

Guru's Guide to Advanced Mechanics

Brett Witty

September 3, 2001

Contents

1	Preliminaries	1
1.1	Systems	1
1.2	Work and Line Integrals	2
1.2.1	Line Integrals	2
1.2.2	Work	4
2	Basic Mechanics	7
2.1	Conservative Systems	7
2.2	Conservative Forces	10
2.3	Central Forces	13
2.4	Conservation of Angular Momentum	14
3	Concepts in Advanced Mechanics	17
3.1	Generalised Coordinates	17
3.2	Constraints	20
3.3	Generalised Forces	21
3.4	Conservative Systems	22
3.5	Equilibrium in Conservative Systems	22
4	Lagrange's Method of Dynamics	25
4.1	Overview of Lagrange's and Hamilton's Methods	25
4.2	Lagrange's Method	25
4.3	Lagrange's Equation	28
5	Hamilton's Method of Dynamics	31
5.1	Hamilton's Method	31
5.2	Derivation of Lagrangian from the Hamiltonian	34
5.3	Conservation of Energy	35

6	Calculus of Variations	37
6.1	Some History and the Euler-Lagrange Equation	37
6.2	Extending Euler-Lagrange	40
6.2.1	Generalisation to n dimensions	40
6.2.2	F independent of x	40
6.3	Calculus of Variations at work	41
6.3.1	Brachistochrone Problem	41
6.3.2	Snell's Law	44
6.4	Principle of Least Action	45
7	Phase Spaces and Canonical Transformations	49
7.1	Phase Space	49
7.2	Canonical Transformations	51
A	Conclusion	55

Introduction

Welcome to the Guru's Guide to Advanced Mechanics. This text primarily covers basic mechanics using calculus, Lagrangian Mechanics and Hamiltonian Mechanics. We assume that the reader has a working knowledge of calculus (differential and integral), some knowledge of basic physics and some mathematical maturity. Typically this subject is taught at a second year level at university, but you don't really need too much to delve into the subject.

As the methods in advanced mechanics are very dependant upon calculus, the reader can expect to do derivatives until the cows come home. Particularly know all the tricks of the trade, such as the product rule and total differentials. Integrals aren't used as much, but still feature, and so the reader should be confident, or motivated enough to revise.

What is Advanced Mechanics?

What we call here Advanced Mechanics is more of an advanced treatment of mechanics rather than anything astonishingly new. The maths is a little more sophisticated, but allows us to easily obtain elegant answers to typically messy systems. For example, if we were to work out all the Newtonian force diagrams for a mass suspended by two springs, it could get very messy. However, by applying powerful general mathematical techniques, we can solve the system relatively quickly and painlessly. Plus the maths isn't all that hard.

Acknowledgements

Thanks to Mark Gould whose lectures were the basis for this text. Thanks to Joel, Mark, Rowan, Duncan and Jenny for reading my stuff and pointing out the many icky errors. Finally, thanks to Sarah, Raphael and the coke

machine down-the-way for moral support. Many thanks to all the readers for going through my bizarre text and finding problems everywhere :) Also, Synge and Griffith deserve special thanks for their excellent book: “Principles of Mechanics” which cleared up a lot of debated issues and helped me shape some of the text.

Chapter 1

Preliminaries

1.1 Systems

In all mechanical problems, we consider **systems**. What is a system exactly? A system is basically our universe of interest. If we are considering a mass on a spring, we consider a mass on a spring. We need not incorporate wind resistance, thermal effects or the economic depreciation of eggs in China, unless we need to. Basically, we simplify the system to look at exactly what we want to analyze. Sometimes these simplifications sound weird. Here is a list of standard assumptions:

- Strings are infinitely thin, massless, unbreakable, unstretchable and usually frictionless.
- Springs are infinitely thin, massless, infinitely stretchable and compressible and never permanently deform.
- Pulleys are frictionless and massless.
- The world is flat and infinitely huge.
- There is no air resistance (unless we want it there).
- Gravity always acts straight down.
- Most objects are infinitesimally small, have finite mass and don't permanently deform.
- Everything acts like in Newtonian mechanics, unless we are studying Relativity or Quantum Mechanics.

Some of these simplifications sound ridiculous and provide a lot of humour against physicists. However, does anyone care about the negligible effect of a curved Earth when we throw a ball ten metres? Do you want to consider the gravitational effect of Alpha Centuri on your pendulum? The correct answer to those two questions is no. However, we must keep our simplifications in mind or we may be confused if we overstep the bounds of our system and get weird results. For example, if we think about throwing a ball ten metres, and calculate everything correctly, we shall get good results. However, if we throw a ball at supersonic speeds and send it into orbit, we shall begin to wonder why our calculations aren't working correctly.

So the moral of this story is to make your system simple enough to analyse it, but make sure you don't overstep your assumptions!

1.2 Work and Line Integrals

1.2.1 Line Integrals

One of the oldest questions is “How long is a piece of string?” While not being able to answer this silly question, it'd be nice if we had a method to mathematically measure things other than straight lines. As you may have expected, line integrals is the method we are looking for.

If you were given the task to measure a wiggly line with a ruler, chances are you'd measure little sections along it, saying that each section is reasonably straight. If I asked you to be more accurate, you'd probably make more segments. Which is basically how we approach it mathematically. We call the length of the line the *arclength* and find it by the following method:

Proposition 1.1. *Given a curve in 3-dimensional space, with parameterisation:*

$$\mathbf{r}(t) = x(t)\hat{i} + y(t)\hat{j} + z(t)\hat{k}$$

Now the length of the arc of r between $t = a$ and $t = b$ is given by:

$$L = \int_a^b |\dot{\mathbf{r}}(t)| dt$$

Sketch of Proof. We segment the curve into n different linear approximations, each of length corresponding to a parameter difference of Δt . Now, the distance between two points $\mathbf{r}(t)$ and $\mathbf{r}(t + \Delta t)$ on the curve, where Δ is really small, is given by:

$$\mathbf{r}(t + \Delta t) - \mathbf{r}(t)$$

Now by Taylor Expansion of $\mathbf{r}(t + \Delta t)$:

$$\begin{aligned} & \mathbf{r}(t + \Delta t) - \mathbf{r}(t) \\ &= \mathbf{r}(t) + \mathbf{r}'(t)\Delta t - \mathbf{r}(t) \\ &= \mathbf{r}'(t)\Delta t \end{aligned}$$

Now if we sum all the segments¹, we get:

$$\begin{aligned} L &= \sum_{i=1}^n |\mathbf{r}_i(t + \Delta t) - \mathbf{r}_i(t)| \\ &= \sum_{i=1}^n |\mathbf{r}'_i(t)\Delta t| \end{aligned}$$

(We take the absolute value to get distance rather than displacement)

If we take the limit as $n \rightarrow \infty$ then $\Delta t \rightarrow dt$ and the length will be given by:

$$L = \int_a^b |\mathbf{r}'(t)| dt$$

□

Example 1.2.1.

Find the circumference of a circle of radius r .

Answer.

$$\mathbf{r}(t) = r \sin t \hat{i} + r \cos t \hat{j}$$

where $0 \leq t \leq 2\pi$. Therefore $\mathbf{r}'(t) = r \cos t \hat{i} - r \sin t \hat{j}$ and then the absolute value is given by:

$$\begin{aligned} |\mathbf{r}'| &= \sqrt{r^2 \cos^2 t + r^2 \sin^2 t} \\ &= r \end{aligned}$$

So the circumference is given by:

$$\begin{aligned} L &= \int_0^{2\pi} r dt \\ &= [rt]_0^{2\pi} \\ &= 2\pi r \end{aligned}$$

which you already know to be true.

¹This is just like the formulation for Riemann Integrals, which you should have seen before.

Example 1.2.2.

Find the arclength of $\mathbf{r}(t) = \cosh t \hat{i} + \sinh t \hat{j} + t \hat{k}$, where $0 \leq t \leq 1$.

Answer.

$$\begin{aligned} \mathbf{r}(t) &= \cosh t \hat{i} + \sinh t \hat{j} + t \hat{k} \\ \mathbf{r}'(t) &= \sinh t \hat{i} + \cosh t \hat{j} + \hat{k} \\ |\mathbf{r}'(t)| &= \sqrt{\sinh^2 t + \cosh^2 t + 1} \\ &= \sqrt{2 \cosh^2 t} \quad (\text{As } \cosh^2 t - \sinh^2 t = 1) \\ &= \sqrt{2} \cosh t \end{aligned}$$

Then the length is given by:

$$\begin{aligned} L &= \int_0^1 |\mathbf{r}'(t)| dt \\ &= \sqrt{2} \int_0^1 \cosh t dt \\ &= \sqrt{2} [\sinh t]_0^1 \\ &= \sqrt{2} \sinh(1) \end{aligned}$$

Exercise 1.2.1.

Find the arclength for the circular helix described by $\mathbf{r}(t) = a \cos \omega t \hat{i} + a \sin \omega t \hat{j} + ct \hat{k}$ for $0 \leq t \leq \theta$.

Answer.

$$\sqrt{a^2 \omega^2 + c^2} \theta$$

1.2.2 Work

You should know from basic dynamics that the work W required to move an object a lineal distance \mathbf{d} is given by:

$$W = \mathbf{F} \cdot \mathbf{d}$$

where F is the force we applied to move it.² To apply this idea to an arbitrary smooth curve in space, we consider the force F acting upon each

²Don't forget that \mathbf{F} and \mathbf{d} are vectors and work (a scalar) is the dot product of \mathbf{F} and \mathbf{d}

segment of our line integral formulation. This is basically just our standard linear work relation, but applied to every segment in our line integral formulation, and then summed together.

Mathematically we say:

$$W = \int_r \mathbf{F} \cdot d\mathbf{r} = \int_a^b \mathbf{F} \cdot \mathbf{r}'(t) dt$$

Example 1.2.3.

Find the work done on a particle travelling in a helix

$$\mathbf{r}(x, y, z) = (\cos t, \sin t, t), \quad 0 \leq t \leq \frac{\pi}{2}$$

with force

$$\mathbf{f}(t) = \hat{i} + \hat{j} + t\hat{k}$$

Answer.

We need to make $\mathbf{r}(x, y, z)$ a function of time, and then find its time derivative to match our formula.

$$\begin{aligned} \mathbf{r}(t) &= \cos t \hat{i} + \sin t \hat{j} + t \hat{k} \\ \dot{\mathbf{r}}(t) &= -\sin t \hat{i} + \cos t \hat{j} + \hat{k} \\ W &= \int \mathbf{f} \cdot d\mathbf{r} = \int_0^{\frac{\pi}{2}} \mathbf{f}(t) \cdot \dot{\mathbf{r}}(t) dt \\ &= \int_0^{\frac{\pi}{2}} (-\sin t + \cos t + t) dt \\ &= \left[\cos t + \sin t + \frac{t^2}{2} \right]_0^{\frac{\pi}{2}} \\ &= \left(0 + 1 + \frac{\pi^2}{8} \right) - (1 + 0 + 0) \\ &= \frac{\pi^2}{8} \end{aligned}$$

So the work done is $\frac{\pi^2}{8}$ Joules.

Now a concept that is very important for the next section is that if we do work on an object, then the object will increase its kinetic energy in exactly the same manner. Mathematically:

$$\text{Work Done} = \text{Increase in Kinetic Energy}$$

How do we show that this holds? Let's consider the work done by a force \mathbf{F} in moving a particle along its path \mathbf{r} from time $t = 0$ to a time t . Now this is expressed by:

$$W = \int_r \mathbf{F} \cdot d\mathbf{r} = \int_0^t \mathbf{F} \cdot \mathbf{v}(t) dt$$

where $\mathbf{v} = \dot{\mathbf{r}}$.

