of the sphere). The components of this three vector are

$$\begin{aligned} J_1 &= J\sin\theta\cos\phi \\ J_2 &= J\sin\theta\sin\phi \\ J_3 &= J\cos\theta \end{aligned} \tag{4.58}$$

We can compute the Poisson brackets between these. We have

$$\begin{aligned} \{J_1, J_2\} &= J^2\{\sin\theta\cos\phi, \sin\theta\sin\phi\} \\ &= -\frac{J}{\sin\theta}\left(-\sin\theta\cos\theta\sin^2\phi - \sin\theta\cos\theta\cos^2\phi\right) \\ &= J\cos\theta = J_3 \end{aligned}$$

Similar calculations show that we have the cyclic Poisson bracket

$$\{J_a, J_b\} = \epsilon_{abc}J_c \tag{4.59}$$

The symmetry of this Poisson bracket structure reflects the underlying $SO(3)$ symmetry of the action. Indeed, the Poisson bracket structure (4.59) is identical to the angular momentum Poisson bracket structure (4.46). This is not a coincidence. As we mentioned previously, the commutation relation (4.59) captures the essence of $SO(3)$ rotation group. (In more mathematical terms, the algebra is closely related to the Lie algebra $so(3)$.)

In fact, the vector $\mathbf{J}$ has a physical connection to angular momentum: it describes the *spin* of a particle. Note that we already used the word "spin" in Section 3 when describing rigid body motion. (It was the ω_3 component of the angular velocity.) But that's not the same thing as our spin $\mathbf{J}$, although both are related to angular momentum. Instead the spin $\mathbf{J}$ is something that is usually first introduced in quantum mechanics and is something rather unfamiliar in classical mechanics. For this reason, it's worth pausing to take a brief diversion into the quantum world.

Elementary particles such as electrons, or quarks, carry an internal angular momentum known as *spin*. In quantum mechanics, this manifests as two internal states of the particles, often denoted "spin up" $|\uparrow\rangle$ and "spin down" $|\downarrow\rangle$. In contrast to other systems, we tend not to introduce spin in quantum mechanics by starting with some classical set-up and then quantising it. Indeed, some authors will tell you that spin is an intrinsically quantum mechanical property with no classical analog. But that's not true. The classical analog is precisely the action (4.57). It turns out that it is precisely because we have a compact phase space $\mathcal{M} = \mathbf{S}^2$ that we get only a finite dimensional Hilbert space after quantisation. (For what it's worth, the classical magnitude of the spin J determines the number of states in the Hilbert space.)

Back to classical mechanics, we can now ask: what is the dynamics of the spin $\mathbf{J}$? If we compute the Hamiltonian that follows from the action (4.57), we find $H = \dot{\phi}p_\phi - L = 0$. There's nothing deep going on here. In particular, the vanishing Hamiltonian is not akin to the covariant relativistic particle that we saw in Section 4.1.5, where we found $H = 0$ because of reparameterisation invariance. Instead, the action (4.57) has $H = 0$ for the obvious reason: the spin $\mathbf{J}$ doesn't move. In fact, this is always the case whenever we have an action that is first order in time derivatives, rather than second order.

To initiate some motion, we need to add a potential term to the action. Physically, we do this by coupling the spin to a magnetic field $\mathbf{B}$. Mathematically, we do this by adding a term to the action (4.57)

$$S = \int dt\ J\dot{\phi}\cos\theta + \mu\mathbf{B}\cdot\hat{\mathbf{J}}$$

Here $\hat{\mathbf{J}} = \mathbf{J}/J$ should be thought of as a function of θ and ϕ μ, i.e. a function on phase space. Meanwhile μ is a parameter that governs the strength of the coupling to to the spin, known as the *magnetic moment*. This now gives the Hamiltonian

$$H = -\mu\mathbf{B}\cdot\hat{\mathbf{J}} = -\mu(B_1\sin\theta\cos\phi + B_2\sin\theta\sin\phi + B_3\cos\theta)$$

The motion of the spin $\mathbf{J}$ can be seen from (4.43),

$$\frac{dJ_a}{dt} = \{J_a, H\} = -\frac{\mu}{J}\{J_a, J_b\}B_b = -\frac{\mu}{J}\epsilon_{abc}B_b J_c \quad \Rightarrow \quad \frac{d\mathbf{J}}{dt} = -\mu\mathbf{B}\times\hat{\mathbf{J}}$$

This tells us that the spin $\mathbf{J}$ only moves if it is not parallel to $\mathbf{B}$. Suppose that $\mathbf{B} = (0, 0, B)$ then the equation of motion reads $J_3 = \text{constant}$, so $\dot{\theta} = \text{constant}$, while

$$\dot{J}_1 = \mu B\hat{J}_2 \quad \text{and} \quad \dot{J}_2 = -\mu B\hat{J}_1 \quad \Rightarrow \quad \phi = \frac{\mu B}{J}t$$

In other words, the spin precesses around the direction of $\mathbf{B}$.

4.4 Canonical Transformations

There is a way to write Hamilton's equations so that they look even more symmetric. Define the $2n$ vector $\mathbf{x} = (q_1, \ldots, q_n, p_1, \ldots, p_n)^T$ and the $2n \times 2n$ matrix J,

$$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{4.60}$$

where each entry is itself an $n\times n$ matrix. Then with this notation, Hamilton's equations read

$$\dot{\mathbf{x}} = J\,\frac{\partial H}{\partial \mathbf{x}} \tag{4.61}$$

Now remember that in the Lagrangian formalism we made a big deal about the fact that we could change coordinates $q^i \to q^i(q)$ without changing the form of the equations. Since we've managed to put q^i and p_i on an equal footing in the Hamiltonian formalism, one might wonder if its possible to make an even larger class of transformations of the form,

$$q^i \to q^i(q, p) \quad \text{and} \quad p_i \to P_i(q, p) \tag{4.62}$$

The answer is yes! But not all such transformations are allowed. To see what class of transformations leaves Hamilton's equations invariant, we use our new symmetric form in terms of $\mathbf{x}$ and write the transformation as

$$x_i \to y_i(x) \tag{4.63}$$

Note that we'll continue to use the index i which now runs over the range $i = 1, \ldots, 2n$. We have

$$\dot{y}_i = \frac{\partial y_i}{\partial x_j}\,\dot{x}_j = \frac{\partial y_i}{\partial x_j}\,J_{jk}\,\frac{\partial H}{\partial y_l}\,\frac{\partial y_l}{\partial x_k} \tag{4.64}$$

or, collating all the indices, we have

$$\dot{\mathbf{y}} = (\mathcal{J}\, J\, \mathcal{J}^T)\, \frac{\partial H}{\partial \mathbf{y}} \tag{4.65}$$

where $\mathcal{J}_{ij} = \partial y_i/\partial x_j$ is the Jacobian that we met in section 4.2. We see that Hamilton's equations are left invariant under any transformation whose Jacobian $\mathcal{J}$ satisfies

$$\mathcal{J}\, J\, \mathcal{J}^T = J \qquad \Rightarrow \qquad \frac{\partial y_i}{\partial x_j}\, J_{jk}\, \frac{\partial y_l}{\partial x_k} = J_{il} \tag{4.66}$$

The Jacobian $\mathcal{J}$ is said to be *symplectic* if this holds. A change of variables with a symplectic Jacobian is said to be a *canonical transformation*.

There is a nice method to construct canonical transformations using "generating functions" which we will mention in section 4.4.3. Before we get to this, let's look at some uses. We start by proving a theorem relating canonical transformations with Poisson brackets.

Theorem: The Poisson bracket is invariant under canonical transformations. Conversely, any transformation which preserves the Poisson bracket structure so that

$$\{q^i, q^j\} = \{P_i, P_j\} = 0 \quad \text{and} \quad \{q^i, P_j\} = \delta_{ij} \tag{4.67}$$

is canonical.

Proof: Let's start by showing that the Poisson bracket is invariant under canonical transformations. Consider two functions $f(x_i)$ and $g(x_i)$. Then,

$$\begin{aligned} \{f, g\} &= \frac{\partial f}{\partial q^i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q^i} \\ &= \frac{\partial f}{\partial x_i}\, J_{ij}\, \frac{\partial g}{\partial x_j} \end{aligned} \tag{4.68}$$

So if $x \to y(x)$, we have

$$\frac{\partial f}{\partial x_i} = \frac{\partial f}{\partial y_k}\, \mathcal{J}_{ki} \tag{4.69}$$

and, assuming the transformation is canonical, the Poisson bracket becomes

$$\{f, g\} = \frac{\partial f}{\partial y_k}\, \mathcal{J}_{ki}\, J_{ij}\, \mathcal{J}_{lj}\, \frac{\partial g}{\partial y_l} = \frac{\partial f}{\partial y_k}\, J_{kl}\, \frac{\partial g}{\partial y_l} \tag{4.70}$$

This means that we can compute our Poisson brackets in any coordinates related by a canonical transformation. Now let's show the converse. Go back to the notation (q^i, p_i) and the new coordinates $(q^i(q,p), P_i(q,p))$. The Jacobian is given by

$$\mathcal{J}_{ij} = \begin{pmatrix} \partial q^i/\partial q^j & \partial q^i/\partial p_j \\ \partial P_i/\partial q^j & \partial P_i/\partial p_j \end{pmatrix} \tag{4.71}$$

If we now compute $\mathcal{J}J\mathcal{J}^T$ in components, we get

$$(\mathcal{J}J\mathcal{J}^T)_{ij} = \begin{pmatrix} \{q^i, q^j\} & \{q^i, P_j\} \\ \{P_i, q^j\} & \{P_i, P_j\} \end{pmatrix} \tag{4.72}$$

So whenever the Poisson bracket structure is preserved, the transformation is canonical. □

Example

In the next section we'll see several non-trivial examples of canonical transformations which mix up q and p variables. But for now let's content ourselves with reproducing the coordinate changes that we had in section 2. Consider a change of coordinates of the form

$$q^i \to q^i(q) \tag{4.73}$$

We know that Lagrange's equations are invariant under this. But what transformation do we have to make on the momenta

$$p_i \to P_i(q,p) \tag{4.74}$$

so that Hamilton's equations are also invariant? We write $\Theta_{ij} = \partial q^i/\partial q^j$ and look at the Jacobian

$$\mathcal{J}_{ij} = \begin{pmatrix} \Theta_{ij} & 0 \\ \partial P_i/\partial q^j & \partial P_i/\partial p_j \end{pmatrix} \tag{4.75}$$

in order for the transformation to be canonical, we require $\mathcal{J}J\mathcal{J}^T = J$. By expanding these matrices out in components, we see that this is true if

$$P_i = (\Theta^{-1})_{ji} p_j \tag{4.76}$$

This is as we would expect, for it's equivalent to $P_i = \partial L/\partial \dot{q}^i$. Note that although $q^i = q^i(q)$ only, $P_i \neq P_i(p)$. Instead, the new momentum P_i depends on both q and p.

