1 Classical Mechanics

These notes are designed to give an overview of some important topics in classical mechanics as briefly as possible, and are by no means complete. There is an emphasis on the underlying concepts and physical principles of the theory, which will serve as the necessary background for a solid understanding of field theory and quantum mechanics. Some prior knowledge is assumed, including a basic knowledge of calculus, and some familiarity with elementary concepts of mechanics. I recommend the books:

- L. D. Landau and E. M. Lifshitz, *Mechanics* (1976).
- H. Goldstein, C. Poole, and J. Safko, *Classical Mechanics* (2001).
- L. Susskind and G. Hrabovsky, *Classical Mechanics* (2014).

1.1 Principles of classical mechanics

The aim of classical mechanics is to describe physical systems, and determine their evolution through time. At the core of classical mechanics are a few basic principles:

1. Determinism: mechanical systems, and the laws of physics that govern them, are deterministic,

2. Homogeneity and isotropy: the laws of physics do not depend explicitly on time, position, or orientation,

3. Galilean relativity: the laws of physics are the same in all inertial reference frames.

A system is deterministic if, given enough information, it is possible (at least in theory) to completely determine its *future*, i.e., to specify the evolution of the system through time. Similarly, a system is considered *reversible* if the same holds in reverse, and we can completely determine its *past*. The assumption that the laws of nature are homogeneous means that they do not depend explicitly on *position*, or in other words, there is no preferred location in space. We also assume the laws are homogeneous in time, meaning they do not depend explicitly on time; this is usually wrapped into the homogeneity assumption. The assumption that the laws of nature are *isotropic* means that they do not depend explicitly on

orientation, or that there is no preferred direction in space. The third assumption is the (Galilean) *principle of relativity* – all motion is relative and the laws of nature do not depend explicitly on the choice of inertial reference frame, or, in other words, there is no special or universal rest frame. We shall return to the precise definition of an *inertial* reference frame soon.

We will focus our initial discussion on the dynamics of *particles*. By particle, we mean a body whose shape and dimension can be neglected in describing its motion. To define a system of particles in regular 3-dimensional space, we may designate each of their locations with a position vector, $\boldsymbol{x}$, with Cartesian coordinates x, y, z. For a system of N particles, there are $3N$ independent coordinate variables (called the *degrees of freedom*). To describe the *motion* of particles, we also consider the rate of change of position, called velocity, which we denote

$$\boldsymbol{v} \equiv \dot{\boldsymbol{x}} \equiv \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{x}.$$

We may also consider the rate of change of velocity (acceleration), $\boldsymbol{a} = \ddot{\boldsymbol{x}}$, and so on. As we shall see, however, the accelerations will be determined if we know the set of $3N$ coordinates $\{\boldsymbol{x}_i\}$, the set of $3N$ velocities $\{\dot{\boldsymbol{x}}_i\}$, and some function which defines the physical laws of the system, known as a *potential function* $V(\boldsymbol{x}, \dot{\boldsymbol{x}})$, the meaning of which will be discussed in the coming sections. The central determinism assumption of classical mechanics is that this information *completely* specifies the system of particles. If we know the $6N$ coordinate/velocity variables at any moment in time (called the *state* of the mechanical system), then we may determine the state of the system at any point in the future, assuming the potential function is known. Further, if the reversibility principle holds, it also means we can determine the state at any point in the past. The set of equations that specifies the relations for the coordinates and velocities of the system are called the *equations of motion*. We shall soon see that these are, in general, a set of second-order differential equations.

For general problems, the set of coordinates chosen need not be the Cartesian coordinates; in many situations, a different choice may be much more convenient. Any set of M quantities $(q_1, q_2, \ldots q_M)$ which completely defines the configuration of a system with M degrees of freedom are called the *generalised coordinates* of the system. It is possible that $M \leq 3N$, for example, when there are constraints on the motions of the system; this will be easier to discuss when the examples arise in the coming sections. Likewise, the set of their time derivatives, $\{\dot{q}_i\}$, are the *generalised velocities*.

We may consider the *phase space* of generalised positions and velocities.[1] We can then imagine the physical system tracing some path through this phase space

[1] Usually, the term phase space actually refers to the space of generalised positions q and *momentums* p, rather than velocities. The distinction is important, though not for our arguments here. We will return the concept of generalised momentum in the coming sections.

