Chapter 3

Lagrangian Mechanics

3.1 Action Principle

In this chapter we will see how the familiar laws of mechanics can be expressed and understood from a very different point of view, which is known as the Lagrangian formulation of mechanics. This is in many ways more elegant than the Newtonian formulation, and it is particularly useful when moving to quantum mechanics. For example, quantum field theories are usually studied in a Lagrangian framework.

The idea is similar to Fermat's principle in optics, according to which light follows the shortest optical path, i.e., the path of shortest time to reach its destination. As a reminder, let us see how Snell's law

$$\frac{\sin\theta_2}{\sin\theta_1} = \frac{n_1}{n_2}, \tag{3.1.1}$$

which tells how a light ray bends at the interface of two materials with refractive indices n_1 and n_2.

Consider a light ray from point (x_a, y_a) to (x_b, y_b). There is a horizontal interface at y, and between y_a and y, the refractive index is n_1 and between y and y_b it is n_2. We now assume that the light follows a straight line from (x_a, y_a) to a point (x, y) on the interface, and then from (x, y) to (x_b, y_b). The only unknown is therefore x. Because of the speed of light in medium is c/n, the optical path length is

$$S(x) = \int_a^b dt = \int_a^b \frac{n}{c} dl = \frac{n_1}{c}\sqrt{(x-x_a)^2 + (y-y_a)^2} + \frac{n_2}{c}\sqrt{(x_b-x)^2 + (y_b-y)^2}. \tag{3.1.2}$$

Note that we have made the simplification $S(x, y) \to S(x)$ since we are assuming straight lines i.e. $y = x$. According to Fermat's principle, we need to find the minimum of this quantity. At the minimum, the derivative with respect to x vanishes, so

$$0 = \frac{\partial S}{\partial x} = \frac{n_1}{c}\frac{x - x_a}{\sqrt{(x-x_a)^2 + (y-y_a)^2}} - \frac{n_2}{c}\frac{x_b - x}{\sqrt{(x_b-x)^2 + (y_b-y)^2}} = \frac{n_1}{c}\sin\theta_1 - \frac{n_2}{c}\sin\theta_2, \tag{3.1.3}$$

from which Snell's law (3.1.1) follows immediately.

Fermat had proposed his principle in 1657, and it motivated Maupertuis to suggest in 1746 that also matter particles would obey an analogous variational principle. He postulated that there exists a quantity called action, which the trajectory of the matter particle would minimise. This idea was later refined by Lagrange and Hamilton, who developed it into its current form, in which the action S is defined as an integral over a function $L(x, \dot{x}, t)$ known as the Lagrangian, as

$$S[x] = \int_{t_a}^{t_b} L(x(t), \dot{x}(t), t)dt. \tag{3.1.4}$$

The Lagrangian is a function of the position x and velocity $\dot{x}$ of the particle, and it may also have some explicit time dependence. We will see later that for conservative systems, the Lagrangian is simply the difference of the kinetic energy T and the potential energy V, i.e., $L = T - V$.

Because the action S is given by an integral over time, it depends on the position and velocity at all times, i.e., on the whole trajectory of the particle. It is therefore a function from the space of functions $x(t)$ to real numbers, and we indicate that by having the argument (i.e. function x) in square brackets. Such function of functions are called functionals.

Given a Lagrangian L, the dynamics is determined by Hamilton's principle (or action principle), which states that *to move from position x_a at time t_a to position x_b at time t_b, the particle follows the trajectory that minimises the action $S[x]$*. In other words, the actual physical trajectory is the function x that minimises the action subject to the boundary conditions $x(t_a) = x_a$ and $x(t_b) = x_b$.

To find this minimising function $x(t)$, we want to calculate the derivative of the action $S[x]$ with respect to the function $x(t)$ and set it to zero. Functional derivatives such as this are studied in the branch of mathematics known as functional analysis. However, here we adopt a slightly simpler approach and consider small variations of the trajectory. This is known as *variational calculus*.

Let us assume that $x(t)$ is the function that minimises the action, and consider a slightly perturbed trajectory

$$\tilde{x}(t) = x(t) + \delta x(t), \tag{3.1.5}$$

where we assume that the perturbation is infinitesimally small and vanishes at the endpoints,

$$\delta x(t_a) = \delta x(t_b) = 0. \tag{3.1.6}$$

This perturbation changes the action by

$$\begin{aligned}
\delta S &= S[x + \delta x] - S[x] \\
&= \int_{t_a}^{t_b} \left[L(x(t) + \delta x(t), \dot{x}(t) + \delta\dot{x}(t), t) - L(x(t), \dot{x}(t), t)\right] dt \\
&= \int_{t_a}^{t_b} \left[\frac{\partial L}{\partial x}\delta x(t) + \frac{\partial L}{\partial \dot{x}}\delta\dot{x}(t)\right] dt \\
&= \int_{t_a}^{t_b} \left[\frac{\partial L}{\partial x}\delta x(t) + \frac{\partial L}{\partial \dot{x}}\frac{d\delta x(t)}{dt}\right] dt \\
&= \int_{t_a}^{t_b} \frac{\partial L}{\partial x}\delta x(t)dt + \left[\frac{\partial L}{\partial \dot{x}}\delta x(t)\right]_{t_a}^{t_b} - \int_{t_a}^{t_b} \left[\frac{d}{dt}\frac{\partial L}{\partial \dot{x}}\right] \delta x(t)dt,
\end{aligned} \tag{3.1.7}$$

where we Taylor expanded to linear order in $\delta x(t)$ and integrated the second term by parts. Because of the boundary conditions (3.1.6), the substitution term vanishes, and we have

$$\delta S = \int_{t_a}^{t_b} \left[\frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}} \right] \delta x(t) dt. \tag{3.1.8}$$