4.4.1 Infinitesimal Canonical Transformations

Consider transformations of the form

$$\begin{aligned} q^i \to q^i &= q^i + \alpha F_i(q,p) \\ p_i \to P_i &= p_i + \alpha E_i(q,p) \end{aligned} \tag{4.77}$$

where α is considered to be infinitesimally small. What functions $F_i(q,p)$ and $E_i(q,p)$ are allowed for this to be a canonical transformation? The Jacobian is

$$\mathcal{J}_{ij} = \begin{pmatrix} \delta_{ij} + \alpha\, \partial F_i/\partial q^j & \alpha\, \partial F_i/\partial p_j \\ \alpha\, \partial E_i/\partial q^j & \delta_{ij} + \alpha\, \partial E_i/\partial p_j \end{pmatrix} \tag{4.78}$$

so the requirement that $\mathcal{J} J \mathcal{J}^T = J$ gives us

$$\frac{\partial F_i}{\partial q^j} = -\frac{\partial E_i}{\partial p_j} \tag{4.79}$$

which is true if

$$F_i = \frac{\partial G}{\partial p_i} \quad \text{and} \quad E_i = -\frac{\partial G}{\partial q^i} \tag{4.80}$$

for some function $G(q,p)$. We say that G *generates* the transformation.

This discussion motivates a slightly different way of thinking about canonical transformations. Suppose that we have a one-parameter family of transformations,

$$q^i \to q^i(q,p;\alpha) \quad \text{and} \quad p_i \to P_i(q,p;\alpha) \tag{4.81}$$

which are canonical for all $\alpha \in \mathbf{R}$ and have the property that $q^i(q,p;\alpha=0) = q^i$ and $P_i(q,p;\alpha=0) = p_i$. Up until now, we've been thinking of canonical transformations in the "passive" sense, with the (q^i, P_i) labelling the same point in phase space as (q^i, p_i), just in different coordinates. But a one-parameter family of canonical transformations can be endowed with a different interpretation, namely that the transformations take us from one point in the phase space (q^i, p_i) to another point in the same phase space $(q^i(q,p;\alpha), P_i(q,p;\alpha))$. In this "active" interpretation, as we vary the parameter α we trace out lines in phase space. Using the results (4.77) and (4.80), the tangent vectors to these lines are given by,

$$\frac{dq^i}{d\alpha} = \frac{\partial G}{\partial p_i} \quad \text{and} \quad \frac{dp_i}{d\alpha} = -\frac{\partial G}{\partial q^i} \tag{4.82}$$

But these look just like Hamilton's equations, with the Hamiltonian replaced by the function G and time replaced by the parameter α. What we've found is that every

one-parameter family of canonical transformations can be thought of as "Hamiltonian flow" on phase space for an appropriately chosen "Hamiltonian" G. Conversely, time evolution can be thought of as a canonical transformation for the coordinates

$$(q^i(t_0), p_i(t_0)) \to (q^i(t), p_i(t)) \tag{4.83}$$

generated by the Hamiltonian. Once again, we see the link between time and the Hamiltonian.

As an example, consider the function $G = p_k$. Then the corresponding infinitesimal canonical transformation is $q^i \to q^i + \alpha\delta_{ik}$ and $p_i \to p_i$, which is simply a translation. We say that translations of q_k are generated by the conjugate momentum $G = p_k$.

4.4.2 Noether's Theorem Revisited

Recall that in the Lagrangian formalism, we saw a connection between symmetries and conservation laws. How does this work in the Hamiltonian formulation?

Consider an infinitesimal canonical transformation generated by G. Then

$$\begin{aligned}\delta H &= \frac{\partial H}{\partial q^i}\,\delta q^i + \frac{\partial H}{\partial p_i}\,\delta p_i \\ &= \alpha\,\frac{\partial H}{\partial q^i}\,\frac{\partial G}{\partial p_i} - \alpha\,\frac{\partial H}{\partial p_i}\,\frac{\partial G}{\partial q^i} + \mathcal{O}(\alpha^2) \\ &= \alpha\,\{H, G\}\end{aligned} \tag{4.84}$$

The generator G is called a *symmetry* of the Hamiltonian if $\delta H = 0$. This holds if

$$\{G, H\} = 0 \tag{4.85}$$

But we know from section 4.3 that $\dot{G} = \{G, H\}$. We have found that if G is a symmetry then G is conserved. Moreover, we can reverse the argument. If we have a conserved quantity G, then we can always use this to generate a canonical transformation which is a symmetry.

Explain that G is the Noether charge that generates the symmetry...

The Runge-Lenz Vector

Another interesting object is the (Hermann-Bernoulli-Laplace-Pauli-) Runge-Lenz vector, defined as

$$\mathbf{A} = \frac{1}{m}\,\mathbf{p} \times \mathbf{L} - \hat{\mathbf{r}} \tag{4.86}$$

where $\hat{\mathbf{r}} = \mathbf{r}/r$. This vector satisfies $\mathbf{A} \cdot \mathbf{L} = 0$. If you're willing to spend some time playing with indices, it's not hard to derive the following expressions for the Poisson bracket structure

$$\{L_a, A_b\} = \epsilon_{abc} A_c \qquad , \qquad \{A_a, A_b\} = -\frac{2}{m}\left(\frac{\mathbf{p}^2}{2m} - \frac{1}{r}\right)\epsilon_{abc} L_c \tag{4.87}$$

The last of these equations suggests something special might happen when we consider the familiar Hamiltonian $H = \mathbf{p}^2/2m - 1/r$ so that the Poisson bracket becomes

$$\{A_a, A_b\} = -\frac{2H}{m}\epsilon_{abc} L_c \tag{4.88}$$

Indeed, for this choice of Hamiltonian is a rather simple to show that

$$\{H, \mathbf{A}\} = 0 \tag{4.89}$$

So we learn that the Hamiltonian with $-1/r$ potential has another constant of motion $\mathbf{A}$ that we'd previously missed! The fact that $\mathbf{A}$ is conserved can be used to immediately derive Kepler's elliptical orbits: dotting $\mathbf{A}$ with $\hat{\mathbf{r}}$ yields $\hat{\mathbf{r}} \cdot \mathbf{A} + 1 = \mathbf{L}^2/r$ which is the equation for an ellipse. Note that the three constants of motion, $\mathbf{L}$, $\mathbf{A}$ and H form a closed algebra under the Poisson bracket.

Noether's theorem tells us that the conservation of $\mathbf{L}$ and H are related to rotational symmetry and time translation respectively. One might wonder whether there's a similar symmetry responsible for the conservation of $\mathbf{A}$. It turns out that there is: the Hamiltonian has a hidden $SO(4)$ symmetry group. You can read more about this in Goldstein.

4.4.3 Generating Functions

There's a simple method to construct canonical transformations between coordinates (q^i, p_i) and (q^i, P_i). Consider a function $F(q, Q)$ of the original q^i's and the final q^i's. Let

$$p_i = \frac{\partial F}{\partial q^i} \tag{4.90}$$

After inverting, this equation can be thought of as defining the new coordinate $q^i = q^i(q, p)$. But what is the new canonical momentum P? We'll show that it's given by

$$P_i = -\frac{\partial F}{\partial q^i} \tag{4.91}$$

The proof of this is a simple matter of playing with partial derivatives. Let's see how it works in an example with just a single degree of freedom. (It generalises trivially to the case of several degrees of freedom). We can look at the Poisson bracket

$$\{Q,P\} = \left.\frac{\partial Q}{\partial q}\right|_p \left.\frac{\partial P}{\partial p}\right|_q - \left.\frac{\partial Q}{\partial p}\right|_q \left.\frac{\partial P}{\partial q}\right|_p \tag{4.92}$$

At this point we need to do the playing with partial derivatives. Equation (4.91) defines $P = P(q,Q)$, so we have

$$\left.\frac{\partial P}{\partial p}\right|_q = \left.\frac{\partial Q}{\partial p}\right|_q \left.\frac{\partial P}{\partial Q}\right|_q \quad \text{and} \quad \left.\frac{\partial P}{\partial q}\right|_p = \left.\frac{\partial P}{\partial q}\right|_Q + \left.\frac{\partial Q}{\partial q}\right|_p \left.\frac{\partial P}{\partial Q}\right|_q \tag{4.93}$$

Inserting this into the Poisson bracket gives

$$\{Q,P\} = -\left.\frac{\partial Q}{\partial p}\right|_q \left.\frac{\partial P}{\partial q}\right|_Q = \left.\frac{\partial Q}{\partial p}\right|_q \frac{\partial^2 F}{\partial q \partial Q} = \left.\frac{\partial Q}{\partial p}\right|_q \left.\frac{\partial p}{\partial Q}\right|_q = 1 \tag{4.94}$$

as required. The function $F(q,Q)$ is known as a generating function of the first kind.

There are three further types of generating function, related to the first by Legendre transforms. Each is a function of one of the original coordinates and one of the new coordinates. You can check that the following expression all define canonical transformations:

$$\begin{aligned} F_2(q,P): &\qquad p_i = \frac{\partial F_2}{\partial q^i} \quad \text{and} \quad q^i = \frac{\partial F_2}{\partial P_i} \\ F_3(p,Q): &\qquad q^i = -\frac{\partial F_3}{\partial p_i} \quad \text{and} \quad P_i = -\frac{\partial F_3}{\partial q^i} \\ F_4(p,P): &\qquad q^i = -\frac{\partial F_4}{\partial p_i} \quad \text{and} \quad q^i = \frac{\partial F_4}{\partial P_i} \end{aligned} \tag{4.95}$$

4.5 Action-Angle Variables

We've all tried to solve problems in physics using the wrong coordinates and seen what a mess it can be. If you work in Cartesian coordinates when the problem really requires, say, spherical polar coordinates, it's always possible to get to the right answer with enough perseverance, but you're really making life hard for yourself. The ability to change coordinate systems can drastically simplify a problem. Now we have a much larger set of transformations at hand; we can mix up q's and p's. An obvious question is: Is this useful for anything?! In other words, is there a natural choice of variables which makes solving a given problem much easier. In many cases, there is. They're called "angle-action" variables.

4.5.1 The Simple Harmonic Oscillator

We'll start this section by doing a simple example which will illustrate the main point. We'll then move on to the more general theory. The example we choose is the simple harmonic oscillator. Notice that as our theory gets more abstract, our examples get easier!

We have the Hamiltonian

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 q^2 \tag{4.96}$$

so that Hamilton's equations are the familiar

$$\dot{p} = -m\omega^2 q \quad \text{and} \quad \dot{q} = \frac{p}{m} \tag{4.97}$$

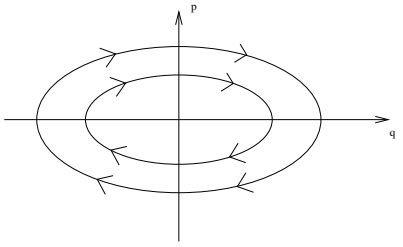


Figure 47:

which has the rather simple solution

$$q = A\cos(\omega(t - t_0)) \quad \text{and} \quad p = -m\omega A\sin(\omega(t - t_0)) \tag{4.98}$$

where A and t_0 are integration constants. The flows in phase space are ellipses as shown in the figure.