Now the change in work is given by:

$$\begin{aligned} \frac{dW}{dt} = \mathbf{F} \cdot \mathbf{v}(t) &= m\dot{\mathbf{v}} \cdot \mathbf{v} \\ &= \frac{1}{2}m\frac{d}{dt}(\mathbf{v} \cdot \mathbf{v}) \quad (\text{from the product rule}) \end{aligned}$$

Now the right side looks strangely like the formula for kinetic energy ($\frac{1}{2}m\mathbf{v} \cdot \mathbf{v}$) except for that derivative. So then we can conclude:

$$\frac{dW}{dt} = \dot{T}$$

where T is the kinetic energy³. Then we can say:

$$\begin{aligned} \dot{T} &= \mathbf{F} \cdot \mathbf{v} \Rightarrow T(t) - T(0) = \int_0^t \mathbf{F} \cdot \mathbf{v} dt \\ \therefore W &= T(t) - T(0) \end{aligned}$$

³ T is used by convention instead of the more logical K . We at Guru aren't quite sure why, but are looking into it.

Chapter 2

Basic Mechanics

In this chapter we shall look at a few basic concepts in mechanics with a very mathematical bent.

2.1 Conservative Systems

A system is conservative if it is such that energy is conserved. Conservation of Energy is an important concept we often use to solve problems, so it'd be nice to know if we can use this property or not. Mathematically, we say that a system is conservative¹ if there is a function $V = V(x, y, z)$ such that:

$$\mathbf{F} \cdot \mathbf{v} = -\frac{dV}{dt}$$

We call V the *potential function* which is basically the potential energy we usually talk about. But given this potential function, how does this show that the energy of a system is conserved? Well with some algebraic manipulation:

$$\begin{aligned} -\frac{dV}{dt} &= \mathbf{F} \cdot \mathbf{v} = \dot{T} \\ \therefore \frac{dT}{dt} + \frac{dV}{dt} &= 0 \\ \text{So } \frac{d}{dt}(T + V) &= 0 \end{aligned}$$

¹We must keep in mind that conservative here applies to the system. Later we shall consider conservative forces, which is quite different.

So $T + V$ is a constant. However, T is the kinetic energy and V is the potential energy, so $T + V$ will be the energy of our system, and thus we have shown that if a system is conservative, its energy is constant.

We must also note that our potential function is unique, except for a constant. So $V + 1, V + 100, V + \pi$ are all acceptable potential functions if V is a potential function. This may strike you as odd, although the maths makes sense (look at the algebra above and see why). In effect, potential could start anywhere. However, this becomes understandable if we think about measuring potential energy in a lab. I want to know how fast a ball is travelling when it hits the ground when being dropped from a certain height. I assume my system conserves energy, and so my potential function is like above. I want to set up my system, but I want to know what I consider a height of zero. Is it the top of my desk, the floor or the ground two storeys down? The fact is that it doesn't matter. The zero height of my desk reference system is the 1 metre height in my floor reference system and is the 9 metre height in my ground reference system. But the fact of the matter is that I'm dropping the ball half a metre, which is the distance from my hand to the desk top. The fact that my hand is 1.5 metres high and the desk top is 1m high in the floor reference system makes little difference. The distance dropped is the same! Where I measure my zero height is the constant that I can change in my potential function.

It is also noteworthy that the requirement for conservative systems applies for one- and two-dimensional systems.

So what is a non-conservative system? Well any system with friction is a prime example of a non-conservative system. This is due to the fact that the work done by the friction will effectively be lost from the system as the energy gets dissipated as heat, sound or other vibrations. Thus the energy won't be constant with time, and so won't be conservative.

Example 2.1.1.

Let's look at a simple pendulum to look at a conservative system at work.

Resolving the forces in x and y :

$$\begin{aligned} m\ddot{x} &= -T \sin \theta \\ m\ddot{y} &= T \cos \theta - mg \end{aligned}$$

Then we have the constraint: $x^2 + y^2 = l^2$ This is obvious as the string on the pendulum is going to stay a constant length, and we can use Pythagoras' Theorem to give the length in terms of x and y .

From trigonometry we can also say: $x = l \sin \theta$ and $y = -l \cos \theta$

Now instead of looking at x and y in our system, we shall consider our system in terms of a vector \mathbf{r} and the angle of tilt, θ . Now \mathbf{r} is a vector with a length l and direction dependent on θ . So it is common practice to say: $\mathbf{r} = l \hat{\mathbf{r}}$ where $\hat{\mathbf{r}} = \sin \theta \hat{\mathbf{i}} - \cos \theta \hat{\mathbf{j}}$

Our equation of motion is given by:

$$\begin{aligned} m\ddot{\mathbf{r}} = \mathbf{F} &= -T \hat{\mathbf{r}} - mg \hat{\mathbf{j}} \\ &= -\frac{T}{l} \mathbf{r} - mg \hat{\mathbf{j}} \end{aligned}$$

We also define a unit rotation vector $\hat{\theta}$, which has the following relation:

$$\hat{\theta} = \frac{d\hat{\mathbf{r}}}{d\theta} = \cos \theta \hat{\mathbf{i}} + \sin \theta \hat{\mathbf{j}}$$

Note that $\hat{\mathbf{r}} \cdot \hat{\theta} = 0$ as $\hat{\mathbf{r}} \perp \frac{d\hat{\mathbf{r}}}{d\theta} = \hat{\theta}$.

Some people have a problem with $\hat{\theta}$. It looks weird and mysterious. But there really isn't anything to it. It's just a unit rotation vector ("Is that all?" some may sarcastically reply). If you recall the definition of a vector, you remember that a vector has a magnitude and a direction. When we define a vector with a hat on it, it means the vector has unit magnitude. So $\hat{\theta}$ has a unit magnitude, so it is a rotation of 1 radian. Our rotation vector has a defined direction, clockwise or anticlockwise (it doesn't really matter in this case, because our system is symmetric).

Then we find the velocity:

$$\mathbf{v} = \dot{\mathbf{r}} = l \frac{d\hat{\mathbf{r}}}{dt} = l \frac{d\hat{\mathbf{r}}}{d\theta} \frac{d\theta}{dt} = l \dot{\theta} \hat{\theta}$$

Now since we know the force and the velocity, we can see whether the simple pendulum is conservative or not:

$$\begin{aligned} \mathbf{F} \cdot \mathbf{v} &= (-T \hat{\mathbf{r}} - mg \hat{\mathbf{j}}) \cdot \mathbf{v} \\ &= -T \hat{\mathbf{r}} \cdot \mathbf{v} - mg \hat{\mathbf{j}} \cdot \mathbf{v} \\ &= -mg \hat{\mathbf{j}} \cdot \mathbf{v} \quad (\text{As } \hat{\mathbf{r}} \perp \mathbf{v}) \\ &= -mg \dot{y} \\ &= -\frac{dV}{dt} \quad (\text{where } V = mgy) \end{aligned}$$

(Note that tension does no work)

Therefore our system is conservative with energy:

$$E = \frac{1}{2}mv^2 + mgy = \frac{1}{2}ml^2\dot{\theta}^2 + mgy$$

Exercise 2.1.1.

Show that, for small θ , the system's equation of motion (in terms of θ) is that of a simple harmonic oscillator, namely:

$$\ddot{\theta} + \omega^2\theta = 0$$

2.2 Conservative Forces

We talked about conservative systems and how if we can find a potential function satisfying some properties, then the motion of the system is conservative, and hence energy is conserved in that system. Another way to look at this is by conservative forces, whose requirements look very similar to the ones for conservative systems. But first, a little vector calculus you will need:

Definition. If we have a function $g : \mathbb{R}^3 \mapsto \mathbb{R}^3$, we call the vector function:

$$\nabla g = \frac{\partial g}{\partial x} \hat{i} + \frac{\partial g}{\partial y} \hat{j} + \frac{\partial g}{\partial z} \hat{k}$$

the *gradient* of g . We also call:

$$\nabla = \frac{\partial}{\partial x} \hat{i} + \frac{\partial}{\partial y} \hat{j} + \frac{\partial}{\partial z} \hat{k}$$

the *gradient operator*.

Given this, we can define a conservative force:

Definition. A force \mathbf{f} is called *conservative* if there exists a function $g(x, y, z)$ such that:

$$\mathbf{f} = \nabla g$$

We usually say more specifically, $\mathbf{f} = -\nabla V$, where $V(x, y, z)$ is called the *potential function*, as before.

So what are the use of conservative forces? Well the following theorem should prove useful:

Theorem 2.1. If a particle of mass m moves under the influence of a conservative force \mathbf{F} and $\mathbf{F} = -\nabla V$, then the system is conservative.

Proof. We said before that if $\mathbf{F} \cdot \mathbf{v} = -\frac{dV}{dt}$ then the system is conservative. So that's our goal.

$$\begin{aligned}
 \mathbf{F} \cdot \mathbf{v} &= -\nabla V \cdot \mathbf{v} \\
 &= -\nabla V \cdot \dot{\mathbf{r}} \\
 &= -\frac{\partial V}{\partial x} \dot{x} - \frac{\partial V}{\partial y} \dot{y} - \frac{\partial V}{\partial z} \dot{z} \\
 &= -\frac{\partial V}{\partial x} \frac{dx}{dt} - \frac{\partial V}{\partial y} \frac{dy}{dt} - \frac{\partial V}{\partial z} \frac{dz}{dt} \\
 &= -\frac{dV}{dt} \quad (\text{By recognizing total differential})
 \end{aligned}$$

□

So a conservative force gives us a conservative system, which is what we want (everybody loves Conservation of Energy). We can generalize this by using the Jacobi matrix, which we shall define:

Definition. The **Jacobi Matrix** of a function $\mathbf{f} = f_1(x_1, x_2, \dots, x_n)\hat{i} + f_2(x_1, x_2, \dots, x_n)\hat{j} + f_3(x_1, x_2, \dots, x_n)\hat{k}$ is the $n \times n$ matrix with elements:

$$[J_{\mathbf{f}}]_{ij} = \frac{\partial f_i}{\partial x_j} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix}$$

Usually n is three, as we only deal with 3D. Then x_1, x_2 and x_3 correspond to x, y and z .

Another useful theorem is:

Lemma 2.1. A force \mathbf{f} is conservative $\iff J_{\mathbf{f}}$ is symmetric.

Proof. \Rightarrow Suppose $\mathbf{f} = -\nabla V$ is conservative.

$$\begin{aligned}
 [J_{\mathbf{f}}]_{ij} = \frac{\partial f_i}{\partial x_j} &= -\frac{\partial^2 V}{\partial x_j \partial x_i} \\
 &= -\frac{\partial^2 V}{\partial x_i \partial x_j} \\
 &\quad (\text{Order of partial differentials does not matter}) \\
 &= [J_{\mathbf{f}}]_{ji}
 \end{aligned}$$

So $J_{\mathbf{f}}$ is symmetric and it is easy to see that the implication the other way follows. \square

Example 2.2.1.

Show whether the following forces are conservative or not:

1. $\mathbf{f} = x^2y\hat{i} + (x^2 - y^2)\hat{j} + xz\hat{k}$

We use Lemma 2.1:

$$J_{\mathbf{f}} = \begin{bmatrix} 2xy & x^2 & 0 \\ 2x & -2y & 0 \\ z & 0 & 1 \end{bmatrix}$$

Obviously this matrix is not symmetric, and so the force is not conservative.