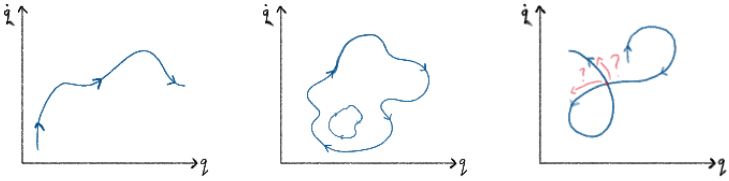

Figure 1.1: Two valid (deterministic and reversible) configuration paths, and one invalid path. (These are not meant to be realistic, just illustrative.)

as time progresses. The central assumptions of determinism and reversibility imply that paths in this phase space *may not cross*. If they could cross, then there would be two possible "futures" (or pasts) from the crossing point, which would violate determinism, and is thus not allowed; see Fig. 1.1. Though the paths cannot cross, it is completely possible to have *closed cycles* in the phase space. As we shall investigate further in the coming sections, the emergence of closed cycles implies some *conservation law* – i.e., it implies that there is some quantity that remains constant throughout the evolution of the system.

As a final note before we begin the study proper, we mention that the specific set of coordinates chosen (called a *reference frame*) is not unique. We can always change reference frames by shifting or rotating our coordinate axis: $\boldsymbol{x} \to \boldsymbol{x}'$. By the homogeneity and isotropy principles, this cannot impact the form of the physical laws or the outcome of any experiment. We may also make more general coordinate transformations that depend on time, in which case a little more care must be taken.

Consider a completely free particle, which is under the influence of no forces. If we chose a completely arbitrary reference frame to describe the particle, then of course, it may appear to be accelerating in our coordinates – for example, consider a reference frame that uniformly accelerates in one direction compared to another: the particle must have non-zero acceleration in at least one of these coordinate systems. However, if the free particle was to accelerate in a given coordinates, this would define a unique direction in space, and thus break the principle of isotropy. As such, the homogeneity and isotropy assumptions imply that there is always *some* particular reference frame in which a free particle does not accelerate. Such a frame is called an *inertial* reference frame.

Consider now a pair of reference frames K and K', specified by coordinates $\boldsymbol{x}$ and $\boldsymbol{x}'$ respectively, where K' moves relative to K with velocity $\boldsymbol{V}$. The relation linking the coordinate systems is:

$$\begin{aligned} \boldsymbol{x}' &= \boldsymbol{x} - \boldsymbol{V}t \\ t' &= t. \end{aligned} \tag{1.1}$$

These equations define a *Galilean transformation*. An inertial reference frame –

one in which free particles do not accelerate – may be linked to any other inertial reference frame by a suitable Galilean transformation. By the Galilean principle of relativity, such a transformation cannot impact the form of the physical laws or the outcome of any experiment. While there is no absolute frame of reference for coordinates, time is an absolute. The constancy of time between frames of reference is of one of the central assumptions in classical mechanics, though does not hold in relativistic mechanics (we will return to relativity in our study of classical field theory).

1.2 Principle of least action

The aim of classical mechanics is to describe the evolution of a system, described by the set of generalised coordinates $\{q_i\}$, and their derivatives $\{\dot{q}_i\}$. We invoke the first of our assumptions (determinism), and presume that there is a unique equation that describes this evolution.

Consider first the motion of a single particle that begins at some location, $q_0 = q(t_0)$, and some time later is found at location $q_f = q(t_f)$. By the determinism assumption, there must be a unique path $q(t)$ linking these points, which depends only on the initial conditions and properties of the system. Our task is to determine the *equations of motion* that uniquely determine this path. Since nature must always choose the same path given the same conditions, it seems reasonable to assume that the path the system will take will be *optimal*, with respect to some quantity (i.e., some quantity should be maximised or minimised for the correct physical solution). For example, we might guess that particles would take the *shortest* path (i.e., optimal with respect to length). It doesn't take much experimentation, however, to see that this is not the case (think of throwing a ball in the air – it certainly doesn't take the shortest path to its destination).