For $x(t)$ to be minimum, the variation of the action (3.1.8) has to vanish for any function $\delta x(t)$. You can see this by noting that if $\delta S < 0$ for any perturbation $\delta x(t)$, then $S[x + \delta x] < S[x]$. Correspondingly, if $\delta S > 0$, then $S[x - \delta x] < S[x]$. In either case, we have found a function that has lower action than $x(t)$. Therefore $x(t)$ can only be the minimum if $\delta S = 0$.

The only way we can have $\delta S = 0$ for every perturbation $\delta x(t)$ is that the expression inside the brackets in Eq. (3.1.8) vanishes, i.e.,

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = 0. \tag{3.1.9}$$

This is known as the *Euler-Lagrange equation*, and it is the equation of motion in the Lagrangian formulation of mechanics. When using Eq. (3.1.9), it is very important to understand the difference between the partial (∂) and total (d) derivatives.

To check that Eq. (3.1.9) really describes the same physics as Newtonian mechanics, let us consider a simple example of a particle in a one-dimensional potential $V(x)$. Because the system is conservative, the Lagrangian is

$$L = T - V = \frac{1}{2} m \dot{x}^2 - V(x), \tag{3.1.10}$$

and the Euler-Lagrange equation is

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = \frac{d}{dt}(m\dot{x}) + \frac{\partial V}{\partial x} = m\ddot{x} + \frac{\partial V}{\partial x} = 0. \tag{3.1.11}$$

This is nothing but Newton's second law

$$m\ddot{x} = -\frac{\partial V}{\partial x}. \tag{3.1.12}$$

It is interesting to note that although Newtonian and Lagrangian formulations of mechanics are mathematically equivalent and describe the same physics, their starting point is very different. Newton's laws describe the evolution of the system as an initial value problem: We know the position and velocity of the particle at the initial time, $x(t_a)$ and $\dot{x}(t_a)$, and we then use Newton's laws to determine the evolution $x(t)$ at later times $t > t_a$.

In contrast, the Lagrangian formulation describes the same physics as a boundary value problem. We know the initial and final positions of the particle $x(t_a)$ and $x(t_b)$, and we use the action principle to determine $x(t)$ for $t_a < t < t_b$, i.e., how the particle travels from one to the other. In particular, we cannot choose the initial velocity because it is determined by the final destination of the particle through the action principle. This may appear very non-local in time because the behaviour of the particle in the far future determines its motion at the current time. However,

because because the two formulations are equivalent, this apparent non-locality in time does not actually affect the physics. For example, it is not possible to use it to send information back in time. In practice, it is usually easier to solve initial value problems, and therefore one usually uses the Lagrangian formulation to set up the problem and derive the equations of motion but then solves them as an initial value problem.

In many ways the Lagrangian formulation is closer to quantum mechanics, which does not allow one to determine the initial position and velocity of the particle either. Furthermore, the principle that the particle chooses one out of all possible trajectories resembles the double slit experiment in quantum mechanics, with the key difference that in the quantum case one has to sum over all possible trajectories rather than just selecting one. This correspondence turns out to be fully accurate and becomes obvious in the path integral formulation in quantum mechanics.

3.2 Generalised Coordinates

One attractive aspect of the Lagrangian formulation is that it is independent of the variables that are used to describe the state of the system. This is because the minimum of the function does not depend on the coordinate system, and the same applies to a functional such as the action S. Therefore, in contrast with Newtonian mechanics, we do not have to use the Cartesian position coordinates, and the Euler-Lagrange equation still has the same form (3.1.9). Instead, we are free to choose whichever set of variables we want to parameterise the state of the system, and which are then called *generalised coordinates* and usually denoted by q. They can be position coordinates, but also angles etc.

Usually we need more than one generalised coordinate, which we label by index i, so that we have some number N generalised coordinates q_i, with $i = 1, \ldots, N$. Each coordinate satisfies the corresponding Euler-Lagrange equation

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0. \tag{3.2.1}$$

As the first example of generalised coordinates, let us consider a simple pendulum that has a mass m at the end of a light rod of fixed length l. The angle of the pendulum from the vertical position is θ, which we choose as the generalised coordinate $q = \theta$. The Lagrangian is[1]

$$L = T - V = \frac{1}{2}ml^2\dot{\theta}^2 + mgl(1 - \cos\theta). \tag{3.2.2}$$

The Euler-Lagrange equation is

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\theta}} - \frac{\partial L}{\partial \theta} = \frac{d}{dt}\left(ml^2\dot{\theta}\right) + mgl\sin\theta = ml^2\ddot{\theta} + mgl\sin\theta = 0, \tag{3.2.3}$$

from which we find

$$\ddot{\theta} = -\frac{g}{l}\sin\theta. \tag{3.2.4}$$

[1] Equation has been corrected.