2. $\mathbf{f} = (x + y)\hat{i} + x\hat{j}$ (Note our method does not change if we are in dimensions other than three)

Using Lemma 2.1:

$$J_{\mathbf{f}} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

This matrix is symmetric and so the force \mathbf{f} is conservative.

Also, given that this force is conservative, we can deduce the potential function. We do this by noting $\mathbf{f} = \nabla g$, which we can solve the exact equation to give us g . And from above, $g = -V$.

3. If we consider Hooke's Law, we can see whether a spring is a conservative force:

$$(\text{Hooke's Law})F = -kx$$

So therefore $F = -\frac{dV}{dx}$ and so integrating gives us:

$$V = \frac{1}{2}kx^2$$

(You may have seen the potential energy in a spring before) Since we have found that $F = -\nabla V$, we've shown that F is conservative. So a spring system is conservative with constant energy given by:

$$E = \frac{1}{2}mv^2 + \frac{1}{2}kx^2$$

Also of note (for later on), we can solve the equation of motion:

$$m\ddot{x} = -kx$$

Solving the differential equation, we get:

$$x = A \cos \omega t + B \sin \omega t$$

where $\omega = \sqrt{\frac{k}{m}}$. A and B are constants based on properties of your system, as is ω .

4. Now if we look at Theorem 2.1, someone who didn't read it carefully may say that a conservative system needs a conservative force. However Theorem 2.1 isn't an if and only if. Here's an example of a normal system where the converse does not apply:

In a pendulum, we have a force $\mathbf{F} = -T\hat{\mathbf{r}} - mg\hat{\mathbf{j}}$. \mathbf{F} is not conservative, but we know the system is conservative as it has a potential energy of $V = mgy$, satisfying the conservative system requirements.

2.3 Central Forces

Simple systems like a mass on a spring or a pendulum are fairly easy to look at with our tools we formulated above. These forces involved are reasonably straightforward, and are mostly linear forces. But what about stranger things, like electromagnetic forces or gravity? These forces are radial, but more specifically, we call them **central forces**.

Definition. A force $\mathbf{f}(x, y, z)$ is called **central** if it has the form:

$$\mathbf{f} = f(r)\hat{\mathbf{r}} = \frac{f(r)\mathbf{r}}{r}$$

where $\mathbf{r} = x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}}$ and $r = \sqrt{x^2 + y^2 + z^2}$.

Central forces act either towards or away from the origin (depending on their sign) with a magnitude dependent only on the distance the object is away from the origin.

The first question you've probably asked is, are central forces conservative? Well, they are:

Theorem 2.2. If there exists a function $V(r)$ such that $\mathbf{f}(r) = -\frac{dV(r)}{dr} = -V'(r)$ then the central force \mathbf{f} is conservative.

Proof.

$$\begin{aligned}
 \mathbf{f} &= -V'(r) \hat{\mathbf{r}} = -V'(r) \nabla r \\
 &= -V'(r) \frac{\partial r}{\partial x} \hat{i} - V'(r) \frac{\partial r}{\partial y} \hat{j} - V'(r) \frac{\partial r}{\partial z} \hat{k} \\
 &= -\nabla V(r) \quad (\text{By the chain rule})
 \end{aligned}$$

□

So if we consider a system with a central force where the potential is directly related to the force, we will have a conservative system. Which is good, because when we send a satellite into orbit, we can safely say that we can depend on our friend, the conservation of energy, which makes calculations relatively straightforward.

2.4 Conservation of Angular Momentum

Another helpful property of a system is *Conservation of Angular Momentum*, especially if we are rotating masses about. However, to make sure we're all talking the same language, that is mathematics, here's the formal definition of angular momentum:

Definition. *If we have a particle with mass m , velocity \mathbf{v} and position \mathbf{r} , then that particle has an **angular momentum** given by:*

$$\mathbf{L} = m(\mathbf{r} \times \mathbf{v})$$

Or, if the particle has linear momentum \mathbf{p} :

$$\mathbf{L} = \mathbf{r} \times \mathbf{p}$$

We can also look at *torque*, which is defined by:

Definition.

$$\begin{aligned}
 \dot{\mathbf{L}} &= m(\dot{\mathbf{r}} \times \mathbf{v}) + m(\mathbf{r} \times \dot{\mathbf{v}}) \quad (\text{Total derivative}) \\
 &= m(\mathbf{v} \times \mathbf{v}) + m(\mathbf{r} \times \dot{\mathbf{v}}) \\
 &= \mathbf{r} \times m\dot{\mathbf{v}} \quad (\text{As } \mathbf{v} \times \mathbf{v} = 0) \\
 &= \mathbf{r} \times \mathbf{F}
 \end{aligned}$$

You've probably seen torque before, but maybe not defined as the change of angular momentum. "What does all this have to do with Conservation of Angular Momentum?" you may cry. Well, we want to show that angular momentum is conserved, and so we want to ensure that the *change* of angular momentum is zero. This indicates that showing that the torque is zero in a system with a central force could be a good way to go about proving it. So let's do it:

Theorem 2.3 (Conservation of Angular Momentum). *Given a central force \mathbf{F} , and :*

$$\mathbf{F} = -\nabla V(r) = -V'(r)\hat{\mathbf{r}} = \frac{-V'(r)}{r}\mathbf{r}$$

Then the change of momentum is zero.

Proof. By simple substitution:

$$\begin{aligned}\dot{\mathbf{L}} &= \mathbf{r} \times \mathbf{F} \\ &= \frac{-V'(r)}{r}(\mathbf{r} \times \mathbf{r}) \\ &= \mathbf{0}\end{aligned}$$

□

Another interesting result follows:

Example 2.4.1.

We look at the two dot products:

$$\begin{aligned}\mathbf{r} \cdot \mathbf{L} &= m\mathbf{r} \cdot (\mathbf{r} \times \mathbf{v}) = 0 \\ \mathbf{v} \cdot \mathbf{L} &= m\mathbf{v} \cdot (\mathbf{r} \times \mathbf{v}) = 0\end{aligned}$$

So \mathbf{r} and \mathbf{v} are perpendicular to \mathbf{L} . Thus the particle is constrained to move in the plane perpendicular to \mathbf{L} , and so a central force restricts rotational movement to a plane.

Chapter 3

Concepts in Advanced Mechanics

3.1 Generalised Coordinates

If we were shown a system and asked to analyse, we often follow through a few standard “problem-solving” steps. We would draw the diagram, assign coordinates and positive directions, write down relevant equations and pull it all together. One question rarely asked is: “Is this standard?” Let’s say you and I are both given the problem of a mass on a spring. Our diagrams will be the same (as we’re considering the same system) and our basic equations should be the same (physical laws would have to be standard). But let’s say you look at the position as being a value between $-a$ and a , where a is the maximum extension and 0 would be when the spring is at its natural length. But let’s say, for some mathematically perverse reason, I decided that the position was dictated by the angle the mass made with some fixed axis (say I had an axis up and to the right from the initial position of the mass). Our system should be moving in the same manner, and should give the same results. But how do we *know* it will. Maybe for some obscure reason, my reference system gives inconsistent results because of some intrinsic property that angles have. For mathematical rigour, we would like to have faith that such differences would not happen, or that we would know when we could avoid them.

This is where *generalised coordinates* come in. Let's look at the definition:

Definition. *The generalised coordinates of a system are the smallest possible number of variables required to describe the configuration of the system.*

Now we can deduce two things from this definition:

- Each generalised coordinate tells us something about the configuration of the system.
- A generalised coordinate must be independent of the other coordinates, unless we have constraints on the system (which we shall talk about in the next section).

These two points are important. The first point says that everything must be relevant. If you bring in a variable that depends on the age of your supervising professor, your results will be different to mine, which they shouldn't. The second point tells us to keep it simple. For example, if I wanted to know the configuration of a ruler on a plane, say, I wouldn't need to know the x -position of both ends of the ruler if I had the length and the y -positions of both ends, because I could figure out the position of one end, from the other, by a quick application of Pythagoras' Theorem.

A perceptive reader may have noticed there hasn't been any mention of time in any of the definitions of generalised coordinates. This is because it leads to a slightly tricky concept: time-dependent generalised coordinates. You see, when we give our generalised coordinates, they could be explicitly in terms of time, which then makes them time-dependent. However, if not, which shall be the case for this text, they are called time-independent and do not explicitly depend on a time parameter.

Which then leads directly onto the concept of generalised velocities and momenta. If I had a standard coordinate x , then normally we would consider the velocity in the x -direction to be $\frac{dx}{dt} = v_x$. Now if we make our x a generalised coordinate, the question is: Is velocity a generalised coordinate? This depends greatly on whether the coordinate is time-dependent or not. If it is, then the velocity $v_x(t)$ is simply the derivative with respect to time, and so the velocity $v_x(t)$ isn't mathematically independent of the position x . However, if the coordinate is independent of time, then the velocity *is* independent of the position, and so can be considered a generalised coordinate. Similarly we can consider momenta in much the same way as velocity, which is important later on with Hamilton's Equations.

It is important that you remember that we only deal with time-independent generalised coordinates here. Our formulations later on rely on this fact and the derivations are slightly different if we allow for time-dependent coordinates.

Another important concept to do with generalised coordinates is the notion of *degrees of freedom*. Quite naturally, the degrees of freedom of a system is the number of generalised coordinates a system can change. Think of it like this: I am a prisoner in a very restrictive prison. My cell is a very narrow but long corridor. It is so narrow I can't turn around, but can move back and forth if I want. I would have two degrees of positional freedom (one for position, one for velocity). Now the wardens have a change of heart and give me a room for me to move about in. The roof is too low for me to jump up and down, and because I'm shifty, they won't let me turn around. Now I have 4 degrees of freedom (two positional, two velocity). One day, it's my birthday and the wardens decide to be very nice and give me a go in a jumping castle. I can now move forwards and backwards (y), left or right (x), up or down (z), and I can even twist and turn in three different directions (twist, somersault and cartwheel). I can also change how fast I change position or direction, so I have my corresponding velocities as well. So now I have a whopping six times two – 12 degrees of freedom.

This idea is easy to apply to m -body systems. For example, if, in my previous example, there was another prisoner who had a birthday the same day as me, what would be the total degrees of freedom in our system? The degrees of freedom is a measure of how many generalised coordinates are required to specify the system. So I have 12 degrees degrees of freedom, and my friend has 12 as well. Naturally the system would be described at any time by the union of our coordinates, giving us 24 degrees of freedom.

Exercise 3.1.1.

What would be the degrees of freedom in our jumping castle system if I had m prisoners?

Exercise 3.1.2.

What is the number of degrees of freedom if I have n dimensions to move about in and m independent particles?

Answer.

$2nm$ degrees of freedom. m particles would each have n positional degrees of freedom plus n for velocities (or momenta). So $m \cdot (n + n) = 2nm$

One last point you should need to know is notation. Generalised coor-

dinates are usually denoted by q_i where $1 \leq i \leq n$, n the dimensions our system is in (for example, a mass-spring system is one-dimensional, whereas a mass-spring pendulum is in two-dimensions). Generalised velocities are usually denoted by \dot{q}_i , with i having the same limits as before. We use the dot notation to show it is a derivative, but we don't actually differentiate q_i to get it, because, as we said before, our generalised coordinates aren't explicitly in terms of t . Generalised momenta are given by p_i . Note that coordinates and momenta/velocities correspond with i . That is, \dot{q}_1 is the velocity in the q_1 direction. We also say our position (or configuration) is given by $r(q_1, \dots, q_n, t)$. So our position is dependent on time but our generalised coordinates aren't.