Now let's do a rather strange change of variables in which we use our freedom to mix up the position and momentum variables. We write

$$(q, p) \to (\theta, I) \tag{4.99}$$

where you can think of θ is our new position coordinate and I our new momentum coordinate. The transformation we choose is:

$$q = \sqrt{\frac{2I}{m\omega}}\,\sin\theta \quad \text{and} \quad p = \sqrt{2Im\omega}\,\cos\theta \tag{4.100}$$

It's an odd choice, but it has advantages! Before we turn to these, let's spend a minute checking that this is indeed a canonical transformation. There's two ways to do this and we'll do both:

1) We can make sure that the Poisson brackets are preserved. In fact, it's easier to work backwards and check that $\{q, p\} = 1$ in (θ, I) coordinates. In other words, we need to show that

$$\{q, p\}_{(\theta,I)} \equiv \frac{\partial q}{\partial \theta}\frac{\partial p}{\partial I} - \frac{\partial q}{\partial I}\frac{\partial p}{\partial \theta} = 1 \tag{4.101}$$

To confirm this, let's substitute the transformation (4.100),

$$\begin{aligned}\{q,p\}_{(\theta,I)} &= \left\{\sqrt{\frac{2I}{m\omega}}\,\sin\theta, \sqrt{2Im\omega}\cos\theta\right\}_{(\theta,I)} \\ &= 2\left\{\sqrt{I}\sin\theta, \sqrt{I}\cos\theta\right\}_{(\theta,I)} = 1\end{aligned} \tag{4.102}$$

where the final equality follows after a quick differentiation. So we see that the transformation (4.100) is indeed canonical.

2) The second way to see that the transformation is canonical is to prove that the Jacobian is symplectic. Let's now check it this way. We can calculate

$$\mathcal{J} = \begin{pmatrix} \partial\theta/\partial q & \partial\theta/\partial p \\ \partial I/\partial q & \partial I/\partial p \end{pmatrix} = \begin{pmatrix} (m\omega/p)\cos^2\theta & -(m\omega q/p^2)\cos^2\theta \\ m\omega q & p/m\omega \end{pmatrix} \tag{4.103}$$

from which we can calculate $\mathcal{J}J\mathcal{J}^T$ and find that it is equal to J as required.

So we have a canonical transformation in (4.100). But what's the point of doing this? Let's look at the Hamiltonian in our new variables.

$$H = \frac{1}{2m}(2m\omega I)\sin^2\theta + \frac{1}{2}\,m\omega^2\,\frac{2I}{m\omega}\cos^2\theta = \omega I \tag{4.104}$$

so the Hamiltonian doesn't depend on the variable θ! This means that Hamilton's equations read

$$\dot{\theta} = \frac{\partial H}{\partial I} = \omega \quad \text{and} \quad \dot{I} = -\frac{\partial H}{\partial \theta} = 0 \tag{4.105}$$

We've managed to map the phase space flow onto a cylinder parameterised by θ and I so that the flows are now all straight lines as shown in the figure. The coordinates (θ, I) are examples of *angle-action variables*.

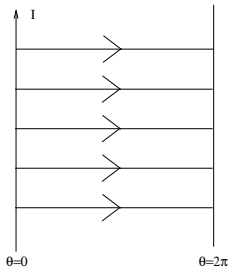


Figure 48:

4.5.2 Integrable Systems

In the above example, we saw that we could straighten out the flow lines of the simple harmonic oscillator with a change of variables, so that the motion in phase space became trivial. It's interesting to ask if we can we do this generally? The answer is: only for certain systems that are known as *integrable*.

Suppose we have n degrees of freedom. We would like to find canonical transformations

$$(q^i, p_i) \to (\theta_i, I_i) \tag{4.106}$$

such that the Hamiltonian becomes $H = H(I_1, \ldots, I_n)$ and doesn't depend on θ_i. If we can do this, then Hamilton's equations tell us that we have n conserved quantities I_i, while

$$\dot{\theta}_i = \frac{\partial H}{\partial I_i} = \omega_i \tag{4.107}$$

where ω_i is independent of θ (but in general depends on I) so that the solutions are simply $\theta_i = \omega_i t$. Whenever such a transformation exists, the system is said to be *integrable*. For bounded motion, the θ_i are usually scaled so that $0 \leq \theta_i < 2\pi$ and the coordinates (θ_i, I_i) are called *angle-action variables*.

Liouville's Theorem on Integrable Systems: There is a converse statement. If we can find n mutually Poisson commuting constants of motion $I_1, \ldots, I_n$ then this implies the existence of angle-action variables and the system is integrable. The requirement of Poisson commutation $\{I_i, I_j\} = 0$ is the statement that we can view the I_i as canonical momentum variables. This is known as Liouville's theorem. (Same Liouville, different theorem). A proof can be found in the book by Arnold.

Don't be fooled into thinking all systems are integrable. They are rather special and precious. It remains an active area of research to find and study these systems. But many – by far the majority – of systems are not integrable (chaotic systems notably among them) and don't admit this change of variables. Note that the question of whether angle-action variables exist is a *global* one. Locally you can always straighten out the flow lines; it's a question of whether you can tie these straight lines together globally without them getting tangled.

Clearly the motion of a completely integrable system is restricted to lie on $I_i =$ constant slices of the phase space. A theorem in topology says that these surfaces must be tori ($S^1 \times \ldots \times S^1$) known as the *invariant tori*.

4.5.3 Action-Angle Variables for 1d Systems

Let's see how this works for a 1d system with Hamiltonian

$$H = \frac{p^2}{2m} + V(q) \tag{4.108}$$

Since H itself is a constant of motion, with $H = E$ for some constant E throughout the motion, the system is integrable. We assume that the motion is bounded so that $q_1 \leq q \leq q_2$ as shown in the figure. Then the motion is periodic, oscillating back and forth between the two end points, and the motion in phase space looks something like the figure 50. Our goal is to find a canonical transformation to variables θ and I that straightens out this flow to look like the second figure in the diagram.

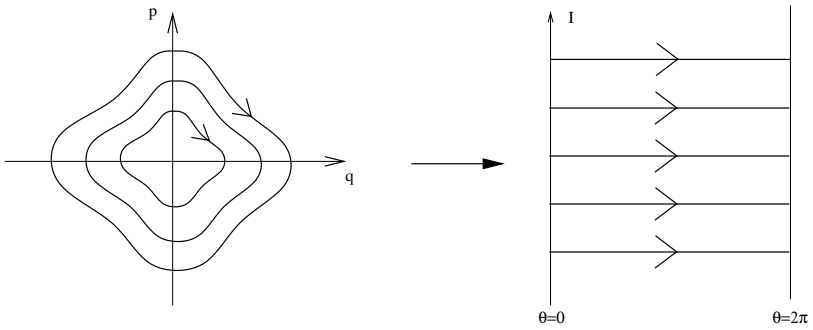


Figure 50: Can we straighten out the flow lines in phase space?

So what are I and θ? Since I is a constant of motion, it should be some function of the energy or, alternatively,

$$H = H(I) = E \tag{4.109}$$

But which choice will have as its canonical partner $\theta \in [0, 2\pi)$ satisfying

$$\dot{\theta} = \frac{\partial H}{\partial I} = \frac{\partial E}{\partial I} \equiv \omega \tag{4.110}$$

for a constant ω which is the frequency of the orbit?

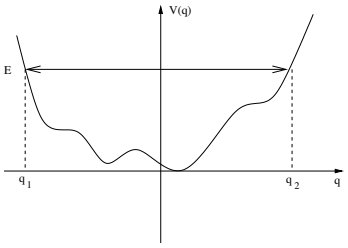


Figure 49:

Claim: The correct choice for I is

$$I = \frac{1}{2\pi} \oint p\, dq \tag{4.111}$$

which is the area of phase space enclosed by an orbit (divided by 2π) and is a function of the energy only.

Proof: Since the Hamiltonian is conserved, we may write the momentum as a function of q and E:

$$p = \sqrt{2m}\,\sqrt{E - V(q)} \tag{4.112}$$

We know that for this system $p = m\dot{q}$ so we have

$$dt = \sqrt{\frac{m}{2}} \frac{dq}{\sqrt{E - V(q)}} \tag{4.113}$$

Integrating over a single orbit with period $T = 2\pi/\omega$, we have

$$\begin{aligned} \frac{2\pi}{\omega} &= \sqrt{\frac{m}{2}} \oint \frac{dq}{\sqrt{E - V(q)}} \\ &= \oint \sqrt{2m} \left(\frac{d}{dE} \sqrt{E - V(q)} \right) dq \end{aligned} \tag{4.114}$$

At this point we take the differentiation d/dE outside the integral. This isn't obviously valid since the path around which the integral is evaluated itself changes with energy E. Shortly we'll show that this doesn't matter. For now, let's assume that this is valid and continue to find

$$\begin{aligned} \frac{2\pi}{\omega} &= \frac{d}{dE} \oint \sqrt{2m} \sqrt{E - V(q)}\, dq \\ &= \frac{d}{dE} \oint p\, dq \\ &= 2\pi \frac{dI}{dE} \end{aligned} \tag{4.115}$$

where in the last line, we've substituted for our putative action variable I. Examining our end result, we have found that I does indeed satisfy

$$\frac{dE}{dI} = \omega \tag{4.116}$$

where ω is the frequency of the orbit. This is our required result, but it remains to show that we didn't miss anything by taking d/dE outside the integral. Let's think about this. We want to see how the area enclosed by the curve changes under a small shift in energy δE. Both the curve itself and the end points $q_1 \leq q \leq q_2$ vary as the energy shifts. The latter change by $\delta q^i = (dV(q^i)/dq)^{-1}\, \delta E$. Allowing the differential d/dE to wander inside and outside the integral is tantamount to neglecting the change in the end points. The piece we've missed is the small white region in the figure. But these pieces are of order δE^2. To see this, note that order δE pieces are given by

$$\int_{q^i + \delta q^i}^{q^i} \sqrt{2m} \sqrt{E - V(q)}\, dq \approx \sqrt{2m} \sqrt{E - V(q)} \left(\frac{\partial V}{\partial q} \right)^{-1} \delta E \tag{4.117}$$

evaluated at the end point $q = q^i$. They vanish because $E = V(q^i)$ at the end points. This completes the proof. □

This tells us that we can calculate the period of the orbit ω by figuring out the area enclosed by the orbit in phase space as a function of the energy. Notice that we can do this without ever having to work out the angle variable θ (which is a complicated function of q and p) which travels with constant speed around the orbit (i.e. satisfies $\theta = \omega t$).

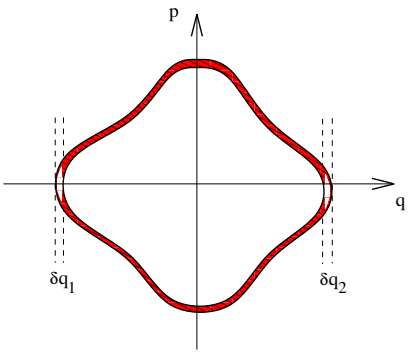


Figure 51:

In fact, it's not too hard to get an expression for θ by going over the above analysis for a small part of the period. It follows from the above proof that

$$t = \frac{d}{dE} \int p\, dq \tag{4.118}$$

but we want a θ which obeys $\theta = \omega t$. We see that we can achieve this by taking the choice

$$\theta = \omega \frac{d}{dE} \int p\, dq = \frac{dE}{dI} \frac{d}{dE} \int p\, dq = \frac{d}{dI} \int p\, dq \tag{4.119}$$

Because E is conserved, all 1d systems are integrable. What about higher dimensional systems? If they are integrable, then there exists a change to angle-action variables given by

$$\begin{aligned} I_i &= \frac{1}{2\pi} \oint_{\gamma_i} \sum_j p_j\, dq^j \\ \theta_i &= \frac{\partial}{\partial I_i} \int_{\gamma_i} \sum_j p_j\, dq^j \end{aligned} \tag{4.120}$$

where the γ_i are the periods of the invariant tori.