As a slightly more complex example, let us consider the motion of a particle in a central potential $V(r)$ in three dimensions. We use the spherical coordinates (r, θ, ϕ) as the generalised coordinates. The Lagrangian is

$$L = T - V = \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\dot{\phi}^2\sin^2\theta\right) - V(r). \tag{3.2.5}$$

The Euler-Lagrange equation for r is

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{r}} - \frac{\partial L}{\partial r} = \frac{d}{dt}\left(m\dot{r}\right) - mr\left(\dot{\theta}^2 + \dot{\phi}^2\sin^2\theta\right) + \frac{dV}{dr} = 0, \tag{3.2.6}$$

which gives

$$m\ddot{r} = mr\left(\dot{\theta}^2 + \dot{\phi}^2\sin^2\theta\right) - \frac{dV}{dr}. \tag{3.2.7}$$

The second term on the right hand side is the force due to the potential, and the first term is the centrifugal force.

The Euler-Lagrange equation for θ is

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\theta}} - \frac{\partial L}{\partial \theta} = \frac{d}{dt}\left(mr^2\dot{\theta}\right) - mr^2\dot{\phi}^2\sin\theta\cos\theta = 0. \tag{3.2.8}$$

Finally, because ϕ does not appear in the Lagrangian (except as a time derivative $\dot{\phi}$), the Euler-Lagrange equation for ϕ is

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\phi}} - \frac{\partial L}{\partial \phi} = \frac{d}{dt}\frac{\partial L}{\partial \dot{\phi}} = 0. \tag{3.2.9}$$

This means that the quantity

$$\frac{\partial L}{\partial \dot{\phi}} = mr^2\sin^2\theta\dot{\phi} \tag{3.2.10}$$

is conserved. We note that this is simply L_z, the z component of the angular momentum vector $\mathbf{L} = m\mathbf{r} \times \dot{\mathbf{r}}$.

We can easily see that this is, in fact, a very general result. For any generalised coordinate q_i, we define the generalised momentum p_i by

$$p_i = \frac{\partial L}{\partial \dot{q}_i}. \tag{3.2.11}$$

The Euler-Lagrange equation implies that whenever the Lagrangian does not depend on q_i, then the corresponding generalised momentum p_i is conserved. This is a simple example of a more general result known as Noether's theorem, which we will come back to later.

As a very simple example, let us consider the Cartesian position coordinate x as our generalised coordinate. With the Lagrangian $L = \frac{1}{2}m\dot{x}^2 - V(x)$, the generalised momentum is simply the conventional momentum $p = \partial L/\partial \dot{x} = m\dot{x}$.

3.3 Precession of a Symmetric Top

Let us now consider an example of the use of Lagrangian mechanics to solve a real problem: a symmetric top. The kinetic energy is given in terms of Euler angles by Eq. (2.8.4) whereas the gravitational potential energy is $V = MgR\cos\theta$. This leaves us with a Lagrangian

$$L = \tfrac{1}{2}I_1\dot{\phi}^2\sin^2\theta + \tfrac{1}{2}I_1\dot{\theta}^2 + \tfrac{1}{2}I_3\left(\dot{\psi}+\dot{\phi}\cos\theta\right)^2 - MgR\cos\theta\,. \tag{3.3.1}$$

The Euler-Lagrange equation (3.2.1) for θ is

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(I_1\dot{\theta}\right) = I_1\dot{\phi}^2\sin\theta\cos\theta - I_3\left(\dot{\psi}+\dot{\phi}\cos\theta\right)\dot{\phi}\sin\theta + MgR\sin\theta\,. \tag{3.3.2}$$

The Lagrangian function (3.3.1) does not contain the other two Euler angles ϕ and ψ so the generalised momenta $p_\phi = \partial L/\partial\dot{\phi}$ and $p_\psi = \partial L/\partial\dot{\psi}$ are constant

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[I_1\dot{\phi}\sin^2\theta + I_3\left(\dot{\psi}+\dot{\phi}\cos\theta\right)\cos\theta\right] = 0 \tag{3.3.3}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[I_3\left(\dot{\psi}+\dot{\phi}\cos\theta\right)\right] = 0\,. \tag{3.3.4}$$

Note that comparison of Eqs. (3.3.4) and (2.8.2) tells us that

$$\omega_3 = \dot{\psi}+\dot{\phi}\cos\theta = \text{constant}. \tag{3.3.5}$$

We are interested in the situation of steady precession at a constant angle θ. In this case we conclude from Eqs. (3.3.3) and (3.3.4) that $\dot{\phi}$ and $\dot{\psi}$ are constant. Hence the axis of the top precesses around the vertical with a constant angular velocity, which we denote by Ω, i.e., $\dot{\phi} = \Omega$. Because we are looking for a solution with fixed θ, the left side of Eq. (3.3.2) must vanish, and we obtain

$$I_1\Omega^2\cos\theta - I_3\omega_3\Omega + MgR = 0\,, \tag{3.3.6}$$

which we can solve for Ω. The general solution is

$$\Omega = \frac{I_3\omega_3 \pm \sqrt{I_3^2\omega_3^2 - 4I_1MgR\cos\theta}}{2I_1\cos\theta}\,. \tag{3.3.7}$$

This only has real roots if

$$\omega_3^2 \geq \omega_c^2 \equiv \frac{4I_1MgR\cos\theta}{I_3^2}\,. \tag{3.3.8}$$

If the top is spinning more slowly than this, there is no solution with constant θ. Instead, the top starts to wobble.

