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

Firstly, 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. Let's 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* C. Each point in C specifies a configuration of the system (i.e. the positions of all N particles). Time evolution gives rise to a curve in C.

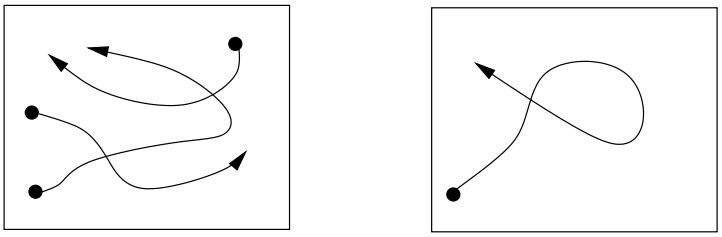

Figure 2: The path of particles in real space (on the left) and in configuration space (on the right).

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 minus sign between T and V! To describe the principle of least action, we consider all smooth paths $x^A(t)$ in C with fixed end points so that

$$x^A(t_i) = x^A_{\text{initial}} \qquad \text{and} \qquad x^A(t_f) = x^A_{\text{final}} \tag{2.3}$$

Of all these possible paths, only one is the true path taken by the system. Which one? To each path, let us assign a number called the *action* S defined as

$$S[x^A(t)] = \int_{t_i}^{t_f} L(x^A(t), \dot{x}^A(t))\, dt \tag{2.4}$$

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:

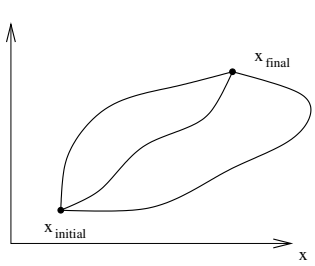

Figure 3:

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

$$x^A(t) \to x^A(t) + \delta x^A(t) \tag{2.5}$$

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

$$\begin{aligned} \delta S &= \delta \left[\int_{t_i}^{t_f} L\, dt \right] \\ &= \int_{t_i}^{t_f} \delta L\, dt \\ &= \int_{t_i}^{t_f} \left(\frac{\partial L}{\partial x^A} \delta x^A + \frac{\partial L}{\partial \dot{x}^A} \delta \dot{x}^A \right) dt \end{aligned} \tag{2.6}$$

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

$$\delta S = \int_{t_i}^{t_f} \left(\frac{\partial L}{\partial x^A} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^A} \right) \right) \delta x^A\, dt + \left[\frac{\partial L}{\partial \dot{x}^A} \delta x^A \right]_{t_i}^{t_f} \tag{2.7}$$

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

$$\frac{\partial L}{\partial x^A} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^A} \right) = 0 \qquad \text{for each } A = 1, \ldots 3N \tag{2.8}$$

These are known as *Lagrange's equations* (or sometimes as the Euler-Lagrange 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 (2.8) are indeed equivalent to (2.1). □

Some remarks on this important result:

- This is an example of a variational principle which you already met in the eponymous "variational principles" course.
- 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".
- *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

$$L = \sqrt{g}\left(R - \tfrac{1}{2}F_{\mu\nu}F^{\mu\nu} + \bar{\psi}\not{D}\psi\right) \tag{2.9}$$

  where the terms carry the names of Einstein, Maxwell (or Yang and Mills) and Dirac respectively, and describe gravity, the forces of nature (electromagnetism and the nuclear forces) and the dynamics of particles like electrons and quarks. 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.
- There is 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. We will describe this in Section 4.8.
- Back to classical mechanics, there are two very important reasons for working with Lagrange's equations rather than Newton's. The first is that Lagrange's equations hold in any coordinate system, while Newton's are restricted to an inertial frame. The second is the ease with which we can deal with constraints in the Lagrangian system. We'll look at these two aspects in the next two subsections.

2.2 Changing Coordinate Systems

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

$$q_a = q_a(x_1, \ldots, x_{3N}, t) \tag{2.10}$$

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

$$\dot{q}_a = \frac{dq_a}{dt} = \frac{\partial q_a}{\partial x^A}\dot{x}^A + \frac{\partial q_a}{\partial t} \tag{2.11}$$

In this equation, and for the rest of this course, we're using the "summation convention" in which repeated indices are summed over. Note also that we won't be too careful about whether indices are up or down - it won't matter for the purposes of this course. To be a good coordinate system, we should be able to invert the relationship so that $x^A = x^A(q_a, t)$ which we can do as long as we have $\det(\partial x^A/\partial q_a) \neq 0$. Then we have,

$$\dot{x}^A = \frac{\partial x^A}{\partial q_a}\dot{q}_a + \frac{\partial x^A}{\partial t} \tag{2.12}$$

Now we can examine $L(x^A, \dot{x}^A)$ when we substitute in $x^A(q_a, t)$. Using (2.12) we have

$$\frac{\partial L}{\partial q_a} = \frac{\partial L}{\partial x^A}\frac{\partial x^A}{\partial q_a} + \frac{\partial L}{\partial \dot{x}^A}\left(\frac{\partial^2 x^A}{\partial q_a \partial q_b}\dot{q}_b + \frac{\partial^2 x^A}{\partial t \partial q_a}\right) \tag{2.13}$$

while

$$\frac{\partial L}{\partial \dot{q}_a} = \frac{\partial L}{\partial \dot{x}^A}\frac{\partial \dot{x}^A}{\partial \dot{q}_a} \tag{2.14}$$

We now use the fact that we can "cancel the dots" and $\partial \dot{x}^A/\partial \dot{q}_a = \partial x^A/\partial q_a$ which we can prove by substituting the expression for $\dot{x}^A$ into the LHS. Taking the time derivative of (2.14) gives us

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_a}\right) = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^A}\right)\frac{\partial x^A}{\partial q_a} + \frac{\partial L}{\partial \dot{x}^A}\left(\frac{\partial^2 x^A}{\partial q_a \partial q_b}\dot{q}_b + \frac{\partial^2 x^A}{\partial q_a \partial t}\right) \tag{2.15}$$

So combining (2.13) with (2.15) we find

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_a}\right) - \frac{\partial L}{\partial q_a} = \left[\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^A}\right) - \frac{\partial L}{\partial x^A}\right]\frac{\partial x^A}{\partial q_a} \tag{2.16}$$

Equation (2.16) is our final result. We see that if Lagrange's equation is solved in the x^A coordinate system (so that [...] on the RHS vanishes) then it is also solved in the q_a coordinate system. (Conversely, if it is satisfied in the q_a coordinate system, so the LHS vanishes, then it is also satisfied in the x^A coordinate system as long as our choice of coordinates is invertible: i.e $\det(\partial x^A/\partial q_a) \neq 0$).

So the form of Lagrange's equations holds in *any* coordinate system. This is in contrast to Newton's equations which are only valid in an inertial frame. Let's illustrate the power of this fact with a couple of simple examples

2.2.1 Example: Rotating Coordinate Systems

Consider a free particle with Lagrangian given by

$$L = \tfrac{1}{2}m\dot{\mathbf{r}}^2 \tag{2.17}$$

with $\mathbf{r} = (x, y, z)$. 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 $\mathbf{r}' = (x', y', z')$ are the coordinates in the rotating system, we have the relationship

$$\begin{aligned} x' &= x\cos\omega t + y\sin\omega t \\ y' &= y\cos\omega t - x\sin\omega t \\ z' &= z \end{aligned} \tag{2.18}$$

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

$$L = \tfrac{1}{2}m[(\dot{x}' - \omega y')^2 + (\dot{y}' + \omega x')^2 + \dot{z}^2] = \tfrac{1}{2}m(\dot{\mathbf{r}}' + \boldsymbol{\omega}\times\mathbf{r}')^2 \tag{2.19}$$

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

$$\begin{aligned} \frac{\partial L}{\partial \mathbf{r}'} &= m(\dot{\mathbf{r}}'\times\boldsymbol{\omega} - \boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{r}')) \\ \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\mathbf{r}}'}\right) &= m(\ddot{\mathbf{r}}' + \boldsymbol{\omega}\times\dot{\mathbf{r}}') \end{aligned} \tag{2.20}$$

so Lagrange's equation reads

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\mathbf{r}}'}\right) - \frac{\partial L}{\partial \mathbf{r}'} = m(\ddot{\mathbf{r}}' + \boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{r}') + 2\boldsymbol{\omega}\times\dot{\mathbf{r}}') = 0 \tag{2.21}$$