Without knowing anything about its form, we call the function to be optimised the *action*, denoted S, which depends on the set of coordinates and their derivatives:

$$S = S[\{q_i(t)\},\, \{\dot{q}_i(t)\}, t_0,\, t_f,\, t].$$

The action may depend *implicitly* on time through the time-dependence of the coordinates; in general, it may also depend *explicitly* on time. Note that the action is not a simple *local* function of coordinates; it depends on the values of q (and its derivatives) along the entire path of motion (see Fig. 1.2). From the principle of *locality* – that physical systems should be influenced only by their immediate surroundings – we expect the physical laws should ultimately be described by only local functions.[2] As an illustration, consider breaking the particle's path into a number of smaller connected paths between intermediate times, $t_0, t_1, t_2, \ldots, t_f$. Since each segment is itself a valid physical trajectory, the action for each segment, δS_i, must be independently optimised. At the same

[2] There's a short discussion on the principle of locality in Appendix 1.7.1.

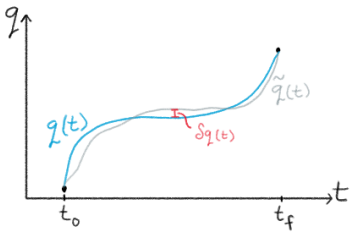

Figure 1.2: Path that minimises the action, $q(t)$, and its variation, $\delta q(t)$.

time, the action for the total path S must also be optimised. This is possible only if the total action is the sum of each smaller action:

$$S = \sum \delta S_i.$$

If we continuously break the paths into infinitesimal segments, then the infinitesimal action for each segment *can* be considered a local function of the position. As such, the action is itself typically written as the integral of another local function, known as the *Lagrangian*, L:

$$S = \int_{t_0}^{t_f} L(q, \dot{q}, t)\, \mathrm{d}t. \tag{1.2}$$

Aside from the locality argument, we may consider this simply a convention for now.[3] The task to completely specify the dynamics of the system, then, is to find the set of functions $q_i(t)$, which *optimise* the action. The overall sign of the action is arbitrary, and by convention we choose it so that the action is minimised.[4] This is called the *principle of least action*, or Hamilton's principle.

We shall now derive the set of equations that satisfy the principle of least action. For simplicity, we will work first in the case of a single particle, so there is only one function, $q(t)$, and generalise to a system of particles later. Suppose $q(t)$ is the function which minimises the action. Define another path between the same initial and final points

$$\tilde{q}(t) = q(t) + \delta q(t), \tag{1.3}$$

where δq (called a *variation*) is a small shift in the path, as shown in Fig. 1.2. By construction, the alternate path is subject to the constraints $\tilde{q}(t_0) = q(t_0)$, and $\tilde{q}(t_f) = q(t_f)$, or in other words,

$$\delta q(t_0) = \delta q(t_f) = 0. \tag{1.4}$$

[3] Note that we wrote the action as the integral over time. In theory, we could define it as the integral over *any* variable that uniquely parametrises the path.

[4] Technically, we seek an extremum (stationary point), which may be minimum or maximum.

The resulting variation in S is

$$\begin{aligned}\delta S &\equiv S(q+\delta q) - S(q) \\ &= \int_{t_0}^{t_f} L(q+\delta q, \dot{q}+\delta\dot{q}, t)\,\mathrm{d}t - \int_{t_0}^{t_f} L(q,\dot{q},t)\,\mathrm{d}t,\end{aligned} \tag{1.5}$$

where $\delta\dot{q} = \mathrm{d}(\delta q)/\mathrm{d}t$. We take δq to be an infinitesimal variation, and expand δS to first-order in δq. For S to be a minimum, the derivative of S must be zero. In other words, the variation δS must vanish:[5]

$$\delta S = \int_{t_0}^{t_f} \left(\frac{\partial L}{\partial q}\delta q + \frac{\partial L}{\partial \dot{q}}\delta\dot{q} \right) \mathrm{d}t = 0. \tag{1.6}$$

Using integration by parts for the second term, we have

$$\delta S = \int_{t_0}^{t_f} \left(\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}} \right) \delta q(t)\,\mathrm{d}t \;+\; \frac{\partial L}{\partial \dot{q}}\delta q \Bigg|_{t_0}^{t_f}. \tag{1.7}$$

From Eq. (1.4), the integrated boundary term is zero. The remaining term must be zero for any arbitrary variation δq, implying the integrand must be zero:

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}} = \frac{\partial L}{\partial q}. \tag{1.8}$$

For the case of N particles, the trajectory of each particle satisfies this equation:

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_i} = \frac{\partial L}{\partial q_i} \quad (i = 1, 2, \ldots, N). \tag{1.9}$$

These equations are not, in general, independent, since the Lagrangian will typically be a function of *all* coordinates, velocities, and time. These are called the *Euler-Lagrange equations*. If the Lagrangian of a system is known, these equations define the path taken by particles in the system, and give the equations of motion. We remind that, so-far, we have said nothing about what the Lagrangian actually is.