For a rapidly spinning top, $\omega_3 \gg \omega_c$, we can expand the square root in Eq. (3.3.7) to obtain

$$\Omega \approx \frac{I_3\omega_3 \pm I_3\omega_3\left(1 - 2\frac{I_1MgR\cos\theta}{I_3^2\omega_3^2}\right)}{2I_1\cos\theta} \to \begin{cases} MgR/I_3\omega_3 & -\text{sign} \\ I_3\omega_3/I_1\cos\theta & +\text{sign} \end{cases} \tag{3.3.9}$$

The first of these is the precession frequency calculated in (2.6.6), while the second is the precession of a free system discussed in section 2.8. Note the absence of any contribution from gravity in the second expression.

Note that if $\theta > \frac{1}{2}\pi$, the top is hanging *below* its point of support, and there is no limit on ω_3. In particular, for $\omega_3 = 0$, we find the possible angular velocities of a compound pendulum swinging in a circle

$$\Omega = \pm\sqrt{\frac{MgR}{I_1 \left|\cos\theta\right|}} \tag{3.3.10}$$

3.4 Constraints

Consider a system of N particles in three dimensions. To specify the position of each particle, you need $3N$ generalised coordinates. However, in many cases the coordinates are not all independent but subject to some constraints, such as the rigidity conditions (see Section 2.2), which reduce the number of generalised coordinates required. For a rigid body, the original $3N$ coordinates may be reduced to six generalised coordinates: three translational, such as the coordinates X, Y, Z of the centre of mass, and three rotational, such as the Euler angles, ϕ, θ, ψ.

We will assume that the constraints can be written in the form $f(x_1, \ldots, x_{3N}, t) = 0$. Constraints like that are called *holonomic*. For a rigid body, the constraints are of this form: The distance between each pair of particles i and j is fixed to a constant d_{ij}, and therefore one has

$$(\mathbf{r}_i - \mathbf{r}_j)^2 - d_{ij}^2 = 0. \tag{3.4.1}$$

Another example is motion on the surface of a sphere of radius R, for which the constraint is

$$f(x, y, z) = x^2 + y^2 + z^2 - R^2 = 0. \tag{3.4.2}$$

Sometimes one has to deal with non-holonomic constraints, for example if the constraint depends on velocities, but these are more complicated to handle, and we will not discuss them in this course.

In principle, solving each constraint equation eliminates one coordinate. If one has initially N coordinates x_i (with $i = 1, \ldots, N$) and C constraints, solving them will allow one to express the original coordinates in terms of $M = N - C$ generalised coordinates q_j, with $j = 1, \ldots, M$,

$$x_i = x_i(q_1, \ldots, q_M, t), \tag{3.4.3}$$

with possibly explicit time-dependence if the constraints are time-dependent. In that case the system is called *forced*, otherwise it is *natural*.

If one can solve the constraints and find the explicit relations (3.4.3), one can then write the Lagrangian in terms of the generalised coordinates q_j and solve the Euler-Lagrange equation. Substituting this solution to Eq. (3.4.3) then gives the solution in terms of the original coordinates.

An alternative approach, which is sometimes useful, is to implement the constraints using *Lagrange multipliers*. Starting with a Lagrangian $L(x_1, \ldots, x_N)$ and a constraint function $f(x_1, \ldots, x_N)$, we define a new Lagrangian L'

$$L'(x_1, \ldots, x_N, \lambda) = L(x_1, \ldots, x_N) + \lambda f(x_1, \ldots, x_N), \tag{3.4.4}$$

which is a function of the original coordinates and a Lagrange multiplier λ. If we now treat λ as the $(N+1)$th coordinate, its Euler-Lagrange equation is

$$\frac{d}{dt}\frac{\partial L'}{\partial \dot{\lambda}} - \frac{\partial L'}{\partial \lambda} = -f(x_1, \ldots, x_N) = 0, \tag{3.4.5}$$

and therefore it satisfies the constraint automatically. The Euler-Lagrange equations for the original coordinates are

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}_i} - \frac{\partial L}{\partial x_i} - \lambda(t)\frac{\partial f}{\partial x_i} = 0, \tag{3.4.6}$$

where the extra term can be interpreted as the (generalised) force that has to be applied to the system to enforce the constraint.