The second and third terms in this expression are the centrifugal and coriolis forces respectively. These are examples of the "fictitious forces" that you were warned about in

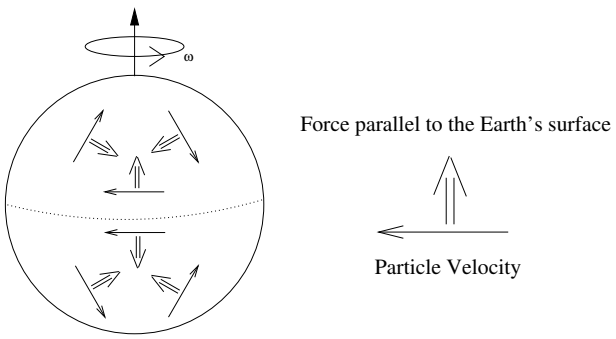

Figure 4: In the northern hemisphere, a particle is deflected in a clockwise direction; in the southern hemisphere in an anti-clockwise direction.

the first year. They're called fictitious because they're a consequence of the reference frame, rather than any interaction. But don't underestimate their importance just because they're "fictitious"! According to Einstein's theory of General Relativity, the force of gravity is on the same footing as these fictitious forces.

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

The coriolis force $\mathbf{F}_{\rm cor} = -2m\boldsymbol{\omega} \times \dot{\mathbf{r}}'$ 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, so has no effect on the particle.

More details on the effect of the Coriolis force in various circumstances can be found in the Dynamics and Relativity lecture notes. Questions discussed include:

- The coriolis force is responsible for the formation of hurricanes. These rotate in different directions in the northern and southern hemisphere, and never form within 500 miles of the equator where the coriolis force is irrelevant. But hurricanes rotate *anti-clockwise* in the northern hemisphere. This is the opposite direction from what we deduced above for a projected particle! What did we miss?

- Estimate the magnitude of the coriolis force. Do you think that it *really* affects the motion of water going down a plughole? What about the direction in which a CD spins?

- Stand on top of a tower at the equator and drop a ball. As the ball falls, the earth turns underneath from west to east. Does the ball land

  1. At the base of the tower?
  2. To the east?
  3. To the west?

2.2.2 Example: Hyperbolic Coordinates

A particle moves in the (x, y) plane with a force directed towards the origin O with magnitude proportional to the distance from O. How does it move? In Cartesian coordinates, this problem is easy. We have the Lagrangian

$$L = \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2) - \tfrac{1}{2}k(x^2 + y^2) \qquad (2.22)$$

Let's set $m = k = 1$ for simplicity. The equation of motion for this system is simply

$$\ddot{x} = -x \quad \text{and} \quad \ddot{y} = -y \qquad (2.23)$$

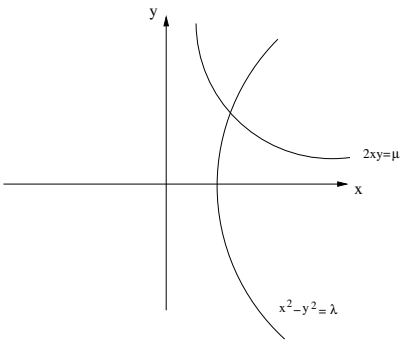

Figure 5: Hyperbolic coordinates.

Now suppose we want to know the motion of the system in *hyperbolic coordinates* defined as

$$2xy = \mu \quad , \quad x^2 - y^2 = \lambda \qquad (2.24)$$

The coordinates μ and λ are curvilinear and orthogonal (i.e. two hyperbolics intersect at 90^o). We could try solving this problem by substituting the change of coordinates directly into the equations of motion. It's a mess. (Try if you don't believe me!). A much simpler way is to derive expressions for $x, y, \dot{x}$ and $\dot{y}$ in terms of the new coordinates and substitute into the Lagrangian to find,

$$L = \tfrac{1}{8}\frac{\dot{\lambda}^2 + \dot{\mu}^2}{\sqrt{\lambda^2 + \mu^2}} - \tfrac{1}{2}\sqrt{\lambda^2 + \mu^2} \qquad (2.25)$$

From which we can easily derive the equation of motion for λ

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\lambda}}\right) - \frac{\partial L}{\partial \lambda} = \frac{d}{dt}\left(\frac{\dot{\lambda}}{4\sqrt{\lambda^2+\mu^2}}\right) + \tfrac{1}{8}(\dot{\lambda}^2+\dot{\mu}^2)\frac{\lambda}{(\lambda^2+\mu^2)^{3/2}} - \tfrac{1}{2}\frac{\lambda}{(\lambda^2+\mu^2)^{3/2}}$$
$$= 0 \tag{2.26}$$

Which is also a mess! But it's a mess that was much simpler to derive. Moreover, we don't need to do any more work to get the second equation for μ: the symmetry of the Lagrangian means that it must be the same as (2.26) with $\lambda \leftrightarrow \mu$ interchanged.

2.3 Constraints and Generalised Coordinates

Now we turn to the second advantage of the Lagrangian formulation. In writing $\dot{\mathbf{p}}_i = -\nabla_i V$, we implicitly assume that each particle can happily roam anywhere in space $\mathbf{R}^3$. What if there are constraints? In Newtonian mechanics, we introduce "constraint forces". 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

The simple pendulum 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 parameterise this as $x = l\sin\theta$ and $y = l\cos\theta$. 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,

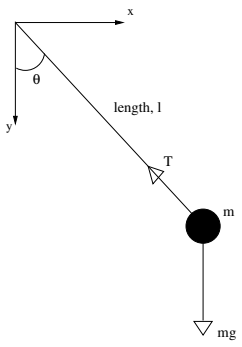

Figure 6:

$$m\ddot{x} = -Tx/l \quad , \quad m\ddot{y} = mg - Ty/l \tag{2.27}$$

To determine the motion of the system, we impose the constraints at the level of the equation of motion, and then easily find

$$\ddot{\theta} = -(g/l)\sin\theta \quad , \quad T = ml\dot{\theta}^2 + mg\cos\theta \tag{2.28}$$

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. Firstly, let's define what we mean by constraints more rigorously.

2.3.1 Holonomic Constraints

Holonomic Constraints are relationships between the coordinates of the form

$$f_\alpha(x_A, t) = 0 \qquad \alpha = 1, \ldots, 3N - n \tag{2.29}$$

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

$$x_A = x_A(q_1, \ldots, q_n) \tag{2.30}$$

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 they are functions of time. We now define a new Lagrangian

$$L' = L(x^A, \dot{x}^A) + \lambda_\alpha f_\alpha(x^A, t) \tag{2.31}$$

We treat λ_α like new coordinates. Since L' doesn't depend on $\dot{\lambda}_\alpha$, Lagrange's equations for λ_α are

$$\frac{\partial L'}{\partial \lambda_\alpha} = f_\alpha(x^A, t) = 0 \tag{2.32}$$

which gives us back the constraints. Meanwhile, the equations for x^A are

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^A}\right) - \frac{\partial L}{\partial x^A} = \lambda_\alpha \frac{\partial f_\alpha}{\partial x^A} \tag{2.33}$$

The LHS is the equation of motion for the unconstrained system. The RHS is the manifestation of 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 for a free particle moving in the plane, augmented by the Lagrange multiplier term for the constraints. It is

$$L' = \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2) + mgy + \tfrac{1}{2}\lambda(x^2 + y^2 - l^2) \tag{2.34}$$

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

$$m\ddot{x} = \lambda x \qquad \text{and} \qquad \ddot{y} = mg + \lambda y \tag{2.35}$$

while the equation of motion for λ reproduces the constraint $x^2+y^2-l^2 = 0$. Comparing with the Newtonian approach (2.27), we again see 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

$$L[q_i, \dot{q}_i, t] = L[x^A(q_i, t), \dot{x}^A(q_i, \dot{q}_i, t)] \tag{2.36}$$

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

$$x_A \to \begin{cases} q_i & i = 1, \ldots, n \\ f_\alpha & \alpha = 1, \ldots 3N - n \end{cases} \tag{2.37}$$

We know that Lagrange's equations take the same form in these new coordinates. In particular, we may look at the equations for q_i,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = \lambda_\alpha \frac{\partial f_\alpha}{\partial q_i} \tag{2.38}$$

But, by definition, $\partial f_\alpha/\partial q_i = 0$. So we are left with Lagrange's equations purely in terms of q_i, with no sign of the constraint forces. 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 (2.36) where we just substitute in $x_A = x_A(q_i, t)$. □

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

$$\begin{aligned} L &= \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2) + mgy \\ &= \tfrac{1}{2}ml^2\dot{\theta}^2 + mgl\cos\theta \end{aligned} \tag{2.39}$$

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

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right) - \frac{\partial L}{\partial \theta} = ml^2\ddot{\theta} + mgl\sin\theta = 0 \tag{2.40}$$

which indeed reproduces the equation of motion for the pendulum (2.28). 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.