3.2 Constraints

One of the important things about generalised coordinates is that they are all independent of each other. Now what happens when we put constraints on our system? For example, let's say I have a goat in a grassy field. At any time it can have any x and y position, and so has 4 degrees of freedom¹. Now let's say we tie the goat to a post with a rigid pole of length l . The goat can only move in a circle now. However, my two coordinates aren't independent because we have the simple constraint equation:

$$x^2 + y^2 = l^2$$

Since l is a constant, y is always in terms of x and so given x , we know what y has to be. So now we only have 1 independent coordinate, x (or equivalently y), so we have only 2 degrees of freedom, instead of four.

Generally we say in a m -particle system with $2n$ degrees of freedom and k constraint equations, we have $2(n - k)m$ degrees of freedom.

Exercise 3.2.1.

How many degrees of freedom do the following systems have:

1. *Simple Pendulum*
2. *Simple Spherical Pendulum*
3. *A free particle in n -dimensional space*
4. *Your index finger*

¹We consider the goat to be a particle so we don't worry about rotations

Answer.

Convince yourself that they have degrees of freedom 2,4, $2n$ and 4 respectively. Note the last answer may be 6 if you let your finger wriggle sideways).

3.3 Generalised Forces

In the mechanics you've done before, we often look at work and forces. We can formulate an equivalent idea, but in relation to generalised coordinates. Suppose we have n particles with m degrees of freedom and generalised coordinates $q_i (1 \leq i \leq m)$. Now the rate at which we do work is given by:

$$\frac{dW}{dt} = \frac{dT}{dt} = \sum_{i=1}^n \mathbf{F}_i \cdot \mathbf{v}_i = \sum_{i=1}^n \mathbf{F}_i \cdot \dot{\mathbf{r}}_i$$

where \mathbf{F}_i is the vector force exerted on the i th particle, and \mathbf{r}_i is the position vector of the i th particle. Of course, this doesn't look too different from what we've done before and doesn't include our generalised coordinates at all. Which is what we look at next.

We say $\mathbf{r}_i = \mathbf{r}_i(q_j, t)$, and so give our position in terms of some of our generalised coordinates. To bring our new \mathbf{r} into the formulas above, we differentiate it giving:

$$\dot{\mathbf{r}}_i = \sum_{j=1}^m \left(\frac{\partial \mathbf{r}_i}{\partial q_j} \dot{q}_j + \frac{\partial \mathbf{r}_i}{\partial t} \right)$$

by total derivation.

So our rate of which work is done is now given by:

$$\begin{aligned} \frac{dW}{dt} = \frac{dT}{dt} &= \sum_{i=1}^n \mathbf{F}_i \cdot \dot{\mathbf{r}}_i \\ &= \sum_{j=1}^m Q_j \dot{q}_j + \sum_{i=1}^n \mathbf{F}_i \cdot \frac{\partial \mathbf{r}_i}{\partial t} \end{aligned}$$

where

$$Q_j = \sum_{i=1}^n \mathbf{F}_i \cdot \frac{\partial \mathbf{r}_i}{\partial q_j}$$

which we called the *generalised force*.

These act pretty much just like the forces you are used to, but in terms of our new generalised coordinates.

3.4 Conservative Systems

If you remember in the second chapter we were interested in conservative systems because it allowed us to use conservation of energy. Well, with our generalised coordinates construction, things aren't much different:

Definition. A system is called **conservative** if there exists a function $V = V(q_i, t)$ called the **potential** such that:

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

and

$$\sum_{i=1}^n \mathbf{F}_i \cdot \frac{\partial \mathbf{r}_i}{\partial t} = -\frac{\partial V}{\partial t}$$

Then

$$\frac{dT}{dt} = -\sum_{j=1}^m \frac{\partial V}{\partial q_j} \dot{q}_j$$

and

$$-\frac{\partial V}{\partial t} = -\frac{dV}{dt}$$

So $T + V = E$, where T is our usual kinetic energy ($T = \sum_{i=1}^n \frac{1}{2} m_i v_i^2$).

Note that we usually say that the potential doesn't change with time, and so $\frac{\partial V}{\partial t} = 0$

3.5 Equilibrium in Conservative Systems

Whenever we talk of equilibrium normally, we often consider the system to be in equilibrium when all the forces on the particles cancel each other out. However because generalised forces aren't quite the same as the forces you are used to, we formulate it differently:

Definition. If all the generalised forces in a system are zero, then the system is said to be in **equilibrium**.

A usual question to ask when analyzing systems at equilibrium is "What happens when I perturb (bump) the system?" We address this in our next theorem:

Theorem 3.1. If we displace a conservative system from equilibrium by an arbitrarily small amount, say $(\delta q_1, \dots, \delta q_m)$, then the change of work is zero.

Proof.

$$\begin{aligned}
 \delta W = \delta T &= -\delta V \\
 &= -\sum_j \frac{\partial V}{\partial q_j} \delta q_j \\
 &= -\sum_j 0 \cdot \delta q_j \quad (\text{because our gen. forces are zero}) \\
 &= 0
 \end{aligned}$$

□

So in perturbing a system at equilibrium, we don't change the energy. But the question is, "Is the equilibrium stable or not?" For example, I can have two systems in equilibrium, the first is a ball in a valley, and the second is a ball on a hill. If I nudge the first system, the ball will move a little, but return to the equilibrium point soon enough. However, if I nudge the second system, the ball will roll off the hill and never come back. The first system is said to be *stable* and the second *unstable*.

How do we formally look at this? Recalling our minima and maxima theory from calculus, we would say our equilibrium with respect to one of the generalised coordinates is stable if our potential is a local minima. Why? Think back to our stable "ball in a valley" example. If I perturb the ball, it will rise a little (because I bumped it) and then fall back to the bottom because it seeks to minimize potential (which is due to gravity). Obviously if the hill was made out of rubber, the ball would continue to drive downwards, in an attempt to minimize potential. But once it has reached the minimum potential, what can it do but sit there *in equilibrium*. Similarly for an unstable equilibrium, our equilibrium point will be a local maxima and the ball will go anywhere but back to the equilibrium because the potential is lower elsewhere.

Now you may have noted that I insisted on saying a *local* minima and maxima. This is because we assume we are only perturbing the system slightly. Back to our valley example, imagine our valley is now actually at the top of a hill. If we perturb the ball a little bit, it hasn't been moved enough to escape the valley and roll down the hill, to the *global* minima. This is very important because we really only want to worry about little perturbations, rather than arbitrary (and possibly large) ones.

So let's put that into our mathematical formalism:

Proposition 3.1. *In a conservative system, an equilibrium is said to be stable with respect to a generalised coordinate q_i if we have a local minimum*

in the potential. Namely, if

$$-\frac{\partial Q_j}{\partial q_i} = \frac{\partial^2 V}{\partial q_i \partial q_j} = \frac{\partial^2 V}{\partial q_j \partial q_i} = -\frac{\partial Q_i}{\partial q_j}$$

is positive for all j .

Chapter 4

Lagrange's Method of Dynamics

4.1 Overview of Lagrange's and Hamilton's Methods

When we are analyzing any dynamical system, what we are basically doing is formulating differential equations of motion. Every time you write $F = ma$, you're actually talking about the differential equation $m\ddot{x}$ being equal to some formulas to do with the forces involved. Then to get anything useful out of the system, like how it behaves, we have to solve these differential equations. When we solve dynamical systems using force diagrams and so forth, we go about it in a fairly messy and undirected way. Which is where Lagrange's and Hamilton's methods for solving dynamical equations come in. They obtain the differential equations and solve them in a very systematic and straightforward way.

Generally speaking, the methods of Lagrange and Hamilton lead us towards obtaining expressions for the kinetic and potential energies of the system. When we get these expressions, these methods usually give them in a differential form that is relatively straightforward to integrate and thus solve for the equations of motion.

4.2 Lagrange's Method

Both Lagrange's and Hamilton's methods utilise generalised coordinates and forces. When we speak of these methods, they are actually just equations in

a certain form. This form allows us to plug in our generalised forces to give us expressions related to the kinetic and potential energies of the system. So our goal in formulating these methods is to relate the generalised forces to the kinetic and potential energies in a form easy to analyze.

So firstly we need to show the relation between Q_i (the generalised force) and T the kinetic energy (note that we already know how the potential energy is related to the generalised forces).

Theorem 4.1. *We relate the generalised force to the kinetic energy of a system with particles $1 \leq j \leq n$ by:*

$$\begin{aligned} Q_i &= \frac{d}{dt} \left\{ \sum_j m_j \dot{\mathbf{r}}_j \cdot \frac{\partial \mathbf{r}_j}{\partial \dot{q}_i} \right\} - \sum_j m_j \dot{\mathbf{r}}_j \cdot \frac{d}{dt} \left(\frac{\partial \mathbf{r}_j}{\partial \dot{q}_i} \right) \\ &= \frac{d}{dt} \left\{ \frac{\partial T}{\partial \dot{q}_i} \right\} - \frac{\partial T}{\partial q_i} \end{aligned}$$

Proof. Now we know the position a particle is given by:

$$\mathbf{r}_i = \mathbf{r}(q_1, q_2, \dots, q_n, t)$$

(Note we only do the derivation for one particle, but for multiple the steps are exactly the same). To find the velocity, we find the derivative with respect to time:

$$\dot{\mathbf{r}} = \frac{d\mathbf{r}}{dt} = \sum_j \frac{\partial \mathbf{r}}{\partial q_j} \dot{q}_j + \frac{\partial \mathbf{r}}{\partial t} \quad (\text{By total differential})$$

If we take the partial derivative of this expression with respect to the derivative of a particular generalised coordinate, say q_i , we get:

$$\frac{\partial \dot{\mathbf{r}}}{\partial \dot{q}_i} = \frac{\partial \mathbf{r}}{\partial q_i} \quad (4.1)$$

We obviously get this because every other term hasn't got a \dot{q}_i term. We call this result the ***cancellation of dots***.

So lets look at the partial derivative of $\frac{d\mathbf{r}}{dt}$ with respect to our q_i to try to associate with what we found above:

$$\frac{\partial}{\partial q_i} \frac{d\mathbf{r}}{dt} = \frac{\partial \dot{\mathbf{r}}}{\partial q_i} = \sum_j \frac{\partial}{\partial q_i} \frac{\partial \mathbf{r}}{\partial q_j} \dot{q}_j + \frac{\partial^2 \mathbf{r}}{\partial q_i \partial t}$$

(If you're working through this, make sure you don't mistake \mathbf{r} for $\dot{\mathbf{r}}$ and vice versa) We look at how this expression changes with time, so we take the derivative with respect to time, giving:

$$\begin{aligned} \frac{d}{dt} \left(\frac{\partial \mathbf{r}}{\partial q_i} \right) &= \sum_j \frac{\partial}{\partial q_j} \left(\frac{\partial \mathbf{r}}{\partial q_i} \right) \frac{dq_j}{dt} + \frac{\partial}{\partial t} \left(\frac{\partial \mathbf{r}}{\partial q_i} \right) \\ &= \sum_j \frac{\partial^2 \mathbf{r}}{\partial q_i \partial q_j} \dot{q}_j + \frac{\partial^2 \mathbf{r}}{\partial t \partial q_i} \end{aligned}$$

This last expression can be recognized as the total time derivative of the derivative of \mathbf{r} with respect to some generalised coordinate q_i , probably easily expressed as:

$$\frac{d}{dt} \left(\frac{\partial \mathbf{r}}{\partial q_i} \right) = \sum_j \frac{\partial^2 \mathbf{r}}{\partial q_i \partial q_j} \dot{q}_j + \frac{\partial^2 \mathbf{r}}{\partial t \partial q_i} = \frac{d}{dt} \left(\frac{\partial \dot{\mathbf{r}}}{\partial q_i} \right) \quad (4.2)$$

After that you may wonder why we did that at all. Well Equations 4.1 and 4.2 are useful "trick" expressions we are to use below.