4.5.4 Action-Angle Variables for the Kepler Problem

Perhaps the simplest integrable system with more than one degree of freedom is the Kepler problem. This is a particle of mass m moving in three dimensions, subject to the potential

$$V(\mathbf{r}) = -\frac{k}{r}$$

We solved this already back in the Dynamics and Relativity course. Recall that we can use the conservation of the (direction of) angular momentum to restrict dynamics to a two-dimensional plane. We'll work in polar coordinates (r, ϕ) in this spatial plane. The associated momenta are $p_r = m\dot{r}$ and $p_\phi = mr^2\dot{\phi}$. The Hamiltonian is

$$H = \frac{1}{2m}p_r^2 + \frac{1}{2mr^2}p_\phi^2 - \frac{k}{r} \tag{4.121}$$

There are two action variables, one associated to the radial motion and one associated to the angular motion. The latter is straightforward: it is the angular momentum itself

$$I_\phi = \frac{1}{2\pi}\int_0^{2\pi} p_\phi d\phi = p_\phi$$

The action variable for the radial motion is more interesting. We can calculating it by using the fact that the total energy, E, and the angular momentum I_ϕ are both conserved. Then, rearranging (4.121), we have

$$p_r^2 = 2m\left(E + \frac{k}{r}\right) - \frac{I_\phi^2}{r^2}$$

and the action variable is

$$I_r = \frac{1}{2\pi}\oint p_r dr = \frac{1}{2\pi}\,2\int_{r_{\rm min}}^{r_{\rm max}} p_r dr = \frac{1}{2\pi}\,2\int_{r_{\rm min}}^{r_{\rm max}} \sqrt{2m\left(E + \frac{k}{r}\right) - \frac{I_\phi^2}{r^2}}dr$$

Here $r_{\rm min}$ and $r_{\rm max}$ are, respectively, the closest and furthest distance to the origin. (If you try to picture this in space, you'll need to recall that in the Kepler problem the origin sits on the focus of the ellipse, rather than the centre; this means that the smallest and furthest distance are opposite each other on the orbit).The factor of 2 in the second equality comes because a complete cycle goes from $r_{\rm min}$ to $r_{\rm max}$ and back again. To do this integral, you'll need the result

$$\int_{r_{\rm min}}^{r_{\rm max}} \sqrt{\left(1 - \frac{r_{\rm min}}{r}\right)\left(\frac{r_{\rm max}}{r} - 1\right)} = \frac{\pi}{2}\left(r_{\rm min} + r_{\rm max}\right) - \pi\sqrt{r_{\rm min}r_{\rm max}}$$

Using this, we find

$$I_r = \sqrt{\frac{m}{2|E|}}k - I_\phi$$

Or, re-arranging,

$$E = -\frac{mk^2}{2(I_r + I_\phi)^2} \tag{4.122}$$

There's something rather nice lurking in this result. The energy is the same as the Hamiltonian in this case and we can use it to compute the speed at which the angular variables change. This follows from Hamilton's equations,

$$\dot{\theta}_r = \frac{\partial H}{\partial I_r} \quad \text{and} \quad \dot{\theta}_\phi = \frac{\partial H}{\partial I_\phi}$$

Here $\theta_\phi = \phi$ while θ_r is some complicated function of r. But we see from (4.122) that the Hamiltonian is symmetric in I_r and I_ϕ. This means that the frequency at which the particle completes a ϕ cycle is the same frequency with which it completes a θ_r cycle. But that's the statement that the orbit is closed: when you go around 2π in space, you come back to the same r value. The existence of closed orbits is a unique feature of the $1/r$ potential. The calculation reveals the underlying reason for this.

4.6 Adiabatic Invariants

Consider a 1d system with a potential $V(q)$ that depends on some parameter λ. If the motion is bounded by the potential then it is necessarily periodic. We want to ask what happens if we *slowly* change λ over time. For example, we may slowly change the length of a pendulum, or the frequency of the harmonic oscillator.

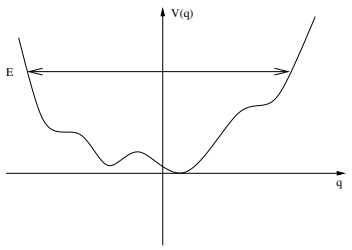


Figure 52:

Since we now have $\lambda = \lambda(t)$, the energy is not conserved. Rather $E = E(t)$ where

$$\dot{E} = \frac{\partial H}{\partial \lambda}\dot{\lambda} \tag{4.123}$$

But there are combinations of E and λ which remain (approximately) constant. These are called *adiabatic invariants* and the purpose of this section is to find them. In fact, we've already come across them: we'll see that the adiabatic invariants are the action variables of the previous section.

For the 1d system, the Hamiltonian is

$$H = \frac{p^2}{2m} + V(q;\lambda(t)) \tag{4.124}$$

and we claim that the adiabatic invariant is

$$I = \frac{1}{2\pi}\oint p\,dq \tag{4.125}$$

where the path in phase space over which we integrate now depends on time and is given by $p = \sqrt{2m}\sqrt{E(t) - V(q;\lambda(t))}$. The purpose of this section is to show that I is indeed an adiabatic invariant. At the same time we will also make clearer what we mean when we say that λ must change slowly.

Let's start by thinking of I as a function of the energy E and the parameter λ. As we vary either of these, I will change. We have,

$$\dot{I} = \left.\frac{\partial I}{\partial E}\right|_{\lambda}\dot{E} + \left.\frac{\partial I}{\partial \lambda}\right|_{E}\dot{\lambda} \tag{4.126}$$

where the subscripts on the partial derivatives tell us what variable we're keeping fixed. For an arbitrary variation of E and λ, this equation tells us that I also changes. But, of course, E and λ do not change arbitrarily: they are related by (4.123). The point of the adiabatic invariant is that when $\dot{E}$ and $\dot{\lambda}$ are related in this way, the two terms in (4.126) approximately cancel out. We can deal with each of these terms in turn. The first term is something we've seen previously in equation (4.116) which tells us that,

$$\left.\frac{\partial I}{\partial E}\right|_\lambda = \frac{1}{\omega(\lambda)} = \frac{T(\lambda)}{2\pi} \tag{4.127}$$

where $T(\lambda)$ is the period of the system evaluated at fixed λ. The second term in (4.126) tells us how the path changes as λ is varied. For example, two possible paths for two different λ's are shown in the figure and the change in I is the change in the area of under the two curves. We have

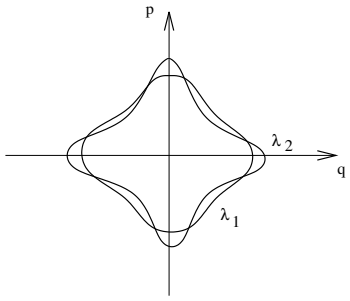


Figure 53:

$$\left.\frac{\partial I}{\partial \lambda}\right|_E = \frac{1}{2\pi}\left.\frac{\partial}{\partial \lambda}\right|_E \oint p dq = \frac{1}{2\pi}\oint \left.\frac{\partial p}{\partial \lambda}\right|_E dq = \frac{1}{2\pi}\int_0^{T(\lambda)} \left.\frac{\partial p}{\partial \lambda}\right|_E \left.\frac{\partial H}{\partial p}\right|_\lambda dt' \tag{4.128}$$

where, in the second equality, we have neglected a contribution arising from the fact that the path around which we integrate changes as λ changes. But this contribution can be safely ignored by the same argument given around (4.117).

We can get a simple expression for the product of partial derivatives by differentiating the Hamiltonian and remembering what depends on what. We have the expression $H(q,p,\lambda) = E$ where, in the left-hand side we substitute $p = \sqrt{2m}\sqrt{E(t) - V(q;\lambda(t))}$. Then differentiating with respect to λ, keeping E (and q) fixed, we have

$$\left.\frac{\partial H}{\partial \lambda}\right|_p + \left.\frac{\partial H}{\partial p}\right|_\lambda \left.\frac{\partial p}{\partial \lambda}\right|_E = 0 \tag{4.129}$$

So substituting this into (4.128) we have

$$\left.\frac{\partial I}{\partial \lambda}\right|_E = -\frac{1}{2\pi}\int_0^{T(\lambda)} \left.\frac{\partial H}{\partial \lambda}\right|_E dt' \tag{4.130}$$

So putting it all together, we have the time variation of I given by

$$\dot{I} = \left[T(\lambda)\left.\frac{\partial H}{\partial \lambda}\right|_E - \left(\int_0^{T(\lambda)} \left.\frac{\partial H}{\partial \lambda}\right|_E dt'\right)\right]\frac{\dot{\lambda}}{2\pi} \tag{4.131}$$

where, in the first term, we've replaced $\dot{E}$ with the expression (4.123). Now we're almost done. So far, each term on the right-hand side is evaluated at a given time t or, correspondingly for a given $\lambda(t)$. The two terms look similar, but they don't cancel! But we have yet to make use of the fact that the change in λ is *slow*. At this point we can clarify what we mean by this. The basic idea is that the speed at which the particle bounces backwards and forwards in the potential is much faster than the speed

at which λ changes. This means that the particle has performed many periods before it notices any appreciable change in the potential. This means that if we compute averaged quantities over a single period,

$$\langle A(\lambda)\rangle = \frac{1}{T}\int_0^T A(t,\lambda)\,dt \tag{4.132}$$

then inside the integral we may treat λ as if it is effectively constant. We now consider the time averaged motion $\langle \dot{I}\rangle$. Since λ can be taken to be constant over a single period, the two terms in (4.131) do now cancel. We have

$$\langle \dot{I}\rangle = 0 \tag{4.133}$$

This is the statement that I is an adiabatic invariant: for small changes in λ, the averaged value of I remains constant[7].

The adiabatic invariants played an important role in the early history of quantum mechanics. You might recognise the quantity I as the object which takes integer values according to the old 1915 Bohr-Sommerfeld quantisation condition

$$\frac{1}{2\pi}\oint p\,dq = n\hbar \qquad n\in\mathbf{Z} \tag{4.134}$$

The idea that adiabatic invariants and quantum mechanics are related actually predates the Bohr-Somerfeld quantisation rule. In the 1911 Solvay conference Einstein answered a question of Lorentz: if the energy is quantised as $E=\hbar n\omega$ where $n\in\mathbf{Z}$ then what happens if ω is changed slowly? Lorentz' worry was that integers cannot change slowly – only by integer amounts. Einstein's answer was not to worry: E/ω remains constant. These days the idea of adiabatic invariants in quantum theory enters into the discussion of quantum computers.