As a final note, consider what happens when we add a term to the Lagrangian that is a total time derivative of some function of coordinates and time:

$$L(q,\dot{q},t) \to L'(q,\dot{q},t) = L(q,\dot{q},t) + \frac{\mathrm{d}}{\mathrm{d}t}K(q,t). \tag{1.10}$$

Notice that the action (1.2) simply changes by a constant term:

$$S' = S + \int_{t_0}^{t_f} \frac{\mathrm{d}K}{\mathrm{d}t}\,\mathrm{d}t = S + K\Big|_{t_0}^{t_f}. \tag{1.11}$$

[5] This is an application of the *calculus of variations*; the formulas follow from the definition of the derivative, and the chain rule. If you're unsure, refer to the appendix in Sec. 1.7.2.

The constant doesn't change the condition for a minimum ($\delta S = 0$ and $\delta S' = 0$ are equivalent). Therefore, the addition of such a term cannot impact the dynamics of the system, and so the Lagrangian is defined only up to the addition of a total time derivative of any function $K(q, t)$.

1.3 The Lagrangian

We shall now consider the form the Lagrangian must take (at least, for particles in an inertial reference frame), resting entirely on a few intuitive assumptions about nature. To do this, we shall invoke our next assumption: the principles of *homogeneity* and *isotropy*.

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

We first consider the case of a single free particle. The homogeneity of space and time implies that there can be no explicit position $\boldsymbol{q}$ or time t dependence in the Lagrangian – the dynamics of a closed system should not depend on where or when they are examined. The Lagrangian must therefore be a function only of the velocity $\boldsymbol{v} = \dot{\boldsymbol{q}}$. Further, the isotropy of space means it must also be independent of the direction, and therefore the Lagrangian for a free particle must be only a function of the *magnitude* of velocity, v^2:

$$L = L(v^2). \tag{1.12}$$

Since the Lagrangian is independent of $\boldsymbol{q}$, we have $\partial L/\partial \boldsymbol{q} = 0$, and the Euler-Lagrange equations (1.9) become particularly simple in this case:[6]

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \boldsymbol{v}}\right) = 0. \tag{1.13}$$

From this, we see that $\partial L/\partial \boldsymbol{v}$ is a constant. Since L is a function only of $|\boldsymbol{v}|$, and the only time-dependence enters through $v(t)$, this implies the *velocity is constant*:

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{v} = 0. \tag{1.14}$$

In other words, in the absence of any interactions, the velocity of a free particle must stay constant. This is the *law of inertia* (or Newton's first law). It's important to note that we have made only the barest of assumptions to arrive at this conclusion: namely that physical systems are deterministic, the dynamics is such that some function (which we called the action) was optimised, and that the universe is homogeneous and isotropic.

[6] The derivative of a scalar with respect to a vector may be defined $\mathrm{d}\phi/\mathrm{d}\boldsymbol{a} \equiv \sum_i (\mathrm{d}\phi/\mathrm{d}a_i)\hat{\boldsymbol{e}}_i$, where $\hat{\boldsymbol{e}}_i$ is the unit vector parallel to the i component of $\boldsymbol{a}$. The special case of spatial derivatives is called the gradient, or 'grad' and is written $\mathrm{d}\phi/\mathrm{d}\boldsymbol{x} \equiv \nabla\phi$.

We so far have determined that the Lagrangian for a free particle must be a function of the magnitude of the velocity only. To go further, we make one more assumption and invoke the (Galilean) principle of relativity – that coordinates change according to Galilean transformations (1.1) when changing between inertial reference frames[7], and that the equations of motion have the same form in every inertial frame. Consider two inertial frames of reference, moving relative to each other by an infinitesimal (constant) velocity $\delta\boldsymbol{v}$, such that the velocities in one frame may be linked to those in the other by $\boldsymbol{v}' = \boldsymbol{v} + \delta\boldsymbol{v}$. Under our assumptions, $L' = L(v'^2)$ must differ from $L(v^2)$ by at most a total time derivative in order to leave the equations of motion invariant. We have, neglecting $(\delta v)^2$ terms, $L(v'^2) = L(v^2 + 2\delta\boldsymbol{v}\cdot\boldsymbol{v})$, which can be expanded as

$$L(v'^2) = L(v^2) + \frac{\partial L}{\partial(v^2)} 2\delta\boldsymbol{v}\cdot\boldsymbol{v}. \tag{1.15}$$