2.3.2 Non-Holonomic Constraints

For completeness, let's quickly review a couple of non-holonomic constraints. There's no general theory to solve systems of this type, although it turns out that both of the examples we describe here can be solved with relative ease using different methods. We won't discuss non-holonomic constraints for the rest of this course, and include a brief description here simply to inform you of the sort of stuff we won't see!

Inequalities

Consider a particle moving under gravity on the outside of a sphere of radius R. It is constrained to satisfy $x^2+y^2+z^2 \geq R^2$. This type of constraint, involving an inequality, is non-holonomic. When the particle lies close to the top of the sphere, we know that it will remain in contact with the surface and we can treat the constraint effectively as holonomic. But at some point the particle will fall off. To determine when this happens requires different methods from those above (although it is not particularly difficult).

Velocity Dependent Constraints

Constraints of the form $g(x^A, \dot{x}^A, t) = 0$ which cannot be integrated to give $f(x^A, t) = 0$ are non-holonomic. For example, consider a coin of radius R rolling down a slope as shown in figure 7. The coordinates (x, y) fix the coin's position on the slope. But the coin has other degrees of freedom as well: the angle θ it makes with the path of steepest

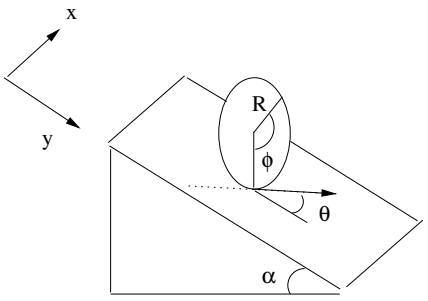

Figure 7: The coin rolling down a slope leads to velocity dependant, non-holonomic constraints.

descent, and the angle ϕ that a marked point on the rim of the coin makes with the vertical. If the coin rolls without slipping, then there are constraints on the evolution of these coordinates. We must have that the velocity of the rim is $v_{\rm rim} = R\dot{\phi}$. So, in terms of our four coordinates, we have the constraint

$$\dot{x} = R\dot{\phi}\sin\theta \qquad , \qquad \dot{y} = R\dot{\phi}\cos\theta \tag{2.41}$$

But these cannot be integrated to give constraints of the form $f(x, y, \theta, \phi) = 0$. They are non-holonomic.

2.3.3 Summary

Let's review what we've learnt so far. A system is described by n *generalised coordinates* q_i which define a point in an n-dimensional *configuration space* C. Time evolution is a curve in C governed by the *Lagrangian*

$$L(q_i, \dot{q}_i, t) \tag{2.42}$$

such that the q_i obey

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0 \tag{2.43}$$

These are n coupled $2^{\rm nd}$ order (usually) non-linear differential equations. Before we move on, let's take this opportunity to give an important definition. The quantity

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \tag{2.44}$$

is called the *generalised momentum* conjugate to q_i. (It only coincides with the real momentum in Cartesian coordinates). We can now rewrite Lagrange's equations (2.43) as $\dot{p}_i = \partial L/\partial q_i$. The generalised momenta will play an important role in Section 4.

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

$$\begin{aligned} & L' = \alpha L \qquad \text{for } \alpha \in \mathbf{R} \\ \text{or} \qquad & L' = L + \frac{df}{dt} \end{aligned} \tag{2.45}$$

for any function f and the equations of motion remain unchanged. To see that the last statement is true, we could either plug L' into Lagrange's equations or, alternatively, recall that Lagrange's equations can be derived from an action principle and the action (which is the time integral of the Lagrangian) changes only by a constant under the transformation. (As an aside: A system no longer remains invariant under these transformations in quantum mechanics. The number α is related to Planck's constant, while transformations of the second type lead to rather subtle and interesting effects related to the mathematics of topology).

2.3.4 Joseph-Louis Lagrange (1736-1813)

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.

Since I started with a quote about Newton's teaching, I'll include here a comment on Lagrange's lectures by one of his more famous students:

> "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*

2.4 Noether's Theorem and Symmetries

In this subsection we shall discuss the appearance of conservation laws in the Lagrangian formulation and, in particular, a beautiful and important theorem due to Noether relating conserved quantities to symmetries.

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 *constant of motion* (or a *conserved quantity*) if the total time derivative vanishes

$$\frac{dF}{dt} = \sum_{j=1}^{n} \left(\frac{\partial F}{\partial q_j} \dot{q}_j + \frac{\partial F}{\partial \dot{q}_j} \ddot{q}_j \right) + \frac{\partial F}{\partial t} = 0 \tag{2.46}$$

whenever $q_i(t)$ satisfy Lagrange's equations. This means that F remains constant along the path followed by the system. Here's a couple of examples:

Claim: If L does not depend explicitly on time t (i.e. $\partial L/\partial t = 0$) then

$$H = \sum_j \dot{q}_j \frac{\partial L}{\partial \dot{q}_j} - L \tag{2.47}$$

is constant. When H is written as a function of q_i and p_i, it is known as the Hamiltonian. It is usually identified with the total energy of the system.

Proof

$$\frac{dH}{dt} = \sum_j \left(\ddot{q}_j \frac{\partial L}{\partial \dot{q}_j} + \dot{q}_j \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_j} \right) - \frac{\partial L}{\partial q_j} \dot{q}_j - \frac{\partial L}{\partial \dot{q}_j} \ddot{q}_j \right) \tag{2.48}$$

which vanishes whenever Lagrange's equations (2.43) hold. □

Claim: Suppose $\partial L/\partial q_j = 0$ for some q_j. Then q_j is said to be *ignorable* (or *cyclic*). We have the conserved quantity

$$p_j = \frac{\partial L}{\partial \dot{q}_j} \tag{2.49}$$

Proof:

$$\frac{dp_j}{dt} = \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_j} \right) = \frac{\partial L}{\partial q_j} = 0 \tag{2.50}$$

where we have used Lagrange's equations (2.43) in the second equality. □

2.4.1 Noether's Theorem

Consider a one-parameter family of maps

$$q_i(t) \to Q_i(s,t) \qquad\qquad s \in \mathbf{R} \tag{2.51}$$

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

$$\frac{\partial}{\partial s} L(Q_i(s,t), \dot{Q}_i(s,t), t) = 0 \tag{2.52}$$

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

Proof of Noether's Theorem:

$$\frac{\partial L}{\partial s} = \frac{\partial L}{\partial Q_i}\,\frac{\partial Q_i}{\partial s} + \frac{\partial L}{\partial \dot{Q}_i}\,\frac{\partial \dot{Q}_i}{\partial s} \tag{2.53}$$

so we have

$$\begin{aligned} 0 = \left.\frac{\partial L}{\partial s}\right|_{s=0} &= \left.\frac{\partial L}{\partial q_i}\,\frac{\partial Q_i}{\partial s}\right|_{s=0} + \left.\frac{\partial L}{\partial \dot{q}_i}\,\frac{\partial \dot{Q}_i}{\partial s}\right|_{s=0} \\ &= \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right)\left.\frac{\partial Q_i}{\partial s}\right|_{s=0} + \left.\frac{\partial L}{\partial \dot{q}_i}\,\frac{\partial \dot{Q}_i}{\partial s}\right|_{s=0} \qquad \text{(By Lagrange)} \\ &= \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\left.\frac{\partial Q_i}{\partial s}\right|_{s=0}\right) \end{aligned} \tag{2.54}$$

and the quantity $\sum_i (\partial L/\partial \dot{q}_i)(\partial Q_i/\partial s)$, evaluated at $s = 0$, is constant for all time. □

Example: Homogeneity of Space

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

$$L = \tfrac{1}{2}\sum_i m_i \dot{\mathbf{r}}_i^2 - V(|\mathbf{r}_i - \mathbf{r}_j|) \tag{2.55}$$

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

$$L(\mathbf{r}_i, \dot{\mathbf{r}}_i, t) = L(\mathbf{r}_i + s\mathbf{n}, \dot{\mathbf{r}}_i, t) \tag{2.56}$$

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

$$\sum_i \frac{\partial L}{\partial \dot{\mathbf{r}}_i} \cdot \mathbf{n} = \sum_i \mathbf{p}_i \cdot \mathbf{n} \tag{2.57}$$

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.