An Example: The Simple Harmonic Oscillator

We saw in section 4.5 that for the simple harmonic oscillator we have $I=E/\omega$. So if we change ω slowly, then the ratio E/ω remains constant. This was Einstein's 1911 point. In fact, for the SHO it turns out that there is an exact invariant that remains constant no matter how quickly you change ω and which, in the limit of slow change, goes over to I. This exact invariant is

$$J = \frac{1}{2}\left[\frac{q^2}{g(t)^2} + (g(t)\dot{q} - q\dot{g}(t))^2\right] \tag{4.135}$$

[7]The proof given above is intuitive, but begins to creak at the seams when pushed. A nice description of these issues, together with a more sophisticated proof using generating functions for canonical transformations is given in in the paper "The Adiabatic Invariance of the Action Variable in Classical Dynamics" by C.G.Wells and S.T.Siklos which can be found at http://arxiv.org/abs/physics/0610084.

where $g(t)$ is a function satisfying the differential equation

$$\ddot{g} + \omega^2(t)g - \frac{1}{g^3} = 0 \tag{4.136}$$

4.6.1 Adiabatic Invariants and Liouville's Theorem

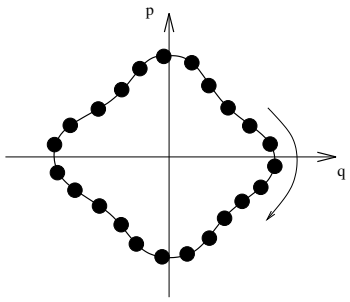


Figure 54:

There's a way to think of adiabatic invariants using Liouville's theorem. Consider first a series of systems, all described by a Hamiltonian with fixed parameter λ. We set off each system with the same energy E or, equivalently, the same action I, but we start them with slightly different phases θ. This means that their dynamics is described by a series of dots, all chasing each other around a fixed curve as shown in the figure. Now let's think about how this train of dots evolves under the Hamiltonian with time dependent $\lambda(t)$. Recall that Liouville's theorem states that the area of phase space is invariant under any Hamiltonian evolution. This holds whether or not $\partial H/\partial t = 0$, so is still valid for the time dependent Hamiltonian with $\lambda(t)$. One might be tempted to say that we're done since all the words sound right: Liouville's theorem implies that the area is conserved which is also the statement that our adiabatic invariant I doesn't change with time. But this is a little too fast! Liouville's theorem says the area of a distribution of particles in phase space is conserved, not the area enclosed by a perimeter ring of particles. Indeed, Liouville's theorem holds for any variation $\lambda(t)$, not just for adiabatic variations. For a fast change of $\lambda(t)$, there is nothing to ensure that the particles that had the same initial energy, but different phases, would have the same final energy and we lose the interpretation of a ring of dots in phase space enclosing some area.

The missing ingredient is the "*adiabatic principle*". In this context it states that for a suitably slow change of the parameter λ, all the systems in the same orbit, with the same energy, are affected in the same manner. If this holds, after some time the dots in phase space will still be chasing each other around another curve of constant energy E'. We can now think of a distribution of particles filling the area I inside the curve. As λ varies slowly, the area doesn't change *and* the outer particles remain the outer particles, all with the same energy. Under these circumstances, Liouville's theorem implies the adiabatic invariant I is constant in time.

4.6.2 An Application: A Particle in a Magnetic Field

We saw in Section 4.1 that a particle in a constant magnetic field $\mathbf{B} = (0, 0, B)$ makes circles with Larmor frequency $\omega = eB/mc$ and a radius R, which depends on the energy

of the particle. But what happens if B is slowly varying over space? i.e. $B = B(x, y)$, but with

$$\partial_i B \ll R \tag{4.137}$$

so that the field is roughly constant over one orbit.

In this example, there is no explicit time dependence of the Hamiltonian so we know that the Hamiltonian itself is an exact constant of motion. For a particle in a constant magnetic field we can calculate H of an orbit by substituting the solutions (**??**) into the Hamiltonian (**??**). We find

$$H = \frac{1}{2}m\omega^2 R^2 = \frac{e^2 R^2 B^2}{2mc^2} \tag{4.138}$$

This quantity is conserved. But what happens to the particle? Does it drift to regions with larger magnetic field B, keeping H constant by reducing the radius of the orbit? Or to regions with smaller B with larger orbits?

We can answer this by means of the adiabatic invariant. We can use this because the motion of the particle is periodic in space so the particle sees a magnetic field which varies slowly over time. The adiabatic invariant is

$$I = \frac{1}{2\pi}\oint p\, dq \tag{4.139}$$

which is now to be thought of as a line integral along the orbit of the electron. We evaluate this on the solution for the uniform magnetic field (**??**)

$$\begin{aligned} I &= \frac{1}{2\pi}\int_0^T (p_x\dot{x} + p_y\dot{y})\, dt \\ &= \frac{1}{2\pi}\int_0^T \left(bR\omega\, \cos\omega t + m\omega^2 R^2 \sin^2\omega t\right)\, dt \\ &= \frac{m\omega R^2}{2\pi}\int_0^{2\pi} \sin^2\theta\, d\theta \end{aligned} \tag{4.140}$$

Setting $\omega = eB/mc$, we see that the adiabatic invariant I is proportional to $(e/c)BR^2$. Since the electric charge e and the speed of light c are not things we can change, we find that BR^2 is constant over many orbits. But as $H \sim B^2R^2$ is also conserved, both the magnetic field B seen by the particle and the radius of the orbit R must be individually conserved. This means the particle can't move into regions of higher or lower magnetic fields: it must move along constant field lines[8].

[8]For results that go beyond the adiabatic approximations, see the paper by the man: E. Witten "*A Slowly Moving Particle in a Two-Dimensional Magnetic Field*", Annals of Physics **120** 72 (1979).

Finally, there's a simple physical way to see that the particle indeed drifts along lines of constant magnetic field. As the particle makes its little Larmor circles, it feels a slightly stronger force when it's, say, at the top of its orbit where the field is slightly larger, compared to when its at the bottom. This net force tends to push the particle to regions of weaker or stronger magnetic field. But we've seen through the use of adiabatic invariants that this isn't possible. The slow drift of the particle acts such that it compensates for this small force, keeping the particle on constant field lines.

There's a slight variant of the above set-up which allows you to trap charged particles using magnetic fields. Consider the particle making its little Larmor circles in the (x, y) plane, but also moving in the z direction and take a magnetic field that's constant in the (x, y)-plane, but ever increasing in the z-direction. The energy of the particle is given by,

$$H = \frac{1}{2}m\dot{z}^2 + \frac{e^2R^2B^2}{2mc^2} = \frac{1}{2}m\dot{z}^2 + \frac{IemB}{mc} \tag{4.141}$$

Both $H > 0$ and $I > 0$ are constant in time. This ensures that there exists a value of the magnetic field $B > 0$ at which we necessarily have $\dot{z} = 0$ and the particle can go no further. At this stage it turns round and goes back again. By creating a magnetic field that increases at two ends, charged particles can be made to bounce back and forth in the z direction, while executing their little circles in the (x, y)-plane. It is this mechanism that traps charged particles in magnetic loops emitted from the sun and is ultimately responsible for solar flares.

4.6.3 Hannay's Angle

Consider a particle bouncing around, with periodic motion, in an arbitrary potential. There are many parameters λ_a describing the shape of the potential. As we slowly vary the λ_a the path in phase space changes although, as we have seen, the area enclosed by the path remains the same. After some time $t_{\rm long}$ (which, by the assumption of adiabiticity, is much longer than the period T of a single orbit) we return to the original parameters so that $\lambda_a(t_{\rm long}) = \lambda_a(0)$. The question we want to ask is: how has the phase angle θ changed?

For any fixed λ_a, the velocity of the angle variable is $\dot{\theta} = \partial H/\partial I = \omega(I, \lambda_a)$. As we slowly vary the parameters, the particle is spinning around its phase space orbits. When we return we therefore expect that the phase has been shifted by $\int \omega dt$. Which is true. But it turns out that there is another, more subtle, contribution to the phase shift as well. We'll now see where this comes from.

As the parameters change, we can write the change in the angle θ as

$$\dot{\theta} = \frac{\partial H}{\partial I} + \frac{\partial \theta}{\partial \lambda_a}\,\dot{\lambda}_a \tag{4.142}$$

Which looks simple enough. But there's a problem. The second term is not well defined. For each set of parameters λ_a we have different action angle variables $I(\lambda_a)$ and $\theta(\lambda_a)$. But there's nothing that stops choosing a different origin $\theta = 0$ for each choice of the parameters. In other words, we could always redefine

$$\theta(\lambda_a) \to \theta(\lambda_a) + \beta(\lambda_a) \tag{4.143}$$

where we shift by a different constant β for each λ_a. What this means is that it doesn't really make any sense to compare the angle variable for different parameters λ_a. This makes the second term — which tells us how the angle variable changes as we change the parameters — ambiguous. One might think this means that we can just ignore it. Or, more precisely, we could choose the shifts β so that the angle variables are defined in such a way that the second term vanishes. But it turns out that this isn't possible. Let's see why. The point is that it does make sense to compare the angle variable for the same parameters λ_a. After such a time $t_{\rm long}$, we have

$$\theta(t_{\rm long}) = \theta(0) + \int_0^{t_{\rm long}} \omega\, dt + \Delta\theta \tag{4.144}$$

The term $\int \omega dt$ is the dynamic term that we anticipated above, arising from the fact that θ is continually making orbits around the curve in phase space. It depends on the time $t_{\rm long}$ that we took to make the change. The other term that we call $\Delta\theta$ is more interesting. From (4.142) it is given by

$$\Delta\theta = \int_0^{t_{\rm long}} \left\langle \frac{\partial \theta}{\partial \lambda_a} \right\rangle \dot{\lambda}_i\, dt = \oint_C \left\langle \frac{\partial \theta}{\partial \lambda_a} \right\rangle d\lambda_a \tag{4.145}$$

where we've used the fact that the change in λ_a is adiabatic to replace the integrand with its average over one period of the orbit. The final expression above means an integration over the curve C that the system traces in parameter space. We see that $\Delta\theta$ is independent of the time $t_{\rm long}$ taken to make the change. However, it does depend on the path that we took through the space of all possible potentials. It is known as the "Hannay angle". Note that the Hannay angle is invariant under the ambiguity (4.144) even though the quantity $\partial\theta/\partial\lambda_a$ that appears in the integrand isn't. This idea of integrating quantities around loops is an example of "holonomy", an important concept in modern physics.

Rather surprisingly, the Hannay angle was first discovered only in 1984. The history is interesting. First Berry discovered a similar phase for the wavefunction in quantum mechanics (now known as the Berry phase). Many physicists were shocked that such a simple and profound idea as Berry's phase had lain hidden in the formulation of quantum mechanics for 50 years and it set off a flurry of theoretical and experimental research. Soon after this, Hannay showed that an analogous phase had lain undiscovered in classical mechanics for 150 years! Although, in certain examples in celestial mechanics, the phase $\Delta\theta$ had been correctly calculated, the general theory lying behind it had not been appreciated. We now describe this theory.

The first step is to use a higher dimensional version of Stokes' theorem to express the contour integral (4.145) as a surface integral

$$\Delta\theta = \int_S \left(\frac{\partial}{\partial\lambda_a} \left\langle \frac{\partial\theta}{\partial\lambda_b} \right\rangle - \frac{\partial}{\partial\lambda_b} \left\langle \frac{\partial\theta}{\partial\lambda_a} \right\rangle \right) dA_{ab} \tag{4.146}$$

where S is a surface in parameter space bounded by the curve C and dA_{ab} is the infinitesimal surface element.