So we want to relate the generalised force to the kinetic energy of the system. From our formulation of generalised forces we can say:

$$\begin{aligned} Q_i &= \sum_j \mathbf{F}_j \cdot \mathbf{v}_j \\ &= \sum_j m_j \ddot{\mathbf{r}}_j \cdot \frac{\partial \mathbf{r}_j}{\partial q_i} \\ &= \frac{d}{dt} \left\{ \sum_j m_j \dot{\mathbf{r}}_j \frac{\partial \mathbf{r}_j}{\partial q_i} \right\} - \sum_j m_j \dot{\mathbf{r}}_j \cdot \frac{d}{dt} \left(\frac{\partial \mathbf{r}}{\partial q_i} \right) \end{aligned} \quad (4.3)$$

Now this expression is too tricky to deal with directly, so we consider the first and last summations separately, then bring them back together in the end.

(First summation)

$$\begin{aligned}
 \sum_j m_j \dot{\mathbf{r}}_j \frac{\partial \mathbf{r}_j}{\partial q_i} &= \sum_j m_j \dot{\mathbf{r}}_j \cdot \frac{\partial \dot{\mathbf{r}}_j}{\partial \dot{q}_i} \quad (\text{Using Equation 4.1}) \\
 &= \sum_j \frac{\partial}{\partial \dot{q}_i} \left(\frac{1}{2} m_j \dot{\mathbf{r}}_j \cdot \dot{\mathbf{r}}_j \right) \quad (\text{By the product rule}) \\
 &= \frac{\partial}{\partial \dot{q}_i} \left(\sum_j \frac{1}{2} m_j \dot{\mathbf{r}}_j \cdot \dot{\mathbf{r}}_j \right) \\
 &= \frac{\partial T}{\partial \dot{q}_i}
 \end{aligned}$$

Which is obviously related to the kinetic energy T .

(Second summation)

(Note this is a similar derivation, so if you're confident you can skip this)

$$\begin{aligned}
 &\sum_j m_j \dot{\mathbf{r}}_j \cdot \frac{d}{dt} \left(\frac{\partial \mathbf{r}_j}{\partial q_i} \right) \\
 &= \sum_j m_j \dot{\mathbf{r}}_j \cdot \frac{\partial \dot{\mathbf{r}}_j}{\partial \dot{q}_i} \quad (\text{Using Equation 4.2}) \\
 &= \sum_j \frac{\partial}{\partial \dot{q}_i} \left(\frac{1}{2} m_j \dot{\mathbf{r}}_j \cdot \dot{\mathbf{r}}_j \right) \\
 &= \frac{\partial T}{\partial \dot{q}_i}
 \end{aligned}$$

So then we get the desired result by substituting the two summations back into Equation 4.3. \square

4.3 Lagrange's Equation

What was the point of all that algebra? Well with a little more juggling, we shall obtain **Lagrange's Equation**, which is the goal of this chapter. So from Theorem 4.1, we can prove the following important theorem:

Theorem 4.2 (Lagrange's Equation). *Given a conservative system with generalised coordinates $q_i, i = 1 \dots m$, kinetic energy $T = T(q_i, \dot{q}_i, t)$ and potential energy $V = V(q_i, t)$, we can write (for all $i = 1 \dots m$):*

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) - \frac{\partial L}{\partial q_i} = 0 \quad (\text{Lagrange's Equation})$$

where $L = T - V$, called the **Lagrangian** of the system.

Proof. We know from Theorem 4.1 that:

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

Also because the system is conservative,

$$Q_i = -\frac{\partial V}{\partial q_i}$$

Also, by our definition of the potential energy, it is independent of all \dot{q}_i , and so:

$$\frac{\partial V}{\partial \dot{q}_i} = 0$$

Combining the first two equations we get:

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

We move $-\frac{\partial V}{\partial q_i}$ to the left hand side and since we can always subtract zero from an equation:

$$\begin{aligned} \frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_i} - 0 \right) - \frac{\partial T}{\partial q_i} + \frac{\partial V}{\partial q_i} &= 0 \\ \frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_i} - \frac{\partial V}{\partial \dot{q}_i} \right) - \frac{\partial T}{\partial q_i} - \frac{\partial V}{\partial q_i} &= 0 \\ \frac{d}{dt} \left(\frac{\partial(T - V)}{\partial \dot{q}_i} \right) - \frac{\partial(T - V)}{\partial q_i} &= 0 \end{aligned}$$

Then defining $T - V$ as L , we have the desired result. □

One of the important things you should note from this is that we get m second order differential equations, giving us a total order of $2m$. When looking at Hamilton's method, we shall find we have $2m$ first order differential equations, which means that Lagrange gives us fewer equations, but are more difficult to solve, whereas Hamilton gives us twice as many equations that are more easy to solve. Regardless, using Lagrange's equation is far easier than previous methods, as you will see in the following examples:

Chapter 5

Hamilton's Method of Dynamics

As we mentioned before, Lagrange's method gives us m second-order differential equations. Which is fine, but second-order equations aren't the easiest we could deal with. A more useful formulation would involve only first order equations, and because we can't lose information, it would have to be of the same total order (order times number of equations). This is where Hamilton's method comes in. It gives us $2m$ first order equations to describe the system.

5.1 Hamilton's Method

Here we shall derive Hamilton's Equations from the Lagrangian giving generalised momenta (which is like standard momentum, but for generalised coordinates).

Theorem 5.1 (Hamilton's Equations). *Given a conservative system of n particles with generalised coordinates q_1, \dots, q_m and corresponding generalised momenta*

$$p_j = \frac{\partial L}{\partial \dot{q}_j} \quad j = 1 \dots m$$

*we can obtain **Hamilton's Equations**:*

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q_i}, \quad \frac{dH}{dt} = -\frac{\partial L}{\partial t} = \frac{\partial H}{\partial t}$$

where $L = L(q_1, \dots, q_m, \dot{q}_1, \dots, \dot{q}_m, t)$ and

$$H = \sum_{i=1}^m p_i \dot{q}_i - L$$

called the **Hamiltonian** of the system.

Proof. We can see that:

$$dL = \sum_{i=1}^m \frac{\partial L}{\partial q_i} dq_i + \sum_{i=1}^m \frac{\partial L}{\partial \dot{q}_i} d\dot{q}_i + \frac{\partial L}{\partial t} dt$$

Now

$$dH = \sum_{i=1}^m p_i d\dot{q}_i + \sum_{i=1}^m \dot{q}_i dp_i - \sum_{i=1}^m \frac{\partial L}{\partial q_i} dq_i - \sum_{i=1}^m \frac{\partial L}{\partial \dot{q}_i} d\dot{q}_i - \frac{\partial L}{\partial t} dt$$

However, we recognise that:

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

So now

$$dH = - \sum_{i=1}^m \dot{p}_i dq_i + \sum_{i=1}^m \dot{q}_i dp_i - \frac{\partial L}{\partial t} dt$$

This suggests that changes in our Hamiltonian depends only on our generalised coordinates, their corresponding momenta and possibly time, namely:

$$H = H(q_i, p_i, t)$$

So then we say

$$dH = \sum_{i=1}^m \frac{\partial H}{\partial p_i} dp_i + \sum_{i=1}^m \frac{\partial H}{\partial q_i} dq_i + \frac{\partial H}{\partial t} dt$$

giving us:

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = - \frac{\partial H}{\partial q_i}, \quad \frac{dH}{dt} = - \frac{\partial L}{\partial t} = \frac{\partial H}{\partial t}$$

(Justification of the assumption of $H = H(q_i, p_i, t)$)

We want to show that H , and therefore L , are functions of q_i, p_i and possibly t . Namely, we want to show \dot{q}_i is a function of q_i, p_i and maybe t .

We assume that our position vector of our j th particle is given only by the generalised coordinates and time, namely:

$$\mathbf{r}_j = \mathbf{r}_j(q_i, t)$$

Therefore we can say the time derivative of \mathbf{r}_j is given by:

$$\dot{\mathbf{r}}_j = \sum_{i=1}^m \frac{\partial \mathbf{r}_j}{\partial q_i} \dot{q}_i + \frac{\partial \mathbf{r}_j}{\partial t}$$

by the total differential.

Now we know the kinetic energy of the system and we try to work our above assumption (and formulae) into it:

$$\begin{aligned} T &= \sum_{j=1}^n \frac{1}{2} m_j \dot{\mathbf{r}}_j \cdot \dot{\mathbf{r}}_j \\ &= \sum_{j=1}^n \frac{1}{2} m_j \left(\sum_{k=1}^m \frac{\partial \mathbf{r}_j}{\partial q_k} \dot{q}_k + \frac{\partial \mathbf{r}_j}{\partial t} \right) \cdot \left(\sum_{l=1}^m \frac{\partial \mathbf{r}_j}{\partial q_l} \dot{q}_l + \frac{\partial \mathbf{r}_j}{\partial t} \right) \\ &= \sum_{k=1}^m \sum_{l=1}^m \left(\sum_{j=1}^n \frac{1}{2} m_j \frac{\partial \mathbf{r}_j}{\partial q_k} \cdot \frac{\partial \mathbf{r}_j}{\partial q_l} \right) \dot{q}_k \dot{q}_l + \sum_{k=1}^m \left(\sum_{j=1}^n m_j \frac{\partial \mathbf{r}_j}{\partial q_k} \cdot \frac{\partial \mathbf{r}_j}{\partial t} \right) \dot{q}_k \\ &\quad + \sum_{j=1}^n \frac{1}{2} m_j \frac{\partial \mathbf{r}_j}{\partial t} \cdot \frac{\partial \mathbf{r}_j}{\partial t} \end{aligned}$$

Now that horrible mess has the nicer form:

$$T = \sum_{k=1}^m \sum_{l=1}^m a_{kl} \dot{q}_k \dot{q}_l + \sum_{k=1}^m b_k \dot{q}_k + c \quad (5.1)$$

where we make a few space-saving substitutions. This is because we can precompute them quite easily, allowing us to get to the meat of the equations without dragging about unnecessary baggage. a_{kl} is a symmetric $m \times m$ matrix¹ which just has all the $\sum_{j=1}^n \frac{1}{2} m_j \frac{\partial \mathbf{r}_j}{\partial q_k} \cdot \frac{\partial \mathbf{r}_j}{\partial q_l}$ values. b_k is another space saver, holding the values of $\sum_{j=1}^n m_j \frac{\partial \mathbf{r}_j}{\partial q_k} \cdot \frac{\partial \mathbf{r}_j}{\partial t}$. Similarly for c , which is just

¹so $a_{kl} = a_{lk}$

$\sum_{j=1}^n \frac{1}{2} m_j \frac{\partial \mathbf{r}_j}{\partial t} \cdot \frac{\partial \mathbf{r}_j}{\partial t}$. Now way back in Section 3.4 (on page 22) we said that the potential energy V is a function dependant solely on the generalised coordinates q_i and possibly time. Basically, $V = V(q_i; t)$.