Example: Isotropy of Space

The isotropy of space is the statement that a closed system, described by the Lagrangian (2.55) 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

$$\begin{aligned} \mathbf{r}_i &\to \mathbf{r}_i + \delta\mathbf{r}_i \\ &= \mathbf{r}_i + \alpha\hat{\mathbf{n}} \times \mathbf{r}_i \end{aligned} \tag{2.58}$$

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

$$L(\mathbf{r}_i, \dot{\mathbf{r}}_i) = L(\mathbf{r}_i + \alpha\hat{\mathbf{n}} \times \mathbf{r}_i, \dot{\mathbf{r}}_i + \alpha\hat{\mathbf{n}} \times \dot{\mathbf{r}}_i) \tag{2.59}$$

which gives rise to the conserved quantity

$$\sum_i \frac{\partial L}{\partial \dot{\mathbf{r}}_i} \cdot (\hat{\mathbf{n}} \times \mathbf{r}_i) = \sum_i \hat{\mathbf{n}} \cdot (\mathbf{r}_i \times \mathbf{p}_i) = \hat{\mathbf{n}} \cdot \mathbf{L} \tag{2.60}$$

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

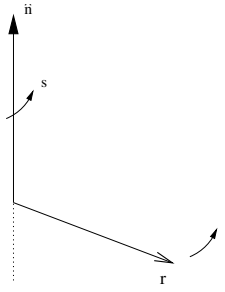

Figure 8:

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

Example: Homogeneity of Time

What about homogeneity of time? 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 earlier in this section that this implies $H = \sum_i \dot{q}_i(\partial L/\partial \dot{q}_i) - L$ is conserved. In the systems we're considering, this is simply the total energy. We see that the existence of a conserved quantity which we call energy can be traced to the homogeneous passage of time. Or

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.

Remarks: It turns out that *all* conservation laws in Nature (as described by 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 (known as baryons).

There are also *discrete* symmetries in Nature which don't depend on a continuous parameter. For example, many theories are invariant under reflection (known as parity) in which $\mathbf{r}_i \to -\mathbf{r}_i$. These types of symmetries do not give rise to conservation laws in classical physics (although they do in quantum physics).

2.5 Applications

Having developed all of these tools, let's now apply them to a few examples.

2.5.1 Bead on a Rotating Hoop

This is an example of a system with a time dependent holonomic constraint. The hoop is of radius a and rotates with frequency ω as shown in figure 9. The bead, of mass m, is threaded on the hoop and moves without friction. We want to determine its motion. There is a single degree of freedom ψ, the angle the bead makes with the vertical. In terms of Cartesian coordinates (x, y, z) the position of the bead is

$$x = a\sin\psi\cos\omega t \quad , \quad y = a\sin\psi\sin\omega t \quad , \quad z = a - a\cos\psi \tag{2.61}$$

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

$$T = \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) = \tfrac{1}{2}ma^2[\dot{\psi}^2 + \omega^2\sin^2\psi] \tag{2.62}$$

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

$$V = mgz = -mga\cos\psi \tag{2.63}$$

So, replacing x, y and z by ψ, we have the Lagrangian

$$L = ma^2\left(\tfrac{1}{2}\dot{\psi}^2 - V_{\rm eff}\right) \tag{2.64}$$

where the effective potential is

$$V_{\rm eff} = \frac{1}{ma^2}\left(-mga\cos\psi - \tfrac{1}{2}ma^2\omega^2\sin^2\psi\right) \tag{2.65}$$

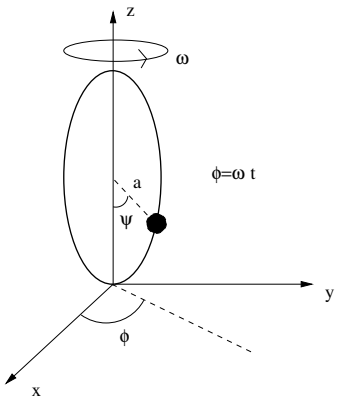

Figure 9:

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

$$\ddot{\psi} = -\frac{\partial V_{\rm eff}}{\partial\psi} \tag{2.66}$$

Let's 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 $\partial V_{\rm eff}/\partial\psi = 0$ to find that the bead can remain stationary at points satisfying

$$g\sin\psi = a\omega^2\sin\psi\cos\psi \tag{2.67}$$

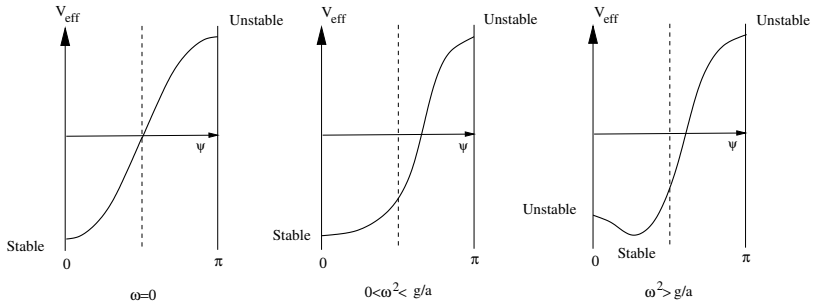

Figure 10: The effective potential for the bead depends on how fast the hoop is rotating

There are at most three such points: $\psi = 0$, $\psi = \pi$ or $\cos\psi = g/a\omega^2$. Note that the first two solutions always exist, while the third stationary point is only there if the hoop is spinning fast enough so that $\omega^2 \geq g/a$. Which of these stationary points is stable depends on whether $V_{\rm eff}(\psi)$ has a local minimum (stable) or maximum (unstable). This

in turn depends on the value of ω. $V_{\rm eff}$ is drawn for several values of ω in figure 10. For $\omega^2 < g/a$, the point $\psi = 0$ at the bottom of the hoop is stable, while for $\omega^2 > g/a$, the position at the bottom becomes unstable and the new solution at $\cos\psi = g/a\omega^2$ is the stable point. For all values of ω the bead perched at the top of the hoop $\psi = \pi$ is unstable.