As an example, consider a mass m hanging from a rope that is wrapped around a pulley of radius R and moment of inertia I. Using the vertical position z of the mass, and the angle θ of the pulley as the coordinates, the Lagrangian is

$$L = T - V = \frac{1}{2}m\dot{z}^2 + \frac{1}{2}I\dot{\theta}^2 - mzg. \tag{3.4.7}$$

If the rope does not slip, the pulley has to rotate as the mass moves, and this imposes a constraint $R\dot{\theta} = -\dot{z}$. Choosing the origin appropriately, we can write this as a holonomic constraint

$$f(\theta, z) = R\theta + z = 0. \tag{3.4.8}$$

Introducing the Lagrange multiplier λ, the new Lagrangian is

$$L' = L + \lambda(R\theta + z) = \frac{1}{2}m\dot{z}^2 + \frac{1}{2}I\dot{\theta}^2 - mzg + \lambda(R\theta + z). \tag{3.4.9}$$

The Euler-Lagrange equations are

$$\begin{aligned} \frac{d}{dt}m\dot{z} + mg - \lambda &= 0 \quad \text{for } z, \\ \frac{d}{dt}I\dot{\theta} - \lambda R &= 0 \quad \text{for } \theta, \\ -(R\theta + z) &= 0 \quad \text{for } \lambda. \end{aligned} \tag{3.4.10}$$

The third equation implements the constraint $\theta = -z/R$, and substituting this to the first two gives

$$\begin{aligned} m\ddot{z} + mg - \lambda &= 0, \\ -\frac{I}{R}\ddot{z} - \lambda R &= 0. \end{aligned} \tag{3.4.11}$$

Solving this pair of equations for $\ddot{z}$ and λ, we obtain

$$\begin{aligned}\left(m + \frac{I}{R^2}\right)\ddot{z} + mg &= 0, \\ \lambda &= \frac{mg}{1 + mR^2/I}.\end{aligned} \tag{3.4.12}$$

The first line shows that the moment of inertia I of the pulley gives the mass extra inertia. The second line gives the force that rope has to apply in order to enforce the constraint. This is just the tension of the rope.

3.5 Normal Modes

3.5.1 Orthogonal Coordinates

Instead of rigid constraints, let us now consider a situation where the constraints are flexible so that the particles can move around their equilibrium positions. We assume that the system is described by N generalised coordinates q_i. We also assume that it is *natural*, which means that the kinetic energy is a quadratic homogeneous function of the generalised velocities. We can then write it as

$$T = \frac{1}{2}\sum_{ij} a_{ij}(q_1, \ldots, q_N)\dot{q}_i\dot{q}_j, \tag{3.5.1}$$

where the coefficients a_{ij} can depend on the coordinates q_i but not on velocities $\dot{q}_i$. They can chosen to be symmetric ($a_{ji} = a_{ij}$) without any loss of generality.

The coordinates are said to be *orthogonal* if there are no cross terms, i.e., $a_{ij} = 0$ if $i \neq j$. Then the kinetic energy is simply

$$T = \frac{1}{2}\sum_{i} a_{ii}(q_1, \ldots, q_N)\dot{q}_i^2. \tag{3.5.2}$$

We can always make our coordinates orthogonal by using the *Gram-Schmidt* procedure. For example, if $N = 2$, the general form of Eq. (3.5.1) is

$$T = \tfrac{1}{2}a_{11}\dot{q}_1^2 + a_{12}\dot{q}_1\dot{q}_2 + \tfrac{1}{2}a_{22}\dot{q}_2^2\,. \tag{3.5.3}$$

Defining a new coordinate

$$q_1' = q_1 + \frac{a_{12}}{a_{11}}q_2\,, \tag{3.5.4}$$

the kinetic energy becomes

$$T = \tfrac{1}{2}a_{11}\dot{q}_1'^2 + \tfrac{1}{2}a_{22}'\dot{q}_2^2 \qquad \text{with} \quad a_{22}' = a_{22} - \frac{a_{12}^2}{a_{11}}\,. \tag{3.5.5}$$

Furthermore, if we rescale the coordinates in Eq. (3.5.2) by

$$q_i' = \sqrt{a_{ii}}q_i, \tag{3.5.6}$$

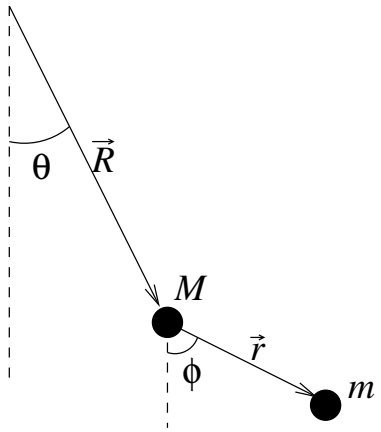

Figure 3.1: Double pendulum

the kinetic energy becomes

$$T = \frac{1}{2}\sum_i \dot{q}_i'^2. \tag{3.5.7}$$

Therefore we can always assume that the kinetic energy has this form.