Given this, we know from its definition the form of p_i , and so:

$$\begin{aligned} p_i = \frac{\partial L}{\partial \dot{q}_i} &= \frac{\partial T - V}{\partial \dot{q}_i} \\ &= \frac{\partial T}{\partial \dot{q}_i} - \underbrace{\frac{\partial V}{\partial \dot{q}_i}}_0 \\ &= \frac{\partial T}{\partial \dot{q}_i} \\ &= 2 \sum_{k=1}^m a_{ik} \dot{q}_k + b_i \end{aligned} \tag{5.2}$$

By rearranging:

$$\dot{q}_j = \frac{1}{2} \sum_{i=1}^m (a^{-1})_{ji} (p_i - b_i)$$

This allows us to write every instance of \dot{q}_i as a function of q_i and p_i . Therefore Equation 5.1 can be written solely as a function of q_i and p_i (and maybe t), and thus L and H are functions of q_i , p_i and possibly t . \square

5.2 Derivation of Lagrangian from the Hamiltonian

(Note: This section is optional and for interest only)

We saw above how we could use the Lagrangian of the system to derive the Hamiltonian, and Hamilton's equations. The question which you may ask is whether the process works in reverse; namely, can we derive the Lagrangian from the Hamiltonian? It turns out that we can. Others may ask, "Why would anyone actually care?" Well, dynamics has been formed on a strong mathematical and physical basis, and so rigour is essential. Additionally, in modern dynamics, many people regard the Hamiltonian to be a more base concept than the Lagrangian, so they would actually start from here rather than the way we did it. Let's look at the derivation:

Imagine that we are given a function $H(q_i, p_i, t)$, and are told that the motion of the system satisfies "Hamilton's Equations":

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

Then we are asked whether we can find a function $L(q_i, \dot{q}_i, t)$ such that it satisfies the equations:

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

(which is Lagrange's equations from before)

Using a similar approach to how we found Equation 5.2, but in reverse, we can find p_i in terms of q_i, \dot{q}_i and t . We then write:

$$L = \sum_{i=1}^m \dot{q}_i p_i - H$$

and replace the p_i , to give us L in terms of q_i, \dot{q}_i and t .

Exercise 5.2.1.

Find the explicit form of p_i above in terms of q_i, \dot{q}_i and t .

5.3 Conservation of Energy

If you recall in Chapters 2 and 3.4, we looked at conservative systems, and recognized that conservation of energy is an important property in dynamics. We defined a conservative system as a system that had a potential energy V . Now under our new Hamiltonian (and Lagrangian) formulation of dynamics, it is possible to look at systems more generally, and consider the entire system as being defined by our Hamiltonian (or Lagrangian). The information present in V is now embedded in H or L , and is no longer required (although we can, in principle, use it to give us H and L). So now we redefine the concept of a conservative system in relation to the Hamiltonian or Lagrangian.

Theorem 5.2. *A system defined by Lagrangian $L(q_i, \dot{q}_i, t)$ or Hamiltonian $H(q_i, p_i, t)$ is called **conservative** if:*

$$\frac{\partial L}{\partial t} = \frac{\partial H}{\partial t} = 0$$

Proof. Assume \mathbf{r}_j is only a function of the generalised coordinates q_i , and so

$$\frac{\partial \mathbf{r}_j}{\partial t} = 0$$

Now b_k and c in Equation 5.1 both depend on $\frac{\partial \mathbf{r}}{\partial t}$. But this term is zero, so both b_k and c are zero, leaving us with:

$$T = \sum_{k=1}^m \sum_{l=1}^m a_{kl} \dot{q}_k \dot{q}_l$$

Now Equation 5.2 becomes:

$$p_i = \frac{\partial L}{\partial \dot{q}_i} = \frac{\partial T}{\partial \dot{q}_i} = 2 \sum_{k=1}^m a_{ik} \dot{q}_k$$

as b_k depends on $\frac{\partial \mathbf{r}_j}{\partial t} = 0$.

Now if we multiply each p_i by its corresponding \dot{q}_i and sum over all i , we get:

$$\sum_i p_i \dot{q}_i = 2 \sum_{i=1}^m \sum_{k=1}^m a_{ik} \dot{q}_k \dot{q}_i$$

which we recognize (from above) to be $2T$.

But we defined the Hamiltonian to be $H = \sum_i p_i \dot{q}_i - L$ and so now

$$\begin{aligned} H &= \sum_i p_i \dot{q}_i - L \\ &= 2T - (T - V) \\ &= T + V \end{aligned}$$

So our Hamiltonian is equivalent to the total energy of the system, and because $\frac{\partial H}{\partial t} = 0$, the total energy won't change, and thus the system is conservative.

□

Chapter 6

Calculus of Variations

6.1 Some History and the Euler-Lagrange Equation

Sometimes when looking at a system, we can have constraints and coordinates set up, but we want to know what path the system takes when we maximise or minimise something. A classic example of this is *Dido's Problem*. Apparently Dido was promised all the land she could enclose with the hide of a bull. Being the resourceful girl she was, she cut the hide into thin strips, making a long curve in which she made into a semicircle around the area that later became Carthage. Now, mathematically Dido's Problem is to find the maximum area enclosed by a curve of length L . It turns out that a semicircle is the best answer, but how can we prove it? This is where the *Calculus of Variations* comes in. Some results obtainable from this method were actually solved by Ancient Greeks, like Dido, amongst others. However, in the 18th century, the master of mathematics, Euler formulated a general approach, quickly followed up by Lagrange. The general approach follows:

Theorem 6.1 (Euler-Lagrange Equation). *Given a known function $F(y, y', x)$ where $y = y(x)$, then we can find a y such that*

$$I(y) = \int_{x_1}^{x_2} F(y, y', x) dx$$

*(called the **functional**), has a minimum or maximum value under the boundary conditions $y(x_1) = y_1$ and $y(x_2) = y_2$, if F satisfies the **Euler-Lagrange***

Equation:

$$\frac{\partial F}{\partial y} - \frac{d}{dx} \left(\frac{\partial F}{\partial y'} \right) = 0$$

Proof. The way we approach this is by imagining we have y (even though we don't), and then making a small variation on the line. We know that $I(y)$ is a maximum or minimum under the boundary conditions, and so $I(y)$ is an extremal curve, and thus any variations on the line should move away from this extreme (consider it to be like a critical point of a function).

Imagining we have y , we make a variation on it by:

$$y \rightarrow y + \delta y = y + \epsilon \eta(x)$$

such that $\delta y = \epsilon \eta(x)$ is the variation, where ϵ is a small parameter, independent of x and $\eta(x)$ is an arbitrary function such that $\eta(x_1) = \eta(x_2) = 0$. This is so we keep the end-points of our new y fixed, but make variations wherever we want along the line.

Now we look at the change in the functional $I(y)$ after this variation of y , which is the changed functional minus the original functional:

$$I(y + \delta y) - I(y) = I(y + \epsilon \eta(x)) - I(y)$$

$I(y)$ is the extremal curve, and thus $I(y + \epsilon \eta(x))$ becomes an extremal curve when $\epsilon = 0$. So necessarily, we have the condition:

$$\left. \frac{d}{d\epsilon} I(y + \epsilon \eta(x)) \right|_{\epsilon=0} = 0$$

which means, for any $\eta(x)$, when we look at the change of ϵ and then set it to zero, our function $I(y + \epsilon \eta(x))$ must have a critical point, and so the derivative must be zero.

So let's work out our new condition in terms of F :

$$\begin{aligned} 0 &= \left. \frac{d}{d\epsilon} I(y + \epsilon \eta(x)) \right|_{\epsilon=0} \\ &= \int_{x_1}^{x_2} \left. \frac{d}{d\epsilon} F(y + \epsilon \eta(x), y' + \epsilon \eta'(x), x) \right|_{\epsilon=0} dx \end{aligned}$$

If we make the substitutions:

$$\begin{aligned} u &= y + \epsilon \eta(x) & \Rightarrow & \frac{du}{d\epsilon} = \eta \\ v &= y' + \epsilon \eta'(x) & \Rightarrow & \frac{dv}{d\epsilon} = \eta' \end{aligned}$$

Then by the chain rule (and total differential of F with respect to ϵ):

$$\begin{aligned}
 & \left. \frac{d}{d\epsilon} F(u, v, x) \right|_{\epsilon=0} \\
 = & \left(\frac{\partial F}{\partial u} \frac{du}{d\epsilon} + \frac{\partial F}{\partial v} \frac{dv}{d\epsilon} \right) \Big|_{\epsilon=0} \\
 = & \left(\frac{\partial F}{\partial u} \eta + \frac{\partial F}{\partial v} \eta' \right) \Big|_{\epsilon=0} \\
 = & \frac{\partial F}{\partial u} \eta + \frac{\partial F}{\partial v} \eta' \quad (\text{As previous line is independent of } \epsilon)
 \end{aligned}$$

So now we substitute this back into our integral above:

$$\int_{x_1}^{x_2} \left. \frac{d}{d\epsilon} F(y + \epsilon\eta(x), y' + \epsilon\eta'(x), x) \right|_{\epsilon=0} dx = \int_{x_1}^{x_2} \left(\frac{\partial F}{\partial u} \eta + \frac{\partial F}{\partial v} \eta' \right) dx$$

By the chain rule and our substitutions for u and v , we have:

$$\begin{aligned}
 \frac{\partial F}{\partial u} &= \frac{\partial F}{\partial y} \frac{\partial y}{\partial u} = \frac{\partial F}{\partial y} \cdot 1 = \frac{\partial F}{\partial y} \\
 \frac{\partial F}{\partial v} &= \frac{\partial F}{\partial y'} \frac{\partial y'}{\partial v} = \frac{\partial F}{\partial y'} \cdot 1 = \frac{\partial F}{\partial y'}
 \end{aligned}$$

Now by the product rule:

$$\frac{d}{dx} \left(\frac{\partial F}{\partial y'} \eta \right) = \frac{d}{dx} \left(\frac{\partial F}{\partial y'} \right) \eta + \frac{\partial F}{\partial y'} \eta'$$

Which we can rearrange to give:

$$\frac{\partial F}{\partial y'} \eta' = -\frac{d}{dx} \left(\frac{\partial F}{\partial y'} \right) \eta + \frac{d}{dx} \left(\frac{\partial F}{\partial y'} \eta \right)$$

Combining all this information and substituting it into our integral, we get:

$$\begin{aligned}
 & \int_{x_1}^{x_2} \left(\frac{\partial F}{\partial u} \eta + \frac{\partial F}{\partial v} \eta' \right) dx \\
 = & \int_{x_1}^{x_2} \left(\frac{\partial F}{\partial y} \eta - \frac{d}{dx} \left(\frac{\partial F}{\partial y'} \eta \right) \right) dx + \int_{x_1}^{x_2} \frac{d}{dx} \left(\frac{\partial F}{\partial y'} \eta \right) dx \\
 = & \int_{x_1}^{x_2} \eta \left(\frac{\partial F}{\partial y} - \frac{d}{dx} \left(\frac{\partial F}{\partial y'} \right) \right) dx + \underbrace{\left[\eta \frac{\partial F}{\partial y'} \right]_{x_1}^{x_2}}_0
 \end{aligned}$$

The last term goes to zero because of the way we defined $\eta(x)$ to act on the boundaries. So now we have our above integral, which must equal to zero, because of the necessary condition we established above. So we have:

$$\int_{x_1}^{x_2} \eta \left(\frac{\partial F}{\partial y} - \frac{d}{dx} \left(\frac{\partial F}{\partial y'} \right) \right) dx = 0 \quad \forall \eta(x)$$

But since this holds for all $\eta(x)$, including non-zero $\eta(x)$, we must have:

$$\frac{\partial F}{\partial y} - \frac{d}{dx} \left(\frac{\partial F}{\partial y'} \right) = 0$$

Which is the Euler-Lagrange equation we wanted. \square

6.2 Extending Euler-Lagrange

6.2.1 Generalisation to n dimensions

You may have noticed that the previous result was only for a F dependent on only one function. We can extend the result for systems with $F(y_1, \dots, y_n, y'_1, \dots, y'_n, x)$, where $y_j(x)$ are n independent functions of x whose values at x_1 and x_2 are specified and we have the functional:

$$I(y_1, \dots, y_n) = \int_{x_1}^{x_2} F(y_1, \dots, y_n, y'_1, \dots, y'_n, x) dx$$

The functional will have a minimum or maximum only if:

$$\frac{d}{dx} \left(\frac{\partial F}{\partial y'_i} \right) - \frac{\partial F}{\partial y_i} = 0 \quad \forall i = 1 \dots n$$

6.2.2 F independent of x

If we have a special situation where F depends only on y, y' , and is independent of x , then we can derive a useful equality.