Claim: The Hannay angle can be written as

$$\Delta\theta = \frac{d}{dI} \int_S W_{ab}\, dA_{ab} \tag{4.147}$$

where the anti-symmetric matrix W_{ab} (known mathematically as a 2-form) is given by

$$W_{ab} = \left\langle \frac{\partial\theta}{\partial\lambda_a}\frac{\partial I}{\partial\lambda_b} - \frac{\partial\theta}{\partial\lambda_b}\frac{\partial I}{\partial\lambda_a} \right\rangle \tag{4.148}$$

Proof: To start with let's think about the averaging procedure a little more. In equation (4.132) we wrote $\langle A\rangle$ as a time average, integrating over a single period. We could equally as well write it as an angle average,

$$\langle A\rangle = \oint A(I,\theta)\, d\theta = \int A(q',p')\, \delta(I'-I) \frac{dq'dp'}{2\pi} \tag{4.149}$$

where in the second equality we integrate over all of phase space and insert the delta function $\delta(I'-I)$ to restrict the integral to the orbit over the curve with action I. It's

this latter formula that we'll use. This allows us to write,

$$\begin{aligned}
\frac{\partial}{\partial\lambda_a}\left\langle\frac{\partial\theta}{\partial\lambda_b}\right\rangle - \frac{\partial}{\partial\lambda_b}\left\langle\frac{\partial\theta}{\partial\lambda_a}\right\rangle &= \left[\frac{\partial}{\partial\lambda_a}\int\frac{\partial\theta}{\partial\lambda_b} - \frac{\partial}{\partial\lambda_b}\int\frac{\partial\theta}{\partial\lambda_a}\right]\,\delta(I'-I)\,\frac{dq'dp'}{2\pi} \\
&= \int\left[\frac{\partial\theta}{\partial\lambda_b}\frac{\partial\delta}{\partial\lambda_a} - \frac{\partial\theta}{\partial\lambda_a}\frac{\partial\delta}{\partial\lambda_b}\right]\,\frac{dq'dp'}{2\pi} \\
&= \int\left[\frac{\partial\theta}{\partial\lambda_b}\frac{\partial I'}{\partial\lambda_a} - \frac{\partial\theta}{\partial\lambda_a}\frac{\partial I'}{\partial\lambda_b}\right]\,\frac{\partial\delta}{\partial I'}\,\frac{dq'dp'}{2\pi} \\
&= -\frac{d}{dI}\int\left[\frac{\partial\theta}{\partial\lambda_b}\frac{\partial I'}{\partial\lambda_a} - \frac{\partial\theta}{\partial\lambda_a}\frac{\partial I'}{\partial\lambda_b}\right]\,\delta(I'-I)\,\frac{dq'dp'}{2\pi} \\
&= \frac{d}{dI}\,W_{ab} \qquad (4.150)
\end{aligned}$$

which completes the proof. I haven't included any examples here of the Hannay angle: some simple ones can be found in the original literature[9] and more advanced applications can be found by googling "Hannay angles". For the pendulum example, in which the length of the pendulum varies, the Hannay angle vanishes. This is because there is only one parameter to vary, while a non-trivial $\Delta\theta$ occurs only if we make a non-trivial loop C in parameter space.

4.7 The Hamilton-Jacobi Equation

In this section we will describe yet another viewpoint on classical dynamics, known as Hamilton-Jacobi theory. It will tie together several concepts that we've met so far. Recall from section 2.1 the principle of least action. We define the action

$$S = \int_0^T L(q^i, \dot{q}^i, t)\,dt \qquad (4.151)$$

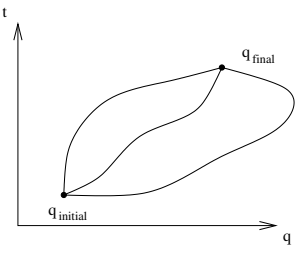


Figure 55:

which we evaluate for *all* paths $q(t)$ with fixed end points

$$q^i(0) = q^i_{\rm initial} \quad , \quad q^i(T) = q^i_{\rm final} \qquad (4.152)$$

Then the true path taken is an extremum of the action: $\delta S = 0$.

Now let's change perspective a little. Consider the action evaluated only along the *true* path $q^i_{\rm classical}(t)$ and define

$$W(q^i_{\rm initial}, q^i_{\rm final}, T) = S[q^i_{\rm classical}(t)] \qquad (4.153)$$

While S is a functional on any path, W is to be considered as a function of the initial and final configurations $q^i_{\rm initial}$ and $q^i_{\rm final}$ as well as the time T it takes to get between them.

[9] J. Hannay, "*Angle Variable Holonomy in Adiabatic Excursion of an Integrable Hamiltonian*" J. Phys A, **18** 221 (1985); M. Berry, "*Classical Adiabatic Angles and Quantal Adiabatic Phase*" J. Phys A, **18** 15 (1985).

Now let's ask what happens if we keep $q^i_{\rm initial}$ fixed but vary the end point $q^i_{\rm final}$. We can go back to the analysis of section 2.1 to see that when the action is varied it looks like

$$\delta S = \int_0^T dt \left[\frac{\partial L}{\partial q^i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right)\right] \delta q^i(t) + \left[\frac{\partial L}{\partial \dot{q}^i}\,\delta q^i(t)\right]_0^T \tag{4.154}$$

If we evaluate this on the classical path the first term vanishes. We're left with

$$\frac{\partial W}{\partial q^i_{\rm final}} = \left.\frac{\partial L}{\partial \dot{q}^i}\right|_{t=T} = p_i^{\rm final} \tag{4.155}$$

The next thing we want to compute is $\partial W/\partial T$. Let's start by considering a classical path with fixed initial configuration $q^i_{\rm initial}$. We'll let the path run on a little longer than before, so $T \to T + \delta T$. Then we have

$$\frac{dW}{dT} = \frac{\partial W}{\partial T} + \frac{\partial W}{\partial q^i_{\rm final}}\,\dot{q}^i_{\rm final} = \frac{\partial W}{\partial T} + p_i^{\rm final}\dot{q}^i_{\rm final} \tag{4.156}$$

But this total derivative is easily calculated since $dS/dT = L$, or

$$\frac{dW}{dT} = L(q^i_{\rm classical}(T), \dot{q}^i_{\rm classical}(T), T) = L(q^i_{\rm final}, \dot{q}^i_{\rm final}, T) \tag{4.157}$$

So we arrive at the equation,

$$\begin{aligned}\frac{\partial W}{\partial T} &= -\left(p_i^{\rm final}\dot{q}^i_{\rm final} - L(q^i_{\rm final}, \dot{q}^i_{\rm final}, T)\right) \\ &= -H(q^i_{\rm final}, p_i^{\rm final}, T)\end{aligned} \tag{4.158}$$

At this stage, the only time in the game is T and the only position in the game is $q^i_{\rm final}$. So we can simply drop the word "final", and relabel $T \to t$. We have found ourselves a time dependent function on configuration space $W = W(q^i, t)$ which satisfies

$$\frac{\partial W}{\partial q^i} = p_i \qquad \text{and} \qquad \frac{\partial W}{\partial t} = -H(q^i, p_i, t) \tag{4.159}$$

or, substituting the first into the second, we have

$$\frac{\partial W}{\partial t} = -H(q^i, \partial W/\partial q^i, t) \tag{4.160}$$

This is the *Hamilton-Jacobi Equation*.

We've shown how a solution to the Hamilton-Jacobi equation can be constructed by looking at the classical action of paths which reach a point q^i at time T starting from some initial reference point $q^i_{\rm initial}$. The starting point $q^i_{\rm initial}$ can be considered integration constants. In fact, there are more general solutions to the Hamilton-Jacobi equation, although all are related to the classical action in a similar way.

Suppose we find a solution to (4.160). What do we do with it? We're now armed with some time-dependent function $W(q^i, t)$ on configuration space. We combine this with the first of Hamilton's equations which reads

$$\dot{q}^i = \left.\frac{\partial H}{\partial p_i}\right|_{p_i = \partial W/\partial q^i} \tag{4.161}$$

where, on the right-hand-side, we've replaced every appearance of the momenta p_i by a function of the coordinates using $p_i = \partial W/\partial q^i$. What we're left with is n first-order differential equations for the evolution of q^i. In this manner the function W determines the path of the classical system: start it off at a point in configuration space and W can be considered as a real valued classical wavefunction which tells it how to evolve. What we need to show is that the evolution dictated by (4.161) does indeed satisfy the equations of motion. In other words, we should prove that the second of Hamilton's equations, $\dot{p}_i = -\partial H/\partial q^i$, is satisfied. We have

$$\dot{p}_i = \frac{d}{dt}\left(\frac{\partial W}{\partial q^i}\right) = \frac{\partial^2 W}{\partial q^i \partial q^j}\dot{q}^j + \frac{\partial^2 W}{\partial t \partial q^i} \tag{4.162}$$

But differentiating the Hamilton-Jacobi equation (4.160) with respect to q^i, we see that we can rewrite the right-hand-side of this equation using

$$\frac{\partial^2 W}{\partial t \partial q^i} = -\frac{\partial H}{\partial q^i} - \frac{\partial H}{\partial p_j}\frac{\partial^2 W}{\partial q^i \partial q^j} = -\frac{\partial H}{\partial q^i} - \dot{q}^j \frac{\partial^2 W}{\partial q^i \partial q^j} \tag{4.163}$$

So that (4.162) becomes $\dot{p}_i = -\partial H/\partial q^i$ as required.

Let's see what we've done. We're used to dealing with second order differential equations for the time evolution on configuration space (i.e. Lagrange's equations) and first order differential equations for time evolution on phase space (Hamilton's equations). But the Hamilton-Jacobi approach allows us to incorporate n of the integration constants in the function $W(q^i, t)$ so that we're left solely with first order differential equations on configuration space given by (4.161).

When we have conservation of the Hamiltonian, so $\partial H/\partial t = 0$, there is solution of the Hamilton-Jacobi equation of a particularly simple form. Define

$$W(q^i, t) = W^0(q^i) - Et \tag{4.164}$$

for some constant E. Then the time dependence drops out and we get the equation

$$H(q^i, \partial W^0/\partial q^i) = E \tag{4.165}$$

W^0 is known as *Hamilton's principal function*. The special property of this solution to the Hamilton-Jacobi equation is that every path in configuration space determined by the function W_0 has the same energy E. With a little thought, we can envisage how to construct solutions to (4.165). Start with a co-dimension one surface in configuration space which we will specify to be a surface of constant W_0. (Co-dimension one means that the surface has dimension $(n-1)$: it splits the configuration space in two). At any point in this surface, the potential energy $V(q)$ is determined. Since $p_i = \partial W_0/\partial q^i$, the momentum is perpendicular to the surface and in the direction of increasing W_0. Its magnitude is fixed by requiring that the total energy is E. But this magnitude then tells us the position of the next surface of constant W_0 (with incremental increase). In this manner, it should be clear that, in multi-dimensional configuration spaces there are many solutions to (4.165). It should also be clear that something singular happens to W_0 in regions where $V(q^i) = 0$. Finally, we note that even when $\partial H/\partial t = 0$, there exist other solutions W to (4.160) which encode families of trajectories on configuration space which have different energies.