As an example, consider a double pendulum, with a second pendulum hanging from the first. The kinetic energy is

$$T = \frac{1}{2}M\dot{\mathbf{R}}^2 + \frac{1}{2}m\left(\dot{\mathbf{R}} + \dot{\mathbf{r}}\right)^2 = \tfrac{1}{2}MR^2\dot{\theta}^2 + \tfrac{1}{2}m\left[R^2\dot{\theta}^2 + r^2\dot{\phi}^2 + 2Rr\dot{\theta}\dot{\phi}\cos(\theta-\phi)\right] \tag{3.5.8}$$

where M, R and θ refer to the upper pendulum and m, r and ϕ to the lower. Note that the kinetic energy of the lower pendulum depends not only on ϕ but also on the motion of the upper pendulum to which it is attached. For small values of θ and ϕ we can set the cosine term to one, so that the kinetic energy is

$$T = \tfrac{1}{2}(M+m)R^2\dot{\theta}^2 + \tfrac{1}{2}mr^2\dot{\phi}^2 + mRr\dot{\theta}\dot{\phi}. \tag{3.5.9}$$

However, because of the last term, the coordinates are not orthogonal. In this case it is obvious that we could get an orthogonal set by simply considering the displacement of the 2 bobs. For small angles this gives

$$x = R\theta \qquad y = R\theta + r\phi\,, \tag{3.5.10}$$

and the kinetic energy, T, becomes

$$T = \tfrac{1}{2}M\dot{x}^2 + \tfrac{1}{2}m\dot{y}^2\,. \tag{3.5.11}$$

Finally, we can reduce this to the standard form (3.5.7) using

$$q_1 = \sqrt{M}x \qquad q_2 = \sqrt{m}y\,. \tag{3.5.12}$$

3.5.2 Small Oscillations

We now consider the potential energy, V. If T is given by (3.5.7), the Lagrangian is

$$L = T - V = \frac{1}{2}\sum_i \dot{q}_i'^2 - V(q_1, \ldots, q_N). \tag{3.5.13}$$

The Euler-Lagrange equations are then simply

$$\ddot{q}_i = -\frac{\partial V}{\partial q_i}\,, \tag{3.5.14}$$

for all i.

If we now assume that the amplitude of the oscillations is small, we can Taylor expand the potential around the origin, which was chosen to correspond to the equilibrium state. To quadratic order we have

$$V(q_1, \ldots, q_N) = V_0 + \sum_i b_i q_i + \tfrac{1}{2}\sum_{ij} k_{ij} q_i q_j + O(q^3), \tag{3.5.15}$$

where V_0, b_i and k_{ij} are constants.

Because we can subtract a constant from the potential without changing the equations of motion (3.5.14), we are free to choose $V_0 = 0$. Because we are assuming that the equilibrium state is $q_i = 0$ for all i, we find that $b_i = 0$. Therefore, we only need to consider the quadratic term

$$V = \tfrac{1}{2}\sum_{ij} k_{ij} q_i q_j, \tag{3.5.16}$$

Note that we can also choose k_{ij} to be symmetric, i.e., $k_{ji} = k_{ij}$ without any loss of generality.

With the potential (3.5.16), the Euler-Lagrange equation (3.5.14) is simply

$$\ddot{q}_i = -\sum_j k_{ij} q_j, \tag{3.5.17}$$

which can be written in matrix form as

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\begin{pmatrix} q_1 \\ q_2 \\ \vdots \\ q_N \end{pmatrix} = -\begin{pmatrix} k_{11} & k_{12} & \cdots & k_{1N} \\ k_{21} & k_{22} & \cdots & k_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ k_{N1} & k_{N2} & \cdots & k_{NN} \end{pmatrix}\begin{pmatrix} q_1 \\ q_2 \\ \vdots \\ q_N \end{pmatrix}. \tag{3.5.18}$$

By defining an N-dimensional coordinate vector $\mathbf{q} = (q_1, \ldots, q_N)$ and an $N \times N$ matrix $\overline{\overline{\mathbf{k}}}$ with elements k_{ij}, we can also write it more compactly as

$$\ddot{\mathbf{q}} = -\overline{\overline{\mathbf{k}}} \cdot \mathbf{q}. \tag{3.5.19}$$

In the same notation, the Lagrangian is

$$L = \frac{1}{2}\dot{\mathbf{q}} \cdot \dot{\mathbf{q}} - \frac{1}{2}\mathbf{q} \cdot \overline{\overline{\mathbf{k}}} \cdot \mathbf{q}. \tag{3.5.20}$$

In our double pendulum example, the potential is

$$\begin{aligned} V(\theta, \phi) &= MgR(1-\cos\theta) + mg\left[R(1-\cos\theta) + r(1-\cos\phi)\right] \\ &\approx \frac{1}{2}(M+m)gR\theta^2 + \frac{1}{2}mgr\phi^2 \\ &= \frac{M+m}{2R}gx^2 + \frac{mg}{2r}(y-x)^2 \\ &= \frac{1}{2}\left(1 + \frac{m}{M}\frac{R+r}{r}\right)\frac{g}{R}q_1^2 + \frac{g}{2r}q_2^2 - \sqrt{\frac{m}{M}}\frac{g}{r}q_1 q_2. \end{aligned} \tag{3.5.21}$$

The coefficient matrix is therefore

$$\overline{\overline{\mathbf{k}}} = \begin{pmatrix} \left(1 + \frac{m}{M}\frac{R+r}{r}\right)\frac{g}{R} & -\sqrt{\frac{m}{M}}\frac{g}{r} \\ -\sqrt{\frac{m}{M}}\frac{g}{r} & \frac{g}{r} \end{pmatrix}. \tag{3.5.22}$$

3.5.3 Eigenvalue Problem

Eq. (3.5.19) is a set of N coupled linear second-order equation. Therefore we should find $2N$ linearly independent solutions.