Theorem 6.2. For $F(y, y')$, where $y(x)$ and $y' = \frac{dy}{dx}$, then we have the specialised Euler-Lagrange equation:

$$F - y' \frac{\partial F}{\partial y'} = k, \quad (k \text{ a constant})$$

Proof. Given the Euler-Lagrange equation, we have the following chain of equalities:

$$\begin{aligned}
 0 &= y' \left[\frac{\partial F}{\partial y} - \frac{d}{dx} \left(\frac{\partial F}{\partial y'} \right) \right] \\
 &= y' \frac{\partial F}{\partial y} - \left[\frac{d}{dx} \left(y' \frac{\partial F}{\partial y'} \right) - y'' \frac{\partial F}{\partial y'} \right] \\
 &\quad \text{(This trick is due to the product rule on } \frac{d}{dx} \left(y' \frac{\partial F}{\partial y'} \right) \text{)} \\
 &= y' \frac{\partial F}{\partial y} + y'' \frac{\partial F}{\partial y'} - \frac{d}{dx} \left(y' \frac{\partial F}{\partial y'} \right) \\
 &= \frac{dF}{dx} - \frac{d}{dx} \left(y' \frac{\partial F}{\partial y'} \right) \quad \text{(Product rule in reverse)} \\
 0 &= \frac{d}{dx} \left[F - y' \frac{\partial F}{\partial y'} \right]
 \end{aligned}$$

Therefore, by integrating, we get:

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

□

6.3 Calculus of Variations at work

Here we shall present a few classic examples of applications of the calculus of variations.

6.3.1 Brachistochrone Problem

The first thing people say when shown this example is: “What in the world is a brachisto-thingy?” Well a brachistochrone is the path taken by a bead sliding along a frictionless wire between two points if it minimises the time it takes to get there. The brachistochrone problem then basically asks, “If I were to shape a wire, what is the shape that minimises the time for the bead to slide along it?”

Anyway, Jacob Bernoulli solved this long-standing problem in 1696 by what turned out to be an application of the Calculus of Variations.

Let’s solve it:

We assume we have a conservative system, we have the conservation of energy:

$$T + V = \frac{1}{2}mv^2 - mgy = E$$

Now, at the start, the particle is at rest, and so energy is zero. Therefore,

$$\frac{1}{2}mv^2 = mgy \quad \Rightarrow \quad v^2 = 2gy$$

We assume velocity is always positive, so now

$$\sqrt{2gy} = v = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}$$

(the last part is by Pythagoras' Theorem with the x and y velocities)

Now:

$$\begin{aligned} & \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2} \\ &= \sqrt{1 + \left(\frac{dy}{dx}\right)^2} \frac{dx}{dt} \\ &= \sqrt{1 + y'^2} \frac{dx}{dt} \end{aligned}$$

Therefore:

$$\sqrt{2gy} = \sqrt{1 + y'^2} \frac{dx}{dt} \quad \Rightarrow \quad \frac{dt}{dx} = \frac{\sqrt{1 + y'^2}}{\sqrt{2gy}}$$

By integrating with respect to x , we can find the time taken,

$$t = \int_0^a F(y, y') dx, \quad \text{where} \quad F(y, y') = \sqrt{\frac{1 + y'^2}{2gy}}$$

By the specialised Euler-Lagrange equation, we know that we can minimise time, and F satisfies:

$$F - y' \frac{\partial F}{\partial y'} = c$$

Substituting:

$$\sqrt{\frac{1 + y'^2}{2gy}} - \frac{y'^2}{\sqrt{(1 + y'^2)2gy}} = c$$

Simplifying:

$$\frac{1}{\sqrt{(1+y'^2)2gy}} = c$$

Rearranging:

$$y(1+y'^2) = 2c$$

$$y'^2 = \frac{2c}{y} - 1 = \frac{2c-y}{y}$$

Now on physical grounds, we say that $y' > 0$ and so therefore,

$$y' = \sqrt{\frac{2c-y}{y}}$$

$$\therefore dx = \sqrt{\frac{y}{2c-y}} dy$$

To solve this, we use a method often used to solve tricky integrations, namely to substitute in a formula that relieves all the problems. In this case, we substitute $y = 2c \sin^2 \theta$, giving:

$$\therefore dy = 4c \sin \theta \cos \theta d\theta$$

$$\text{Then } dx = \sqrt{\frac{2c \sin^2 \theta}{2c \cos^2 \theta}} 4c \sin \theta \cos \theta d\theta$$

$$dx = 4c \sin^2 \theta d\theta$$

$$= 2c(1 - \cos 2\theta) d\theta \quad (\text{By trigonometric identity})$$

$$(\text{Integrating}) x = 2c\left(\theta - \frac{1}{2} \sin 2\theta\right) \quad (\text{provided } x = 0 \text{ at } \theta = 0)$$

We make the last assumption based on how we defined our system (as in, it goes from $x = 0$ to $x = x_1$). Thus now we can say:

$$x = c(2\theta - \sin 2\theta)$$

$$y = 2c \sin^2 \theta = c(1 - \cos 2\theta) \quad (\text{By trig. identity})$$

By making the substitution $\phi = 2\theta$, we now have the parametric equation:

$$x = c(\phi - \sin \phi)$$

$$y = c(1 - \cos \phi)$$

Thus given an endpoint, we can solve for c and then find our brachistochrone.

6.3.2 Snell's Law

In elementary physics, you probably met **Snell's Law**. It basically tells us how much light bends when going from one refractive medium to another. Let's derive it!

Example 6.3.1 (Snell's Law).

A beam of light moves through a medium from the origin to a point (a, b) . Find a path so that time taken is least.

Answer.

Since we're talking about finding optimal paths, we know we're talking about using Calculus of Variations. So let's try to get our functional, the thing we want to minimise, which happens to be time.

Let $c(x)$ be the velocity of light so that:

$$\begin{aligned} c(x) &= \sqrt{x'^2 + y'^2} \\ &= \sqrt{1 + y'^2} \frac{dx}{dt} \end{aligned}$$

Therefore

$$\frac{dt}{dx} = \frac{\sqrt{1 + y'^2}}{c(x)}$$

So, by integrating, the time taken is given by:

$$t = \int_0^a \frac{\sqrt{1 + y'^2}}{c(x)} dx$$

Since F , the function inside our functional's integral, is just a function of y' and x , we can find the minimum time by the Euler-Lagrange eqn:

$$0 = \frac{\partial F}{\partial y'} - \frac{d}{dx} \left(\frac{\partial F}{\partial y'} \right)$$

Now we know that the speed of light $c(x)$ is a constant in a medium. Therefore our second term is zero, and we get:

$$\frac{\partial F}{\partial y'} = 0$$

Now by doing the partial derivatives, we get:

$$\frac{y'}{c(x)\sqrt{1 + y'^2}} = 0$$

If we substitute $\tan \phi$ in the expression for y' then with a little manipulation, we get:

$$\frac{\sin \phi}{c(x)} = 0$$

So regardless of the system, our equation above holds. But $c(x)$ is a constant of the medium, and so we can say:

$$\frac{\sin \phi_1}{c_1} = 0 = \frac{\sin \phi_2}{c_2} \Rightarrow \frac{\sin \phi_1}{\sin \phi_2} = \frac{c_1}{c_2}$$

which is Snell's Law.

Now you may ask, "We went through all that trouble to find the *path* the light took, but we didn't end up talking about it. What gives?" Well it turns out that if you look back a few steps, we have the equation:

$$\frac{y'}{c(x)\sqrt{1+y'^2}} = 0$$

Since the fraction equals zero, we have two things that could be going on: the top could be zero, or the bottom could be infinity. Now we know (from measurements) that the speed of light isn't infinite (plus Einstein may have a few points to talk to you about if it were), which leaves the possibility that $\sqrt{1+y'^2}$ is infinite. But if it is, then the top is infinite too, of about the same magnitude, not giving us zero. So the only alternative is for y' to be zero. Now what does this mean? When we differentiate y with respect to time, it gives us zero. So y is constant with time. So looking back to our equation for time, we can see that it depends only on how c relates to x . Say we were in one medium only (that is, we don't change), $c(x)$ is a constant, and so our time is given by:

$$t = \frac{a}{c} \quad \text{which is just} \quad \text{time} = \frac{\text{distance}}{\text{velocity}}$$

which is kind of obvious. With two mediums, we know our change in y is constant, so with changing mediums, our velocity in the x -direction is the only thing that changes. From this you can see why our light travels along a straight line, until it bends when changing media.

6.4 Principle of Least Action

In investigating dynamics, we often strive to look for basic concepts, things which are at the heart of our theory and encompass it all. When you are

first presented dynamics, we often consider energy to be a fundamental concept. However, there turns out to be a *more* fundamental concept than energy: action. The action of a system is related to how the system follows a particular trajectory. It gives a reason why a projectile travels in a parabola, instead of doing a few corkscrews in between.

Now action is given by the **action functional**:

$$A = \int_{t_1}^{t_2} L(q_i, \dot{q}_i, t) dt$$

where L is your usual Lagrangian. This action functional looks strikingly like the functional I in the Calculus of Variations formulation.

Which is where we come to the **Principle of Least Action**.

Theorem 6.3 (Principle of Least Action). *A system will evolve according to the trajectory that minimises the action functional A .*

This kind of makes sense, as the action functional looks like the functional that is minimised in the Calculus of Variations formulation, and the functional depends on the Lagrangian, which dictates trajectory and energy.

Example 6.4.1 (Projectile Motion).

Let's consider a particle launched from the origin with a particular velocity and angle. From this angle and velocity, we can work out our horizontal and vertical initial velocities. Now the question is, what path does it travel? By the Principle of Least Action, we have the functional:

$$A = \int_{t_1}^{t_2} L(q_i, \dot{q}_i, t) dt$$

We want to minimise this functional, and so we can use the Euler-Lagrange Equation. But before we can do this, we need the Lagrangian of the system.

We assume there is no horizontal forces on the particle (like wind or air resistance). The only force acting is the gravitation force, which gives a potential of $-mgy$

So our Lagrangian is given by:

$$L = T - V = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}m\dot{y}^2 + mgy$$

Now the Principle of Least Action states that the system will evolve according to the trajectory that minimises the action functional A . Therefore

the Euler-Lagrange equations (for two generalised coordinates in this case) must hold. Therefore:

$$\begin{aligned}\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}} \right) - \frac{\partial L}{\partial x} &= 0 \\ \frac{d}{dt} (m\dot{x}) - 0 &= \\ m\ddot{x} &= \\ \therefore \ddot{x} &= 0\end{aligned}$$

This is true by the setup of the system (no horizontal forces). Now let's look at the Euler-Lagrange equations for y :

$$\begin{aligned}\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{y}} \right) - \frac{\partial L}{\partial y} &= 0 \\ \frac{d}{dt} (m\dot{y}) + mg &= \\ m\ddot{y} + mg &= \\ \therefore \ddot{y} &= -g\end{aligned}$$

This we can also see is consistent with our system (and experience).