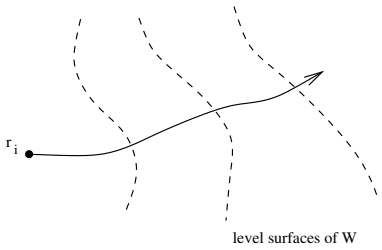


Figure 56:

4.7.1 Action and Angles from Hamilton-Jacobi

For the majority of examples the Hamilton-Jacobi approach doesn't give a particularly useful way for solving a problem; its utility really lies in the structure it reveals about classical dynamics. So rather than go through the gymnastics of solving a complicated problem using this method, let us focus on a rather simple example which which illustrates connections between the different ideas we've seen. A system with a single degree of freedom has Hamiltonian

$$H = \frac{p^2}{2m} + V(q) \tag{4.166}$$

Since the Hamiltonian is time independent, energy is conserved and the solution to the Hamilton-Jacobi equation has a single integration constant, let's call it β, which is necessarily some function of the energy. In the above discussion we were a little lax about showing these integration constants explicitly, but let's do it now: we'll write $W = W(q, t; \beta)$ with $\beta = \beta(E)$. Now we ask a somewhat strange question: suppose we try to perform a canonical transformation from (q, p) to new coordinates (α, β) such that β is the new momentum. What is the new coordinate α?

Since we wish the change of coordinates to be canonical, we must be able to write $q = q(\alpha, \beta)$ and $p = p(\alpha, \beta)$ such that

$$\{q, p\}_{(\alpha,\beta)} \equiv \frac{\partial q}{\partial \alpha}\frac{\partial p}{\partial \beta} - \frac{\partial q}{\partial \beta}\frac{\partial p}{\partial \alpha} = 1 \tag{4.167}$$

Using $p = \partial W/\partial q$ and remembering what all depends on what all ($W = W(q, \beta)$ and $q = q(\alpha, \beta)$ and $p = p(\alpha, \beta)$) we can write this as,

$$\{q, p\}_{(\alpha,\beta)} = \frac{\partial q}{\partial \alpha}\left(\frac{\partial^2 W}{\partial \beta \partial q} + \frac{\partial^2 W}{\partial q^2}\frac{\partial q}{\partial \beta}\right) - \frac{\partial q}{\partial \beta}\frac{\partial^2 W}{\partial q^2}\frac{\partial q}{\partial \alpha} = \frac{\partial q}{\partial \alpha}\frac{\partial}{\partial q}\left(\frac{\partial W}{\partial \beta}\right) \tag{4.168}$$

and we find that the transformation is canonical if we take $\alpha = \partial W/\partial \beta$. Note the nice symmetry here: we have a solution $W(q, t; \beta)$ to the Hamilton Jacobi equation and we can think in terms of canonical coordinates (q, p) or alternatively (α, β) where

$$p = \frac{\partial W}{\partial q} \quad , \quad \alpha = \frac{\partial W}{\partial \beta} \tag{4.169}$$

The function W is an example of a generating function of the second kind (4.95).

So what to do with this? Let's look at some examples. Take $\beta = E$, so that our new momentum variable is the energy itself. What is the canonical coordinate? If we write $W(q, t; E) = W_0(q, E) - Et$ then the coordinate canonically dual to E is

$$\alpha = \frac{\partial W_0}{\partial E}(q, E) - t \tag{4.170}$$

Taking the time dependence over the left-hand-side, we see that α has the interpretation of $-t_0$, the initial starting time. This tells us that we may parameterise every trajectory in a one-dimensional system in terms of the energy and starting time of the path, and that these are canonical variables. Again we see the dual relationship between energy and time. Note that both E and t_0 are independent of time; we've found canonical variables for which neither the coordinate nor the momentum vary along the path.

As another example consider the case of $\beta = I$, our action variable of Section 4.5. What is the canonical coordinate α in this case? We expect that it will be related to the angle variable θ. To see this, we use the fact that W is the classical action to write

$$W_0 = \int L\, dt + Et = \int (L + H) dt = \int \dot{q}p\, dt = \int p\, dq \tag{4.171}$$

So we have that

$$\alpha = \frac{\partial W}{\partial \beta} = \frac{d}{dI}\int p\, dq - \frac{dE}{dI}t = \theta - \omega t \tag{4.172}$$

where we've used our expression (4.119) for the angle variable, as well as the equation (4.116) for the frequency of motion ω. So we see that α is not quite equal to θ, but is shifted by a term linear in time. In fact this means that α itself does not change in time. Once again, we've arrived at a way to parameterise the orbits of motion by canonical variables which do not themselves change with time. In fact, in most presentations, this is the starting motivation for the Hamilton-Jacobi approach to classical dynamics and, even for higher dimensional systems, the function W can be thought of as a way to generate new, time independent, canonical variables. More discussion on the relationship between canonical transformations, angle-action variables and the Hamilton-Jacobi formulation can be found in the book by Hand and Finch, or deep within Goldstein.

4.8 Quantum Mechanics

One of the primary reasons for studying the rather formal aspects of classical mechanics discussed in this course is to make contact with quantum mechanics. For this reason, in this last section of the course we will illustrate the connection between the classical and quantum world and point out a few analogies that might just make the quantum behaviour look a little less weird. (Just a little less: after all, it really is weird!)

In classical mechanics the state of a system is described by a point (q^i, p_i) in phase space. In quantum mechanics the state is described by a very different object: a complex valued wavefunction $\psi(q)$ over the configuration space. The observables are operators on the space of wavefunctions. The standard representations of the position operator $\hat{q}^i$ and momentum operator $\hat{p}_i$ are

$$\begin{aligned} \hat{q}^i\psi(q) &= q^i\psi(q) \\ \hat{p}_i\psi(q) &= i\hbar\frac{\partial\psi}{\partial q^i} \end{aligned} \tag{4.173}$$

which leads to the well known Heisenberg commutation relations

$$\begin{aligned} [\hat{p}_i, \hat{p}_j] &= 0 \\ [\hat{q}^i, \hat{q}^j] &= 0 \\ [\hat{q}^i, \hat{p}_j] &= i\hbar\,\delta_{ij} \end{aligned} \tag{4.174}$$

Of course, we've already seen something very familiar in section 4.3 on Poisson brackets as summarised in equation (4.42). Although the bilinear, antisymmetric operators $[\,,\,]$ and $\{\,,\,\}$ act on very different spaces, they carry the same algebraic structure. Heuristically the relations (4.42) and (4.174) both represent the mathematical fact

that momentum p_i generates infinitesimal translations of q^i: in classical mechanics we saw this in when we studied infinitesimal canonical transformations in section 4.4.1; in quantum mechanics it follows from the representation (4.173) and Taylor's expansion.

In general the map between a classical system and a quantum system goes via the Poisson brackets and was formulated by Dirac:

$$\{\ ,\ \}_{\text{classical}} \leftrightarrow -\frac{i}{\hbar}[\ ,\ ]_{\text{quantum}} \tag{4.175}$$

This prescription for going between the classical and quantum theories is known as *canonical quantisation*. It also gives rise to the quantum equations of motion. In the Poisson bracket language, we have seen that the classical equation of motion for an arbitrary function $f(q,p)$ is

$$\dot{f} = \{f, H\} \qquad \longrightarrow \qquad i\hbar\dot{\hat{f}} = [\hat{f}, \hat{H}] \tag{4.176}$$

which is the equation of motion in the Heisenberg picture, in which the time dependence is assigned to the operator rather than the state.

While a great physicist, Dirac was never much of a storyteller. It shows in the following anecdote recounting his graduate student days:

> *"I went back to Cambridge at the beginning of October 1925, and resumed my previous style of life, intense thinking about these problems during the week and relaxing on Sunday, going for a long walk in the country alone. The main purpose of these long walks was to have a rest so that I would start refreshed on the following Monday.*
>
> *It was during one of the Sunday walks in October 1925, when I was thinking about this $(uv - vu)$, in spite of my intention to relax, that I thought about Poisson brackets. I remembered something which I had read up previously, and from what I could remember, there seemed to be a close similarity between a Poisson bracket of two quantities and the commutator. The idea came in a flash, I suppose, and provided of course some excitement, and then came the reaction "No, this is probably wrong".*
>
> *I did not remember very well the precise formula for a Poisson bracket, and only had some vague recollections. But there were exciting possibilities there, and I thought that I might be getting to some big idea. It was really a very disturbing situation, and it became imperative for me to brush up on*

my knowledge of Poisson brackets. Of course, I could not do that when I was right out in the countryside. I just had to hurry home and see what I could find about Poisson brackets.

I looked through my lecture notes, the notes that I had taken at various lectures, and there was no reference there anywhere to Poisson brackets. The textbooks which I had at home were all too elementary to mention them. There was nothing I could do, because it was Sunday evening then and the libraries were all closed. I just had to wait impatiently through that night without knowing whether this idea was really any good or not, but I still think that my confidence gradually grew during the course of the night.

The next morning I hurried along to one of the libraries as soon as it was open, and then I looked up Poisson brackets in Whitackers Analytical Dynamics*, and I found that they were just what I needed.*"

4.8.1 Hamilton, Jacobi, Schrödinger and Feynman

While the Poisson bracket structure of quantum mechanics dovetails nicely with Heisenberg's approach, the Hamilton-Jacobi equation is closely tied to Schrödinger's wave equation. Let's first recall what Schrödinger's equation looks like for a one-dimensional system with a Hamiltonian operator $\hat{H} = \hat{p}^2/2m + V(\hat{q})$ acting on wavefunctions $\psi(q)$,

$$i\hbar \frac{\partial \psi}{\partial t} = \hat{H}\psi = -\frac{\hbar^2}{2m}\frac{\partial^2 \psi}{\partial q^2} + V(q)\psi \tag{4.177}$$

where we have used the representation (4.173) for the position and momentum operators. To see the relationship of this equation to classical dynamics we decompose the wavefunction into the modulus and phase,

$$\psi(q,t) = R(q,t)\, e^{iW(q,t)/\hbar} \tag{4.178}$$

where R and W are both real functions. We know that R is related to the probability P for finding a particle at position q at time t: $P(q,t) = |\psi(q,t)|^2 = R(q,t)^2$. But what is the interpretation of the phase W? Let's substitute this decomposition of ψ into the Schrödinger equation to find

$$i\hbar\left[\frac{\partial R}{\partial t} + \frac{iR}{\hbar}\frac{\partial W}{\partial t}\right] = -\frac{\hbar^2}{2m}\left[\frac{\partial^2 R}{\partial q^2} + \frac{2i}{\hbar}\frac{\partial R}{\partial q}\frac{\partial W}{\partial q} - \frac{R}{\hbar^2}\left(\frac{\partial W}{\partial q}\right)^2 + \frac{iR}{\hbar}\frac{\partial^2 W}{\partial q^2}\right] + VR$$

At this stage, we take the classical limit $\hbar \to 0$. Or, more precisely, we consider a situation with

$$\hbar \left| \frac{\partial^2 W}{\partial q^2} \right| \ll \left| \frac{\partial W}{\partial q} \right| \tag{4.179}$$

which can be understood physically as the requirement that the de Broglie wavelength of the particle is much smaller than any other length scale around. Either way, collecting together the terms above to leading order in $\hbar$ we find

$$\frac{\partial W}{\partial t} + \frac{1}{2m}\left(\frac{\partial W}{\partial q}\right)^2 + V(q) = \mathcal{O}(\hbar) \tag{4.180}$$

which we recognise as the Hamilton-Jacobi equation (4.160). So in the classical limit the phase of the wavefunction is understood as the classical action of the path taken by the particle.