We look for solutions of the form

$$\mathbf{q}(t) = \mathbf{A}e^{i\omega t}, \tag{3.5.23}$$

where $\mathbf{A}$ and ω are constants. We will check later that we have found all $2N$ solutions.

Substituting the Ansatz (3.5.23) into Eq. (3.5.19) gives

$$-\omega^2 \mathbf{A}e^{i\omega t} = -\overline{\overline{\mathbf{k}}} \cdot \mathbf{A}e^{i\omega t}, \tag{3.5.24}$$

which is equivalent to

$$\overline{\overline{\mathbf{k}}} \cdot \mathbf{A} = \omega^2 \mathbf{A}. \tag{3.5.25}$$

This has the form of the eigenvalue equation: It shows that $\mathbf{A}$ is an eigenvector of the matrix $\overline{\overline{\mathbf{k}}}$ with eigenvalue ω^2.

We know from linear algebra that the eigenvalues of a matrix are given by solutions of the characteristic equation

$$\det(\overline{\overline{\mathbf{k}}} - \omega^2\overline{\overline{\mathbf{1}}}) \equiv \begin{vmatrix} k_{11}-\omega^2 & k_{12} & \cdots & k_{1N} \\ k_{21} & k_{22}-\omega^2 & \cdots & k_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ k_{N1} & k_{N2} & \cdots & k_{NN}-\omega^2 \end{vmatrix} = 0. \tag{3.5.26}$$

Once we have found the eigenvalue ω^2, we can substitute it to Eq. (3.5.25) to find the eigenvector.

Because $\overline{\overline{\mathbf{k}}}$ is a symmetric $N \times N$ matrix, it has N real eigenvalues ω_α^2, where $\alpha = 1, \ldots, N$, corresponding to N eigenvectors $\mathbf{A}_\alpha$, which we can choose to be orthonormal,

$$\mathbf{A}_\alpha \cdot \mathbf{A}_\beta = \delta_{\alpha\beta}. \tag{3.5.27}$$

For each eigenvalue ω_α^2, there are two independent solutions,

$$\mathbf{q}_\alpha^+(t) = \mathbf{A}_\alpha e^{i\omega_\alpha t} \quad \text{and} \quad \mathbf{q}_\alpha^-(t) = \mathbf{A}_\alpha e^{-i\omega_\alpha t}. \tag{3.5.28}$$

For N eigenvalues, this takes the total number of linearly independent solutions to $2N$, proving that we have the complete solution.

The frequencies ω_α are real because the eigenvalues ω_α^2 have to be non-negative: If we had $\omega_\alpha^2 < 0$, then for the corresponding eigenvector $\mathbf{A}_\alpha$ and small ϵ we would have

$$V(\epsilon \mathbf{A}_\alpha) = \frac{1}{2}\epsilon^2 \mathbf{A}_\alpha \cdot \overline{\overline{\mathbf{k}}} \cdot \mathbf{A}_\alpha = \frac{1}{2}\epsilon^2 \omega_\alpha^2 < 0, \tag{3.5.29}$$

meaning that $\mathbf{q} = 0$ could not be the minimum of the potential as we assumed.

The individual solutions $\mathbf{q}_\alpha^\pm$ are complex, so physical solutions have to be real linear combinations of them,

$$\begin{aligned} \mathbf{q}_\alpha(t) &= a_+ \mathbf{q}_\alpha^+(t) + a_- \mathbf{q}_\alpha^-(t) = \left(a_+ e^{i\omega_\alpha t} + a_- e^{-i\omega_\alpha t}\right) \mathbf{A}_\alpha \\ &= \left[(a_+ + a_-)\cos\omega_\alpha t + i(a_+ - a_-)\sin\omega_\alpha t\right] \mathbf{A}_\alpha. \end{aligned} \tag{3.5.30}$$

Because we want a real solution, the coefficients $a_1 \equiv (a_+ + a_-)$ and $a_2 = i(a_+ - a_-)$ have to be real, and we can equally well write

$$\mathbf{q}_\alpha(t) = (a_1 \cos\omega_\alpha t + a_2 \sin\omega_\alpha t)\, \mathbf{A}_\alpha, \tag{3.5.31}$$

which we can also write a pure cosine term with a phase shift,

$$\mathbf{q}_\alpha(t) = c\cos(\omega_\alpha t + \phi)\mathbf{A}_\alpha, \tag{3.5.32}$$

where c and ϕ are real constants that have to be determined from the initial conditions.

Finally, the general solution is a linear combination of solutions of the form (3.5.32),

$$\mathbf{q}(t) = \sum_{\alpha=1}^{N} c_\alpha \cos(\omega_\alpha t + \phi_\alpha)\mathbf{A}_\alpha. \tag{3.5.33}$$

The modes of vibration of the system, i.e., the individual solutions (3.5.32) are known as *normal modes.* For a forced system there are resonances at the frequencies of the normal modes.