2.5.2 Double Pendulum

A double pendulum is drawn in figure 11, consisting 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

$$T_1 = \tfrac{1}{2}m_1 l_1^2\dot\theta_1^2 \quad \text{and} \quad V_1 = -m_1 g l_1\cos\theta_1 \tag{2.68}$$

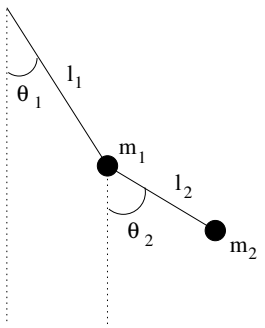

Figure 11:

For the second particle it's a little more involved. Consider the position of the second particle in the (x, y) plane in which the pendulum swings (where we take the origin to be the pivot of the first pendulum with y increasing downwards)

$$x_2 = l_1\sin\theta_1 + l_2\sin\theta_2 \quad \text{and} \quad y_2 = l_1\cos\theta_1 + l_2\cos\theta_2 \tag{2.69}$$

Which we can substitute into the kinetic energy for the second particle

$$\begin{aligned} T_2 &= \tfrac{1}{2}m_2(\dot x^2 + \dot y^2) \\ &= \tfrac{1}{2}m_2\left(l_1^2\dot\theta_1^2 + l_2^2\dot\theta_2^2 + 2l_1 l_2\cos(\theta_1-\theta_2)\dot\theta_1\dot\theta_2\right) \end{aligned} \tag{2.70}$$

while the potential energy is given by

$$V_2 = -m_2 g y_2 = -m_2 g\left(l_1\cos\theta_1 + l_2\cos\theta_2\right) \tag{2.71}$$

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

$$\begin{aligned} L = {} & \tfrac{1}{2}(m_1+m_2)l_1^2\dot\theta_1^2 + \tfrac{1}{2}m_2 l_2^2\dot\theta_2^2 + m_2 l_1 l_2\cos(\theta_1-\theta_2)\dot\theta_1\dot\theta_2 \\ & +(m_1+m_2)g l_1\cos\theta_1 + m_2 g l_2\cos\theta_2 \end{aligned} \tag{2.72}$$

The equations of motion follow by simple calculus using Lagrange's two equations (one for θ_1 and one for θ_2). The solutions to these equations are complicated. In fact, above a certain energy, the motion is chaotic.

2.5.3 Spherical Pendulum

The spherical pendulum is allowed to rotate in three dimensions. The system has two degrees of freedom drawn in figure 12 which cover the range

$$0 \leq \theta < \pi \qquad \text{and} \qquad 0 \leq \phi < 2\pi \tag{2.73}$$

In terms of cartesian coordinates, we have

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

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

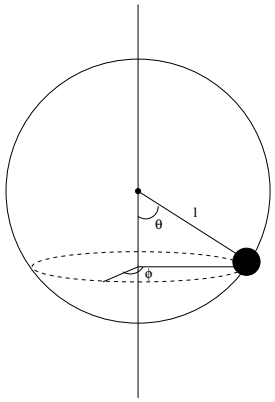

Figure 12:

$$\begin{aligned} L &= \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - mgz \\ &= \tfrac{1}{2}ml^2(\dot{\theta}^2 + \dot{\phi}^2\sin^2\theta) + mgl\cos\theta \end{aligned} \tag{2.74}$$

Notice that the coordinate ϕ is ignorable. From Noether's theorem, we know that the quantity

$$J = \frac{\partial L}{\partial \dot{\phi}} = ml^2\dot{\phi}\sin^2\theta \tag{2.75}$$

is constant. This is the component of angular momentum in the ϕ direction. The equation of motion for θ follows from Lagrange's equations and is

$$ml^2\ddot{\theta} = ml^2\dot{\phi}^2\sin\theta\cos\theta - mgl\sin\theta \tag{2.76}$$

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

$$\ddot{\theta} = -\frac{\partial V_{\text{eff}}}{\partial\theta} \tag{2.77}$$

where the effective potential is defined to be

$$V_{\text{eff}}(\theta) = -\frac{g}{l}\cos\theta + \frac{J^2}{2m^2l^4}\frac{1}{\sin^2\theta} \tag{2.78}$$

An important point here: we must substitute for J into the equations of motion. If you substitute J for $\dot{\phi}$ directly into the Lagrangian, you will derive an equation that looks like the one above, but you'll get a minus sign wrong! This is because Lagrange's equations are derived under the assumption that θ and ϕ are independent.

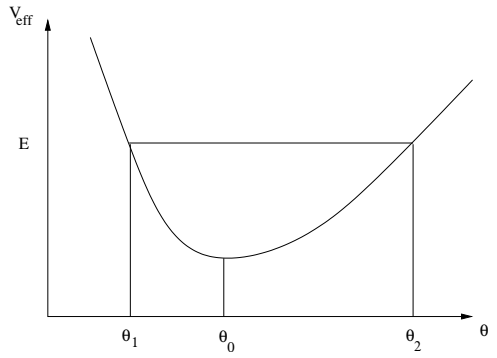

Figure 13: The effective potential for the spherical pendulum.

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

$$E = \tfrac{1}{2}\dot{\theta}^2 + V_{\text{eff}}(\theta) \tag{2.79}$$

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

$$t - t_0 = \frac{1}{\sqrt{2}} \int \frac{d\theta}{\sqrt{E - V_{\text{eff}}(\theta)}} \tag{2.80}$$

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 on a computer.

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 $\phi = \phi(\theta) = \phi(t)$. Let's 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 $V_{\text{eff}} \leq E$ (which follows from (2.79)). 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 figure 14, in which the bob oscillates between the two extremes: $\theta_1 \leq \theta \leq \theta_2$. Note that we could make more progress in understanding the motion of the spherical pendulum than for the

double pendulum. The reason for this is the existence of two conservation laws for the spherical pendulum (energy and angular momentum) compared to just one (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_{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 (2.77), we have

$$\delta\ddot{\theta} = -\left(\frac{\partial^2 V_{\text{eff}}}{\partial\theta^2}\bigg|_{\theta=\theta_0}\right)\delta\theta + \mathcal{O}(\delta\theta^2) \tag{2.81}$$

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

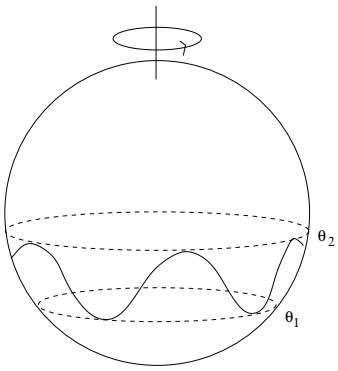

Figure 14:

2.5.4 Two Body Problem

We now turn to the study of two objects interacting through a central force. The most famous example of this type is the gravitational interaction between two bodies in the solar system which leads to the elliptic orbits of planets and the hyperbolic orbits of comets. Let's see how to frame this famous physics problem in the Lagrangian setting. We start by rewriting the Lagrangian in terms of the centre of mass $\mathbf{R}$ and the separation $\mathbf{r}_{12} = \mathbf{r}_1 - \mathbf{r}_2$ and work with an arbitrary potential $V(|\mathbf{r}_{12}|)$

$$\begin{aligned} L &= \tfrac{1}{2}m_1\dot{\mathbf{r}}_1^2 + \tfrac{1}{2}m_2\dot{\mathbf{r}}_2^2 - V(|\mathbf{r}_{12}|) \\ &= \tfrac{1}{2}(m_1+m_2)\dot{\mathbf{R}}^2 + \tfrac{1}{2}\mu\dot{\mathbf{r}}_{12}^2 - V(|\mathbf{r}_{12}|) \end{aligned} \tag{2.82}$$

where $\mu = m_1m_2/(m_1+m_2)$ is the *reduced mass*. The Lagrangian splits into a piece describing the centre of mass $\mathbf{R}$ and a piece describing the separation. This is familiar from Section 1.3.2. From now on we neglect the centre of mass piece and focus on the separation. We know from Noether's theorem that $\mathbf{L} = \mathbf{r}_{12} \times \mathbf{p}_{12}$ is conserved, where $\mathbf{p}_{12}$ is the momentum conjugate to $\mathbf{r}_{12}$. Since $\mathbf{L}$ is perpendicular to $\mathbf{r}_{12}$, the motion of the orbit must lie in a plane perpendicular to $\mathbf{L}$. Using polar coordinates (r, ϕ) in that plane, the Lagrangian is

$$L = \tfrac{1}{2}\mu(\dot{r}^2 + r^2\dot{\phi}^2) - V(r) \tag{2.83}$$

To make further progress, notice that ϕ is ignorable so, once again using Noether's theorem, we have the conserved quantity

$$J = \mu r^2\dot{\phi} \tag{2.84}$$

This is also conservation of angular momentum: to reduce to the Lagrangian (2.83), we used the fact that the direction of $\mathbf{L}$ is fixed; the quantity J is related to the magnitude of $\mathbf{L}$. To figure out the motion we calculate Lagrange's equation for r from (2.83)

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{r}}\right) - \frac{\partial L}{\partial r} = \mu\ddot{r} - \mu r\dot{\phi}^2 + \frac{\partial V}{\partial r} = 0 \tag{2.85}$$

We can 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,

$$\mu\ddot{r} = -\frac{\partial}{\partial r}V_{\text{eff}}(r) \tag{2.86}$$

where the effective potential is given by

$$V_{\text{eff}}(r) = V(r) + \frac{J^2}{2\mu r^2} \tag{2.87}$$

The last term is known as the "angular momentum barrier". Let me reiterate the warning of the spherical pendulum: 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.