Finally, let us finish on the same topic that we started: the principle of least action. Recall from section 2.1 that we can determine the true path of a system by assigning a number, called the action S, to every possible path. The equations of motion are then equivalent to insisting that the true path is an extremum of S. But what about all the other paths? Do they play any role in Nature? The answer is that, in the quantum world, they do. Suppose a particle is observed to be at position q^i at time $t = 0$. Then the probability P that it will later be observed to be at position q_f at time $t = T$ is encapsulated in the wavefunction $\psi(q_f, T)$. The Feynman path integral formula for the wavefunction is

$$\psi(q_f, T) = N \int_{q^i}^{q_f} \mathcal{D}q(t)\, e^{iS[q(t)]/\hbar} \tag{4.181}$$

The N here is just a normalisation constant to ensure that probabilities add up to one: i.e. $\int |\psi(q)|^2 dq = 1$. The tricky part of this formula is the integral: it is a sum over all possible paths from $q = q^i$ at time 0 to $q = q_f$ at time T. These paths are weighted with their action. It's as if the particle really does take every possible path, but with a particular phase. In the limit $\hbar \to 0$, the phases oscillate wildly for any path away from the classical equation of motion $\delta S = 0$ and they cancel out in the integral. But for situations where $\hbar$ is important, the other paths are also important.

Let's prove that the wavefunction defined by (4.181) satisfies the Schrödinger equation. Firstly we need to understand this integral over paths a little better. We do this by splitting the path into n small segments, each ranging over a small time $\delta t = t/n$.

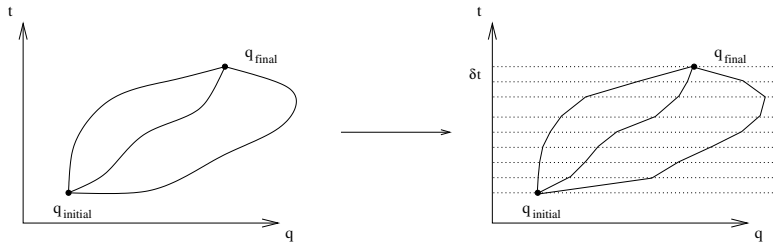


Figure 57: Discretising the paths.

Then we define

$$\int \mathcal{D}q(t) = \lim_{n\to\infty} \prod_{k=1}^{n} \int_{-\infty}^{+\infty} \frac{dq_k}{C} \tag{4.182}$$

where q_k is the position of the particle at time $t = k\delta t$. In this expression C is a constant that we're going to figure out shortly that will be required to make sense of this infinite number of integrals. In any given segment, we treat the path as straight lines as shown in the figure and replace the action with the appropriate quantity,

$$S = \int_0^T dt \ \left(\frac{1}{2}m\dot{q}^2 - V(q)\right) \to \sum_{k=1}^{n} \left(\frac{m}{2}\frac{(q_{k+1}-q_k)^2}{\delta t} - \delta t\, V\left(\frac{q_{k+1}+q_k}{2}\right)\right)$$

Then to prove that ψ defined in (4.181) satisfies Schrödinger's equation, let's consider adding a single extra time step to look at the wavefunction at time $t + \delta t$. We can Taylor expand the left hand side of (4.181) happily

$$\psi(q_f, T+\delta t) = \psi(q_f, T) + \frac{\partial \psi}{\partial T}\,\delta t + \mathcal{O}(\delta t^2) \tag{4.183}$$

while the right hand side requires us to do one extra integral over the penultimate position of the path q'. But the first n integrals simply give back the original wavefunction, now evaluated at q'. We get

$$\int_{-\infty}^{+\infty} \frac{dq'}{C} \ \exp\left[\frac{im}{2\hbar}\frac{(q_f-q')^2}{\delta t} - \frac{i}{\hbar}\delta t\, V\left(\frac{q_f+q'}{2}\right)\right] \psi(q', t) \tag{4.184}$$

The term in the exponent means that the integral oscillates wildly whenever q' is far from q_f and these regions of the integral will all cancel out. We can therefore Taylor expand around $(q_f - q')$ to rewrite this as

$$\int_{-\infty}^{+\infty} \frac{dq'}{C} \ \exp\left[\frac{im}{2\hbar}\frac{(q_f-q')^2}{\delta t}\right] \left(1 - \frac{i\delta t}{\hbar}V(q_f) + \ldots\right) \left(1 + (q'-q_f)\frac{\partial}{\partial q_f} + \tfrac{1}{2}(q'-q_f)^2\frac{\partial^2}{\partial q_f^2} + \ldots\right) \psi(q_f, T) \tag{4.185}$$

At this stage we do the integral over q'. We'll use the formulae for Gaussian integration

$$\int dy\, e^{-ay^2} = \sqrt{\frac{\pi}{a}} \quad , \quad \int dy\, y e^{-ay^2} = 0 \quad , \quad \int dy\, y^2 e^{-ay^2} = \frac{1}{2a}\sqrt{\frac{\pi}{a}} \tag{4.186}$$

Then equating the lefthand side (4.183) with the righthand side (4.185), we have

$$\psi(q_f, T) + \frac{\partial \psi}{\partial T}\,\delta t = \frac{1}{C}\sqrt{\frac{2\pi\hbar\delta t}{-im}}\left[1 - \frac{i\delta t}{\hbar}V(q_f) + \frac{i\hbar\delta t}{2m}\frac{\partial^2}{\partial q_f^2} + \mathcal{O}(\delta t^2)\right]\psi(q_f, T)$$

At this stage we see what the constant C has to be to make sense of this whole calculation: we should take

$$C = \sqrt{\frac{2\pi\hbar\delta t}{-im}} \tag{4.187}$$

so that $C \to 0$ as $\delta t \to 0$. Then the terms of order $\mathcal{O}(\delta t^0)$ agree. Collecting the terms of order $\mathcal{O}(\delta t)$, and replacing the time T at the end point with the general time t, we see that we have

$$i\hbar\frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2 \psi}{\partial q_f^2} + V(q_f)\psi = \hat{H}\psi \tag{4.188}$$

and we recover Schrödinger's equation as promised.

4.8.2 Nambu Brackets

Throughout this section, we've seen that several of the structures appearing in quantum mechanics were anticipated, in some form, within the framework of classical dynamics. You just need to know where to look. One might wonder whether classical mechanics also contains other structures which will prove to be important in future discoveries. Or, alternatively, whether there are ways to extend the framework of classical dynamics that hints at new ways to formulate the laws of physics. In this section, I'll briefly describe such an extension due to Nambu in 1973. I should confess immediately that there's no known use for Nambu's formalism! And it's not at all clear that one will be found! But then again, maybe it holds the clue that will prove crucial in the search for the ideas beyond the known laws of Nature.

We've seen that the Hamiltonian framework deals with canonical pairs of coordinates and momenta (q^i, p_i) with $i = 1, \ldots, n$. Nambu's idea was to extend this to triplets of objects (q^i, p_i, r_i) with $i = 1, \ldots, n$. We don't say what this extra variable r_i is: just that it is necessary to define the state of a system. This means that the phase space

has dimension $3n$. The Nambu bracket, which replaces the Poisson bracket, acts on three functions f, g and h in phase space,

$$\{f,g,h\} = \sum_i \frac{\partial f}{\partial q^i}\frac{\partial g}{\partial p_i}\frac{\partial h}{\partial r_i} - \frac{\partial f}{\partial q^i}\frac{\partial g}{\partial r_i}\frac{\partial h}{\partial p_i} + \frac{\partial f}{\partial r_i}\frac{\partial g}{\partial q^i}\frac{\partial h}{\partial p_i} - \frac{\partial f}{\partial r_i}\frac{\partial g}{\partial p_i}\frac{\partial h}{\partial q^i} + \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial r_i}\frac{\partial h}{\partial q^i} - \frac{\partial f}{\partial p_i}\frac{\partial h}{\partial q^i}\frac{\partial h}{\partial r_i} \tag{4.189}$$

This satisfies similar properties to the Poisson bracket, including linearity and

- Anti-symmetry: $\{f,g,h\} = -\{g,f,h\} = \{g,h,f\}$.
- Leibniz: $\{fg,h,l\} = f\{g,h,l\} + \{f,h,l\}g$.
- "Jacobi": $\{\{f,g,h\},l,m\} + \{h,\{f,g,l\},m\} + \{h,l,\{f,g,m\}\} = \{f,g,\{h,l,m\}\}$.

In order to specify time evolution, we need two "Hamiltonians". We call them $H(q,p,r)$ and $G(q,p,r)$. Then any function f over phase space evolves as

$$\frac{df}{dt} = \{f,G,H\} \tag{4.190}$$

In particular, the new version Hamilton's equations read

$$\begin{aligned} \dot{q}^i &= \frac{\partial G}{\partial p_i}\frac{\partial H}{\partial r_i} - \frac{\partial G}{\partial r_i}\frac{\partial H}{\partial p_i} \\ \dot{p}_i &= \frac{\partial G}{\partial r_i}\frac{\partial H}{\partial q^i} - \frac{\partial G}{\partial q^i}\frac{\partial H}{\partial r_i} \\ \dot{r}_i &= \frac{\partial G}{\partial q^i}\frac{\partial H}{\partial p_i} - \frac{\partial G}{\partial p_i}\frac{\partial H}{\partial q^i} \end{aligned} \tag{4.191}$$

where there's no sum over i on the right-hand side of these equations. By the anti-symmetry of the Nambu bracket, we learn that both H and G are conserved (as long as neither have explicit time dependence).

Many of the key features of classical dynamics are retained in Nambu's formalism. For example, Liouville's theorem still holds. (This was Nambu's original motivation for suggesting this framework). Similarly, canonical transformations can be defined as a change of variables $q^i \to q^i(q,p,r)$ and $p_i \to P_i(q,p,r)$ and $r_i \to R_i(q,p,r)$ such that the Nambu bracket structure is preserved, for example

$$\{q^i, P_j, R_k\} = \begin{cases} 1 & \text{if } i=j=k \\ 0 & \text{otherwise} \end{cases} \tag{4.192}$$

together with similar equations involving other combinations of Q's, P's and R's. "Hamilton's" equations (4.191) are invariant under these canonical transformations.

The Nambu bracket provides a generalisation of classical dynamics. But can we quantise it? In other words, can we find some operators which reproduce the Nambu bracket structure (up to a factor of $\hbar$ and perhaps an i) in much the same way that the usual quantum theory ties in with the Poisson bracket? This turns out to be pretty tricky. In particular, it seems difficult to keep all three conditions: anti-symmetry, Leibniz and Jacobi. Perhaps this suggests that the correct mathematical structure has not yet been uncovered. Perhaps it suggests that the Nambu bracket is just not useful!

Chances are that you won't ever have any use for the Nambu bracket. But you never know. Perhaps one day, like Dirac, you'll return from wandering the fields around Cambridge and desperately need to recall this concept. But, unlike Dirac, you'll be able to find a reference in your lecture notes (the notes you had taken at various lectures).