It is often useful to use the normal modes to define a set of generalised coordinates known as the *normal coordinates*, which we denote by $\tilde{q}_\alpha$. They are defined by expressing the original coordinates q_i in terms of the eigenvectors $\mathbf{A}_\alpha$ as

$$\mathbf{q} = \sum_\alpha \tilde{q}_\alpha \mathbf{A}_\alpha. \tag{3.5.34}$$

Substituting this to Eq. (3.5.20) we find that the Lagrangian becomes simply

$$L = \sum_\alpha \left(\frac{1}{2}\dot{\tilde{q}}_\alpha^2 - \frac{1}{2}\omega_\alpha^2 \tilde{q}_\alpha^2 \right). \tag{3.5.35}$$

The Euler-Lagrange equations are

$$\ddot{\tilde{q}}_\alpha + \omega_\alpha^2 \tilde{q}_\alpha = 0, \tag{3.5.36}$$

which means that each normal coordinate $\tilde{q}_\alpha$ oscillates independently of all others with its own normal frequency ω_α.

3.6 Continuous Systems

In addition to mechanical systems consisting of a finite number of degrees of freedom, one is often also interested in continuous systems, for example waves propagation in continuous media or field theories describing particle physics or electromagnetism.

To see how continuous systems are described in the Lagrangian formulation, consider a stretched string. We assume that in equilibrium the string is stretched to length ℓ_0 and has tension k. The displacement of the string from its equilibrium position is given by the continuous function $y(x,t)$, where $x \in \{0, \ell_0\}$. To describe the time and space derivatives, we use the notation

$$\dot{y} = \frac{\partial y}{\partial t}, \quad y' = \frac{\partial y}{\partial x}. \tag{3.6.1}$$

The Lagrangian is still given by the difference of the kinetic and potential energies, $L = T - V$. The kinetic energy can be calculated by considering an infinitesimal segment of length dx. The mass of such a segment is $dm = \mu dx$ where the constant μ is the mass per unit length. The velocity of the segment is simply $\dot{y}$, and therefore the kinetic energy of the infinitesimal segment is

$$dT = \frac{1}{2}\mu\, dx\, \dot{y}^2. \tag{3.6.2}$$

Integrating over the who distance ℓ_0, we find the total kinetic energy

$$T = \int_0^{\ell_0} dx\, \frac{1}{2}\mu\, \dot{y}^2. \tag{3.6.3}$$

The potential energy is V of the string is due to its tension k,

$$V = k(\ell - \ell_0), \tag{3.6.4}$$

where ℓ is the length of the displaced string. Again, this can be calculated by considering an infinitesimal segment of length dx. According to Pythagoras theorem, the length of the segment is

$$d\ell = \sqrt{dx^2 + dy^2} = \sqrt{1 + y'^2}dx \approx \left(1 + \frac{1}{2}y'^2\right) dx, \tag{3.6.5}$$

where we have assumed that the displacement is small and smooth so that $y' \ll 1$, and Taylor expanded to quadratic order. The length of the string is then obtained by summing over all the infinitesimal segments

$$\ell = \int_0^{\ell_0} dx \left(1 + \frac{1}{2}y'^2\right) = \ell_0 + \int_0^{\ell_0} dx \frac{1}{2}y'^2, \tag{3.6.6}$$

and therefore the potential energy is

$$V = \int_0^{\ell_0} dx \frac{1}{2}ky'^2. \tag{3.6.7}$$

We can now write the whole Lagrangian,

$$L = T - V = \int_0^{\ell_0} dx \left(\frac{1}{2}\mu\dot{y}^2 - \frac{1}{2}ky'^2\right). \tag{3.6.8}$$

The integrand is called the Lagrangian density and denoted by $\mathcal{L}$, i.e., $L = \int dx\mathcal{L}$, where

$$\mathcal{L} = \frac{1}{2}\mu\dot{y}^2 - \frac{1}{2}ky'^2. \tag{3.6.9}$$

The action is then given by an integral over both time and space,

$$S = \int dt \int dx \mathcal{L}(\dot{y}, y', y). \tag{3.6.10}$$

Variation of the action is

$$\delta S = \int dt \int dx \left[\frac{\partial\mathcal{L}}{\partial\dot{y}}\delta\dot{y} + \frac{\partial\mathcal{L}}{\partial y'}\delta y' + \frac{\partial\mathcal{L}}{\partial y}\delta y\right] = \int dt \int dx \left[-\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{y}} - \frac{d}{dx}\frac{\partial\mathcal{L}}{\partial y'} + \frac{\partial\mathcal{L}}{\partial y}\right]\delta y. \tag{3.6.11}$$

The action principle $\delta S = 0$ therefore leads to the Euler-Lagrange equation

$$\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{y}} + \frac{d}{dx}\frac{\partial\mathcal{L}}{\partial y'} - \frac{\partial\mathcal{L}}{\partial y} = 0. \tag{3.6.12}$$

Substituting the Lagrangian density (3.6.9) for the string, we find the equation of motion

$$\frac{d}{dt}(\mu\dot{y}) + \frac{d}{dx}(-ky') = \mu\ddot{y} - ky'' = 0, \tag{3.6.13}$$

which is the wave equation, as one would expect.