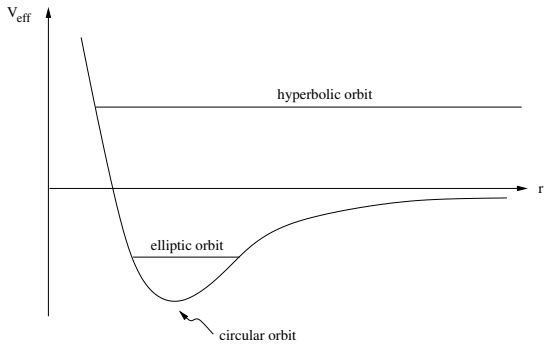

Figure 15: The effective potential for two bodies interacting gravitationally.

So far, you may recognise that the analysis has been rather similar to that of the spherical pendulum. Let's continue following that path. Since $\partial L/\partial t = 0$, Noether tells us that energy is conserved and

$$E = \tfrac{1}{2}\mu\dot{r}^2 + V_{\text{eff}}(r) \tag{2.88}$$

is constant throughout the motion. We can use this fact to "reduce to quadrature",

$$t - t_0 = \sqrt{\frac{\mu}{2}} \int \frac{dr}{\sqrt{E - V_{\text{eff}}(r)}} \tag{2.89}$$

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

$$V(r) = -\frac{Gm_1m_2}{r} \tag{2.90}$$

where G is Newton's constant. For this potential, the different solutions were studied in your Part I mechanics 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 15.

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. Eliminating the centre of mass reduced this to 3 degrees of freedom describing the separation. We then used conservation of the direction of $\mathbf{L}$ to reduce to 2 degrees of freedom (r and ϕ), and conservation of the magnitude of $\mathbf{L}$ to reduce to a single variable r. Finally conservation of E allowed us to solve the problem. You might now be getting an idea about how important conservation laws are to help us solve problems!

2.5.5 Restricted Three Body Problem

Consider three masses m_1, m_2 and m_3 interacting gravitationally. In general this problem does not have an analytic solution and we must resort to numerical methods (i.e. putting it on a computer). However, suppose that $m_3 \ll m_1$ and m_2. Then it is a good approximation to first solve for the motion of m_1 and m_2 interacting alone, and then solve for the motion of m_3 in the time dependent potential set up by m_1 and m_2. Let's see how this works.

For simplicity, let's assume m_1 and m_2 are in a circular orbit with $\phi = \omega t$. 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

$$\omega^2 = \frac{G(m_1 + m_2)}{r^3} \tag{2.91}$$

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 (which is a 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. Let's go to a frame which rotates with m_1 and m_2 with the centre of mass at the origin. The particle m_1 is a distance $r\mu/m_1$ from the origin, while m_2 is a distance $r\mu/m_2$ from the origin.

Then, from the example of Section 2.2.1, the Lagrangian for m_3 in the rotating frame is

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

where V is the gravitational potential for m_3 interacting with m_1 and m_2

$$V = -\frac{Gm_1m_3}{r_{13}} - \frac{Gm_2m_3}{r_{23}} \qquad (2.92)$$

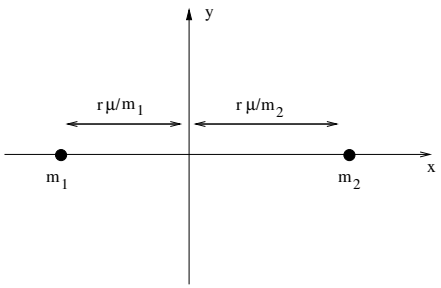

Figure 16:

The separations are given by

$$r_{13}^2 = (x + r\mu/m_1)^2 + y^2 \quad , \quad r_{23}^2 = (x - r\mu/m_2)^2 + y^2 \qquad (2.93)$$

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

$$\begin{aligned} m_3\ddot{x} &= 2m_3\omega\dot{y} + m_3\omega^2 x - \frac{\partial V}{\partial x} \\ m_3\ddot{y} &= -2m_3\omega\dot{x} + m_3\omega^2 y - \frac{\partial V}{\partial y} \end{aligned} \qquad (2.94)$$

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 (and, in the process, won 2,500 krona and lost 3,500 krona). We'll be a little less ambitious here 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, the centrifugal force of the third body exactly cancels its gravitational force. The equations we have to solve are

$$m_3\omega^2 x = \frac{\partial V}{\partial x} = Gm_1m_3\frac{x + r\mu/m_1}{r_{13}^3} + Gm_2m_3\frac{x - r\mu/m_2}{r_{23}^3} \qquad (2.95)$$

$$m_3\omega^2 y = \frac{\partial V}{\partial y} = Gm_1m_3\frac{y}{r_{13}^3} + Gm_2m_3\frac{y}{r_{23}^3} \qquad (2.96)$$

There are five solutions to these equations. Firstly suppose that $y = 0$ so that m_3 sits on the same line as m_1 and m_2. Then we have to solve the algebraic equation

$$\omega^2 x = Gm_1\frac{x + r\mu/m_1}{|x + r\mu/m_1|^3} + Gm_2\frac{x - r\mu/m_2}{|x - r\mu/m_2|^3} \qquad (2.97)$$

In figure 17, we have plotted the LHS and RHS of this equation to demonstrate the three solutions, one in each of the regimes:

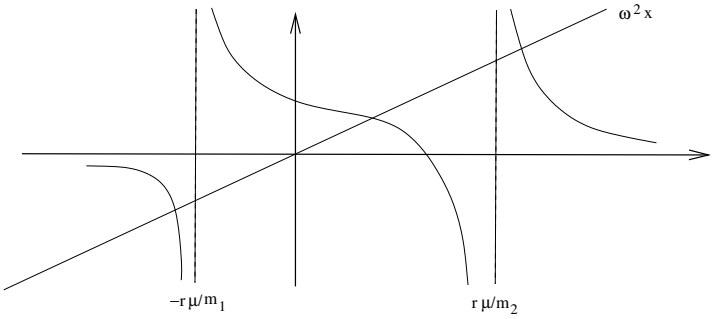

Figure 17: The three solutions sitting on $y = 0$.

$$x < -\frac{r\mu}{m_1} \quad , \quad -\frac{r\mu}{m_1} < x < \frac{r\mu}{m_2} \quad , \quad x > \frac{r\mu}{m_2} \tag{2.98}$$

Now let's look for solutions with $y \neq 0$. From (2.96) we have

$$\frac{Gm_2}{r_{23}^3} = \omega^2 - \frac{Gm_1}{r_{13}^3} \tag{2.99}$$

which we can substitute into (2.95) and, after a little algebra, we find the condition for solutions to be

$$\omega^2 = \frac{G(m_1 + m_2)}{r_{13}^3} = \frac{G(m_1 + m_2)}{r_{23}^3} \tag{2.100}$$

which means that we must have $r_{13} = r_{23} = r$. There are two such points.

In general there are five stationary points drawn in the figure. These are called Lagrange points. It turns out that L_1, L_2 and L_3 are unstable, while L_4 and L_5 are stable as long as m_2 is sufficiently less than m_1.

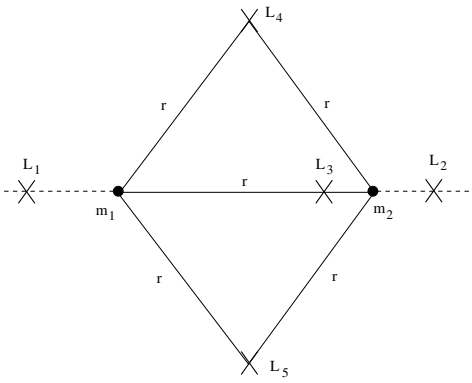

Figure 18: The five Lagrange points. X marks the spots.