So what about our trajectory? It's easy to see (by integrating twice), that we get:

$$\begin{aligned}x &= v_x t \\ y &= y_0 + v_y t - gt^2\end{aligned}$$

These two parametric equations will trace out a parabola, as per our experience with basic mechanics.

NOTE:

You may wonder how we justify the labelling of constants in our equations above. With x it's obvious as the constant we get from integrating \ddot{x} is equal to \dot{x} , the velocity in x . Similarly for y . The y_0 is an initial condition (we can find it by putting $t = 0$), which is equal to zero by our assumption that the particle is launched from the origin.

Chapter 7

Phase Spaces and Canonical Transformations

7.1 Phase Space

It is neat that we can use such powerful mathematical techniques to analyse these systems. However, given some complicated formula, it's reasonably tricky to say qualitatively what the system is doing. Everyone loves to visualise what their systems are doing, which is why we use *phase spaces*.

So what is a phase space? Well a phase space is basically just a plot of all the generalised coordinates versus the generalised momenta¹. So for a system with n degrees of freedom, we shall have a $2n$ -dimensional phase space. If our system is complicated, since we won't have, say, 6-dimensional graph paper, it is sufficient to plot individual graphs of a particular q_i versus its corresponding p_i .

In looking at phase spaces, we shall assume we only have conservative systems, for simplicity. So:

$$H = H(q, p) \quad \frac{\partial H}{\partial t} = 0$$

and H is a constant of the motion. We can look at arbitrary systems if we like, but conservative systems are easier to understand the first time you use phase spaces. Now Hamilton's equations give:

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

¹We are only concerned with Hamiltonian dynamics here. Lagrangians work exactly the same, but most people prefer the Hamiltonian method.

Since H is a constant of the motion, \dot{q} and \dot{p} give the change in each variable, and thus giving us a direction on our phase portrait². Now for a given H value, our equations will trace out a line as time evolves. Which is pretty cool. Each of the individual lines are called **stream lines** and define a particular configuration and evolution of the system. So if I have a system with given energy, it will move along a particular line. If I instead choose a bigger energy, it will follow a stream line further out perhaps. This is handy for a system with some equilibrium envelope, such that for certain energies, the particle acts similarly for all energies below it (for example, a ball in a valley). However, if we exceed this value, the system evolves differently (because the ball has enough momentum to get it out of the valley and thus not just oscillate inside of it). Let's look at some simple examples:

Example 7.1.1.

Given a particle of mass $m = 1$ moving along a straight line with no forces applied, then our Hamiltonian is given by: $H = \frac{1}{2}p^2$. If p and q are rectangular Cartesian coordinates, then our stream lines are parallel to the q -axis, moving towards positive infinity for $0 < p$ and moving towards negative infinity for $p < 0$. Our motion depends on how our initial momentum, which selects a particular stream line. If $p = 0$ then our stream line is just a point on the origin (because it's not moving). Otherwise it moves with a velocity proportional to p .

Example 7.1.2.

Our old favourite, the simple harmonic oscillator, will have a Hamiltonian:

$$H = \frac{1}{2m}p^2 + \frac{1}{2}kq^2$$

Since H is constant, you should be able to recognise that the above equation describes ellipses for our stream lines (the radii given by our H value). We need to find in which direction, clockwise or counter-clockwise, the stream lines move. From Hamilton's equations we have:

$$\frac{dq}{dt} = \frac{p}{m} \quad \frac{dp}{dt} = -kq$$

To figure out the direction, we look at when $p = q$, both positive. Then the change in q is positive, and so the upper right quadrant moves to the right (also the change in p is negative, meaning the particle moves down as well).

²phase space and phase portrait are generally the same thing. Phase space is more of the mathematical idea and portrait the actual graph.

This gives us the direction (clockwise) of our particle and so we have our phase space.

7.2 Canonical Transformations

If you recall way back when we defined generalised coordinates and had the debate over which system of coordinates to use. Generally our theory is invariant (as in, it doesn't change) under coordinate transformations. For example, any general statement in the theory (such as equilibrium and behaviour) is the same regardless of whether we use one system of generalised coordinates or another. In using Hamilton's approach, we can actually go one step better and look at transforming one set of generalised coordinates *and* momenta, to another. However, we must make sure the canonical form of the equations of motion in the system aren't ruined by the transformation. A simple way of looking at this is imagine changing currency. I have a particular formula from converting Australian Dollars to English Pounds. Let's say I want to convert Australian Dollars to English Pounds, but now I want to (for some strange reason) change them into Japanese Yen first then into pounds. So I work out a nice formula that does that for me. A canonical transformation means that my conversion from Australian dollars to English Pounds is not changed because I'm changing them to Japanese Yen first. Plus, it makes sure I don't change my dollar into, say, $4i$ pounds. In mathematical terms, we need to make sure all the transformations are **canonical**. What does that mean exactly?

Well if I have my Hamilton's equations for my set of coordinates q_i and momenta p_i :

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

I transform my (q, p) into (Q, P) given a certain transformation equation. I want to ensure that I get Hamilton's equations with my new coordinates:

$$\dot{Q}_i = \frac{\partial H}{\partial P_i} \quad \dot{P}_i = -\frac{\partial H}{\partial Q_i}$$

This happens when I have a **canonical transformation**. Any other transformations won't allow us to work under the Hamiltonian method of dynamics, and thus isn't much use to us.

So how do we get such a transformation? Well one way to think of it is imagine we have all our generalised coordinates and momenta presented in a Euclidean phase space of $2n$ dimensions with our standard rectangular Cartesian axes (which may sound more complicated than it actually is,

however you consider representing a $2n$ -dimensional graph is probably the way I'm trying to relate). So at a certain point in time, we would have all our data as little dots in $2n$ -space. Now we want to keep the data we have, but just represent it differently. So imagine keeping our data points fixed as we *rotate the axes*. If we align how we view these new axes, it will seem that our data points have moved, although in reality they haven't. You can visualise this for yourself: plot some arbitrary points on a piece of paper and then use a transparent sheet to draw the axes on. By rotating the axes about we change how our points are measured, but not the actual data - how each point relates to another. So the important thing is that the *values* of the dynamical data do not change, just how we represent them.

Which is all well and good, but we need an actual mathematical method to see whether our transformation are canonical or not. There is a method, unsurprisingly, which relies heavily on what is known as a **Poisson Bracket**.

Definition. Given functions $F(q, p)$ and $G(q, p)$, the **Poisson Bracket**³ of F and G is defined to be:

$$\{F, G\} = \frac{\partial F}{\partial q} \frac{\partial G}{\partial p} - \frac{\partial F}{\partial p} \frac{\partial G}{\partial q}$$

Note that $\{F, G\} = -\{G, F\}$

Now we can state the theorem we require:

Theorem 7.1 (Canonical Transformation Requirement). A transformation $(q, p) \rightarrow (Q, P)$ is called **canonical** if $\{Q, P\} = 1$.

Proof. Now we want our Hamiltonians in both our regular and transformed systems to be the same so:

$$H(q, p) = H'(Q(q, p), P(q, p))$$

We can then say:

$$\frac{\partial H}{\partial p} = \frac{\partial H'}{\partial Q} \frac{\partial Q}{\partial p} + \frac{\partial H'}{\partial P} \frac{\partial P}{\partial p} \tag{7.1}$$

and similarly for $\frac{\partial H}{\partial q}$.

³Poisson means *fish* in French. So in good, silly humour, some people call them Fish Brackets, without meaning disrespect to the great mathematician Siméon Denis Poisson, responsible for Poisson Distributions amongst other things.

The goal of this proof is to try to restate Hamilton's equations, but with respect to our transformed coordinates. So we say:

$$\begin{aligned}\dot{Q} &= \frac{\partial Q}{\partial q} \dot{q} + \frac{\partial Q}{\partial p} \dot{p} && \text{(Total differential)} \\ &= \frac{\partial Q}{\partial q} \frac{\partial H}{\partial p} - \frac{\partial Q}{\partial p} \frac{\partial H}{\partial q} && \text{(By Hamilton's eqns, which we know hold)} \\ &= \{Q, H\}\end{aligned}$$

Now using Equation 7.1, we can write \dot{Q} as:

$$\begin{aligned}\dot{Q} &= \frac{\partial Q}{\partial q} \left(\frac{\partial H'}{\partial Q} \frac{\partial Q}{\partial p} + \frac{\partial H'}{\partial P} \frac{\partial P}{\partial p} \right) - \frac{\partial Q}{\partial p} \left(\frac{\partial H'}{\partial Q} \frac{\partial Q}{\partial q} + \frac{\partial H'}{\partial P} \frac{\partial P}{\partial q} \right) \\ &= \frac{\partial H'}{\partial P} \left(\frac{\partial Q}{\partial q} \frac{\partial P}{\partial p} - \frac{\partial Q}{\partial p} \frac{\partial P}{\partial q} \right) \\ \dot{Q} &= \{Q, P\} \frac{\partial H'}{\partial P}\end{aligned}$$

Similarly (which you can prove for yourself),

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

Now our expressions for \dot{Q} and \dot{P} look suspiciously like Hamilton's equations, except with the pesky Poisson Bracket out the front. Therefore, we can deduce that for our transformed coordinates obey Hamilton's equations if and only if:

$$\{Q, P\} = 1$$

Note that we can also look at the Jacobian of the transformation:

$$J = \frac{\partial(Q, P)}{\partial(q, p)} = \begin{vmatrix} \frac{\partial Q}{\partial q} & \frac{\partial Q}{\partial p} \\ \frac{\partial P}{\partial q} & \frac{\partial P}{\partial p} \end{vmatrix}$$

So alternatively, $J = 1$ for a canonical transformation. \square

Example 7.2.1.

We have the transformation:

$$\begin{aligned}Q &= \cos \alpha q + \sin \alpha p \\ P &= -\sin \alpha q + \cos \alpha p\end{aligned}$$

Show that this transformation is canonical.

Answer.

We have to show that the Poisson Bracket $\{Q, P\}$ equals 1.

$$\begin{aligned}\{Q, P\} &= \frac{\partial Q}{\partial q} \frac{\partial P}{\partial p} - \frac{\partial Q}{\partial p} \frac{\partial P}{\partial q} \\ &= \cos^2 \alpha + \sin^2 \alpha \\ &= 1 \quad (\text{By trigonometry})\end{aligned}$$

\therefore the transformation is canonical.

Appendix A

Conclusion

Thus ends the Guru's Guide to Advanced Mechanics. Although we covered a lot of material, there is quite a lot more that we didn't. If you're interested, you may investigate:

- Chaos
- Perturbation theory
- Vibrational dynamics
- Fluid mechanics
- Special Relativity
- Quantum Mechanics

All of the above topics make use of the concepts developed here, and extend upon them greatly. You may also investigate particular systems, such as nonconservative systems or rheonomic systems (where the constraints of a system are "moving"). In particular, we suggest looking at chaotic systems as they give weird results, but nonetheless just apply Hamilton's equations and phase space plots to their full extent (plus chaos is cool).

We hope this text has provided some help and/or insight into dynamics. If you have any questions, suggestions or you've found errors, please don't hesitate to contact the author at witty@physics.uq.edu.au or visit the website at <http://www.maths.uq.edu.au/courses/guru/>.