For the earth-sun system, NASA and ESA make use of the Lagrange points L_2 and L_3 to place satellites. There are solar observatories at L_3; satellites such as WMAP and PLANCK which measure the cosmic microwave background radiation (the afterglow of the big bang) gather their data from L_2. Apparently, there is a large collection of cosmic dust which has accumulated at L_4 and L_5. Other planetary systems (e.g. the sun-jupiter and sun-mars systems) have large asteroids, known as trojans, trapped at their L_4 and L_5.

2.5.6 Purely Kinetic Lagrangians

Often in physics, one is interested in systems with only kinetic energy and no potential energy. For a system with n dynamical degrees of freedom q^a, $a = 1, \ldots, n$, the most general form of the Lagrangian with just a kinetic term is

$$L = \tfrac{1}{2}\, g_{ab}(q_c)\, \dot{q}^a \dot{q}^b \tag{2.101}$$

The functions $g_{ab} = g_{ba}$ depend on all the generalised coordinates. Assume that $\det(g_{ab}) \neq 0$ so that the inverse matrix g^{ab} exists ($g^{ab} g_{bc} = \delta^a{}_c$). It is a short exercise to show that Lagrange's equation for this system are given by

$$\ddot{q}^a + \Gamma^a_{bc} \dot{q}^b \dot{q}^c = 0 \tag{2.102}$$

where

$$\Gamma^a_{bc} = \tfrac{1}{2}\, g^{ad} \left(\frac{\partial g_{bd}}{\partial q^c} + \frac{\partial g_{cd}}{\partial q^b} - \frac{\partial g_{bc}}{\partial q^d} \right) \tag{2.103}$$

The functions g_{ab} define a *metric* on the configuration space, and the equations (2.102) are known as the *geodesic* equations. They appear naturally in General Relativity where they describe a particle moving in curved spacetime. Lagrangians of the form (2.101) also appear in many other areas of physics, including the condensed matter physics, the theory of nuclear forces and string theory. In these contexts, the systems are referred to as *sigma models*.

2.5.7 Particles in Electromagnetic Fields

We saw from the beginning that the Lagrangian formulation works with conservative forces which can be written in terms of a potential. It is no good at dealing with friction forces which are often of the type $F = -k\dot{x}$. But there are other velocity dependent forces which arise in the fundamental laws of Nature. It's a crucial fact about Nature that all of these can be written in Lagrangian form. Let's illustrate this in an important example.

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

$$\mathbf{B} = \nabla \times \mathbf{A} \quad , \quad \mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t} \tag{2.104}$$

Let's study the Lagrangian for a particle of electric charge e of the form,

$$L = \tfrac{1}{2} m \dot{\mathbf{r}}^2 - e\,(\phi - \dot{\mathbf{r}} \cdot \mathbf{A}) \tag{2.105}$$

The momentum conjugate to $\mathbf{r}$ is

$$\mathbf{p} = \frac{\partial L}{\partial \dot{\mathbf{r}}} = m\dot{\mathbf{r}} + e\mathbf{A} \tag{2.106}$$

Notice that the momentum is not simply $m\dot{\mathbf{r}}$; it's modified in the presence of electric and magnetic fields. Now we can calculate Lagrange's equations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\mathbf{r}}}\right) - \frac{\partial L}{\partial \mathbf{r}} = \frac{d}{dt}\left(m\dot{\mathbf{r}} + e\mathbf{A}\right) + e\nabla\phi - e\nabla(\dot{\mathbf{r}}\cdot\mathbf{A}) = 0 \tag{2.107}$$

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

$$m\ddot{r}^a = -e\left(\frac{\partial\phi}{\partial r^a} + \frac{\partial A_a}{\partial t}\right) + e\left(\frac{\partial A_b}{\partial r^a} - \frac{\partial A_a}{\partial r^b}\right)\dot{r}^b \tag{2.108}$$

Now we use our definitions of the $\mathbf{E}$ and $\mathbf{B}$ fields (2.104) which, in terms of indices, read

$$E_a = -\frac{\partial\phi}{\partial r^a} - \frac{\partial A_a}{\partial t} \quad , \qquad B_c = \epsilon_{cab}\frac{\partial A_b}{\partial r^a} \tag{2.109}$$

The last of these can be equivalently written as $\partial A_b/\partial r^a - \partial A_a/\partial r^b = \epsilon_{abc}B_c$. The equation of motion then reads

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

or, reverting to vector notation,

$$m\ddot{\mathbf{r}} = e\left(\mathbf{E} + \dot{\mathbf{r}}\times\mathbf{B}\right) \tag{2.110}$$

which is the Lorentz force law.

Gauge Invariance: The scalar and vector potentials are not unique. We may make a change of the form

$$\phi \to \phi - \frac{\partial\chi}{\partial t} \quad , \qquad \mathbf{A} \to \mathbf{A} + \nabla\chi \tag{2.111}$$

These give the same $\mathbf{E}$ and $\mathbf{B}$ fields for any function χ. This is known as a *gauge transformation*. Under this change, we have

$$L \to L + e\frac{\partial\chi}{\partial t} + e\dot{\mathbf{r}}\cdot\nabla\chi = L + e\frac{d\chi}{dt} \tag{2.112}$$

but we know that the equations of motion remain invariant under the addition of a total derivative to the Lagrangian. This concept of gauge invariance underpins much of modern physics. You can learn more about this in the lectures on Electromagnetism.

2.6 Small Oscillations and Stability

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

Peskin doesn't say this to knock physics. He's just a fan of harmonic oscillators. And rightly so. By studying the simple harmonic oscillator and its relatives in ever more inventive ways we understand why the stars shine and why lasers shine and, thanks to Hawking, why even black holes shine.

In this section we'll see one reason why the simple harmonic oscillator is so important to us. 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. We've already seen several examples where we get an equation of the form

$$\ddot{x} = f(x) \tag{2.113}$$

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

$$x = x_0 \qquad \text{and} \qquad \dot{x} = 0 \tag{2.114}$$

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

$$x(t) = x_0 + \eta(t) \tag{2.115}$$

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

$$\ddot{\eta} = f'(x_0)\,\eta \qquad + \mathcal{O}(\eta^2) \tag{2.116}$$

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

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

   $$\eta(t) = A\cos(\omega(t - t_0)) \tag{2.117}$$

   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.

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

$$\eta(t) = Ae^{\lambda t} + Be^{-\lambda t} \tag{2.118}$$

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*.

Now let's generalise this discussion to n degrees of freedom with equations of motion of the form,

$$\ddot{q}_i = f_i(q_1, \ldots, q_n) \qquad i = 1, \ldots, n \tag{2.119}$$

An equilibrium point q_i^0 must satisfy $f_i(q_1^0, \ldots, q_n^0) = 0$ for all $i = 1 \ldots, n$. Consider small perturbations away from the equilibrium point

$$q_i(t) = q_i^0 + \eta_i(t) \tag{2.120}$$

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

$$\ddot{\eta}_i \approx \left.\frac{\partial f_i}{\partial q_j}\right|_{q_k = q_k^0} \eta_j \tag{2.121}$$

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

$$\boldsymbol{\eta} = \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix} \quad , \quad F = \begin{pmatrix} \frac{\partial f_1}{\partial q_1} & \cdots & \frac{\partial f_1}{\partial q_n} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial q_1} & \cdots & \frac{\partial f_n}{\partial q_n} \end{pmatrix} \tag{2.122}$$

where each partial derivative in the matrix F is evaluated at $q_i = q_i^0$. The equation now becomes simply

$$\ddot{\boldsymbol{\eta}} = F\boldsymbol{\eta} \tag{2.123}$$

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, we can't assume that F is symmetric. Nonetheless, it is true that for equations

of the form (2.123) 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,

$$F\boldsymbol{\mu}_a = \lambda_a^2 \boldsymbol{\mu}_a \qquad , \qquad \boldsymbol{\zeta}_a^T F = \lambda_a^2 \, \boldsymbol{\zeta}_a^T \qquad a = 1, \ldots, n \tag{2.124}$$

where there's no sum over a in these equations. The left and right eigenvectors satisfy $\boldsymbol{\zeta}_a \cdot \boldsymbol{\mu}_b = \delta_{ab}$. Note that although the eigenvectors differ, the eigenvalues λ_a^2 for $a = 1, \ldots, n$ are the same. Although λ_a^2 are real for the physical systems of interest (to be proved shortly) they are not always positive. The most general solution to $\ddot{\boldsymbol{\eta}} = F\boldsymbol{\eta}$ is

$$\boldsymbol{\eta}(t) = \sum_a \boldsymbol{\mu}_a \left[A_a e^{\lambda_a t} + B_a e^{-\lambda_a t} \right] \tag{2.125}$$

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

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

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

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 is the case the system will typically oscillate around the equilibrium point as a linear superposition of all the normal modes, each at a different frequency.

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

$$\boldsymbol{\eta}(t) = \sum_{a, \lambda_a^2 > 0} \boldsymbol{\mu}_a \left[A_a e^{\lambda_a t} + B_a e^{-\lambda_a t} \right] + \sum_{a, \lambda_a^2 < 0} \boldsymbol{\mu}_a A_a \cos(\omega_a (t - t_a)) \tag{2.126}$$

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,

$$L = \tfrac{1}{2} T_{ij}(q) \dot{q}_i \dot{q}_j - V(q) \tag{2.127}$$

We will require that $T_{ij}(q)$ is invertible and positive definite for all q. Expanding about an equilibrium point as in (2.120), to linear order in η_i the equations read

$$T_{ij}\ddot{\eta}_j = -V_{ij}\eta_j \tag{2.128}$$

where $T_{ij} = T_{ij}(q^0)$ and $V_{ij} = \partial^2 V/\partial q_i \partial q_j$, evaluated at $q_i = q_i^0$. Then in the matrix notation of (2.123), 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

$$F\boldsymbol{\mu} = \lambda^2\boldsymbol{\mu} \quad \Rightarrow \qquad V\boldsymbol{\mu} = -\lambda^2\, T\boldsymbol{\mu} \tag{2.129}$$

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 $\bar{\boldsymbol{\mu}}$. We have $\bar{\boldsymbol{\mu}}\cdot V\boldsymbol{\mu} = \lambda^2\,\bar{\boldsymbol{\mu}}\cdot T\boldsymbol{\mu}$. But for any symmetric matrix S, the quantity $\bar{\boldsymbol{\mu}}\cdot S\boldsymbol{\mu}$ is real. (This follows from expanding $\boldsymbol{\mu}$ in the complete set of real, orthogonal eigenvectors of S, each of which has a real eigenvalue). Therefore both $\bar{\boldsymbol{\mu}}\cdot V\boldsymbol{\mu}$ and $\bar{\boldsymbol{\mu}}\cdot T\boldsymbol{\mu}$ are both real. Since we have assumed that T is invertible and positive definite, we know that $\bar{\boldsymbol{\mu}}\cdot T\boldsymbol{\mu} \neq 0$ so, from (2.129), we conclude that the eigenvalue λ^2 is indeed real.

2.6.1 Example: The Double Pendulum

In section 2.5.2, we derived the Lagrangian for the double pendulum. 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 (2.72) 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 dubious). 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

$$L \approx ml^2\dot{\theta}_1^2 + \tfrac{1}{2}ml^2\dot{\theta}_2^2 + ml^2\dot{\theta}_1\dot{\theta}_2 - mgl\theta_1^2 - \tfrac{1}{2}mgl\theta_2^2 \tag{2.130}$$

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

$$\begin{aligned} 2ml^2\ddot{\theta}_1 + ml^2\ddot{\theta}_2 &= -2mgl\theta_1 \\ ml^2\ddot{\theta}_2 + ml^2\ddot{\theta}_1 &= -mgl\theta_2 \end{aligned} \tag{2.131}$$

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

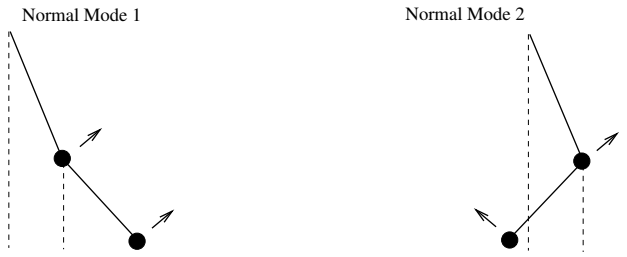

Figure 19: The two normal modes of the double pendulum

$$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \ddot{\boldsymbol{\theta}} = -\frac{g}{l} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \boldsymbol{\theta} \qquad \Rightarrow \qquad \ddot{\boldsymbol{\theta}} = -\frac{g}{l} \begin{pmatrix} 2 & -1 \\ -2 & 2 \end{pmatrix} \boldsymbol{\theta} \tag{2.132}$$

We have two eigenvectors. They are

1. $\boldsymbol{\mu}_1 = \begin{pmatrix} 1 \\ \sqrt{2} \end{pmatrix}$ which has eigenvalue $\lambda_1^2 = -(g/l)(2-\sqrt{2})$. This corresponds to the motion shown in figure 19 for the first normal mode.

2. $\boldsymbol{\mu}_2 = \begin{pmatrix} 1 \\ -\sqrt{2} \end{pmatrix}$ which has eigenvalue $\lambda_2^2 = -(g/l)(2+\sqrt{2})$. This corresponds to the motion shown in figure 19 for the second normal mode.

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 Example: The Linear Triatomic Molecule

Consider the molecule drawn in the figure. It's 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 object takes the form,

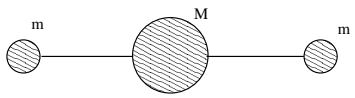

Figure 20:

$$L = \tfrac{1}{2}m\dot{x}_1^2 + \tfrac{1}{2}M\dot{x}_2^2 + \tfrac{1}{2}m\dot{x}_3^2 - V(x_1 - x_2) - V(x_2 - x_3) \tag{2.133}$$

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

$$x_i(t) = x_i^0 + \eta_i(t) \tag{2.134}$$

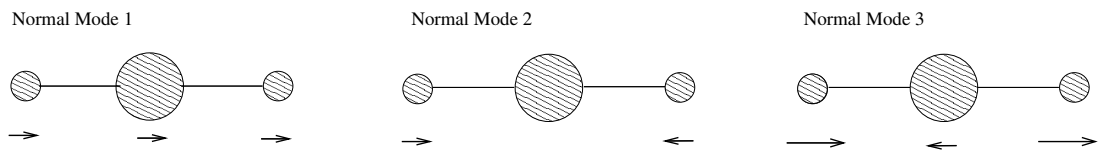

Figure 21: The three normal modes of the triatomic molecule

Taylor expanding the potential about the equilibrium point,

$$V(r) = V(r_0) + \left.\frac{\partial V}{\partial r}\right|_{r=r_0}(r-r_0) + \tfrac{1}{2}\left.\frac{\partial^2 V}{\partial r^2}\right|_{r=r_0}(r-r_0)^2 + \ldots \tag{2.135}$$

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

$$L \approx \tfrac{1}{2}m\dot{\eta}_1^2 + \tfrac{1}{2}M\dot{\eta}_2^2 + \tfrac{1}{2}m\dot{\eta}_3^2 - \frac{k}{2}\left[(\eta_1-\eta_2)^2 + (\eta_2-\eta_3)^2\right] \tag{2.136}$$

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

$$\begin{pmatrix} m\ddot{\eta}_1 \\ M\ddot{\eta}_2 \\ m\ddot{\eta}_3 \end{pmatrix} = -k\begin{pmatrix} \eta_1-\eta_2 \\ (\eta_2-\eta_1)+(\eta_2-\eta_3) \\ \eta_3-\eta_2 \end{pmatrix} \tag{2.137}$$

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

$$F = \begin{pmatrix} -k/m & k/m & 0 \\ k/M & -2k/M & k/M \\ 0 & k/m & -k/m \end{pmatrix} \tag{2.138}$$

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.

All motions are drawn in figure 21. For small deviations from equilibrium, the most general motion is a superposition of all of these modes.

$$\boldsymbol{\eta}(t) = \boldsymbol{\mu}_1(A + Bt) + \boldsymbol{\mu}_2 C\cos(\omega_2(t - t_2)) + \boldsymbol{\mu}_3 D\cos(\omega_3(t - t_3)) \tag{2.139}$$