The equations of motion will remain unchanged only if the final term of this equation is zero (evidently it is not), or if it's a total time derivative, as we saw in Eq. (1.11). This final term is a total time derivative *only* if it is linear in $\boldsymbol{v}$ (since $\boldsymbol{v} = \dot{\boldsymbol{q}}$, and $\delta\boldsymbol{v}$ is independent of time):

$$L(v'^2) = L(v^2) + \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial(v^2)} 2\delta\boldsymbol{v}\cdot\boldsymbol{q}\right).$$

As such, $\partial L/\partial(v^2)$ must be independent of velocity, and therefore be a constant. We arbitrarily choose this constant to be $m/2$, and integrate to find

$$L = \frac{1}{2}mv^2. \tag{1.16}$$

Any integration constant will not affect the equations of motion, and can be discarded. The factor of $m/2$ is arbitrary so far, but we will call m the *mass* of the particle. For a single free particle, the mass has no physical significance. However, for a *system* of particles, each may have a different mass, and the ratios will be physically meaningful (particularly as we discuss *forces* below).

Notice that, in this free particle case, we have

$$\frac{\partial L}{\partial \boldsymbol{v}} = m\boldsymbol{v}, \tag{1.17}$$

which you will recognise as momentum. We will generalise this now, because it will turn out to be a very useful construction. If the system is defined by a set of generalised coordinates $\{q_i\}$, then we can define a *canonical momentum*, p,

$$p_i \equiv \frac{\partial L}{\partial \dot{q}}, \tag{1.18}$$

[7] This final assumption does not hold in relativistic mechanics.

which is also called the *generalised momentum*, or the *momentum conjugate* to q. For regular linear coordinates ($q = x$, and $\boldsymbol{v} = \dot{\boldsymbol{x}}$), then this corresponds to the regular linear momentum, though in general, this quantity is not simply $m\dot{\boldsymbol{x}}$. For now, we may consider this simply a definition; we will return to this in our discussion of symmetries. With this, the Euler-Lagrange equations may be written:

$$\frac{\mathrm{d}p_i}{\mathrm{d}t} = \frac{\partial L}{\partial q_i}. \tag{1.19}$$

1.3.1 Particle in an external field

The above discussion, in particular the homogeneity and isotropy arguments, were based on the assumption of a *closed* system. A *closed* system is one in which all important variables are considered part of the system; nothing outside the considered system may impact its dynamics.

On the other hand, an *open* system is one in which external influences may impact the dynamics of the studied particles. Consider for example, a particle moving in an electric field. In theory, the electric field is sourced by particles and everything on the whole would form part of a closed system. In practice, it is often more convenient to treat the background electric field as *external* to the system, and consider the particles inside as part of an open system. This is particularly useful in the case where the action of the considered particles does not significantly impact the external field, so may be considered independently.

If we consider an open system, some of the arguments we used above to derive the form of the Lagrangian must be revisited. Namely, the homogeneity and isotropy arguments no longer hold. As such, we now *may* have terms entering the Lagrangian that depend explicitly on coordinates. We therefore add an extra term, which we call the *potential*:

$$L = \frac{1}{2}m\boldsymbol{v}^2 - V(t, \boldsymbol{x}), \tag{1.20}$$

where the negative sign is by convention. We have written the potential as though it depends only on the coordinate, though in general it may also depend on the velocity. The specific form of V will depend on the details of the external field.

The equations of motion can be determined from an application of Eq. (1.9) without much difficulty. We find:

$$m\boldsymbol{a} = -\frac{\partial V}{\partial \boldsymbol{x}} \equiv \boldsymbol{F}, \tag{1.21}$$

where $\boldsymbol{a} = \dot{\boldsymbol{v}} = \ddot{\boldsymbol{x}}$ is the *acceleration*, and we call the derivative term on the right-hand-side the *force* on the particle, due to the potential $\boldsymbol{F} = -\nabla V$. You will recognise this as *Newton's second law*, which justifies our choice for the definition of the mass constant, m.

1.3.2 Systems of particles

We will now consider a *closed* system of several particles. In this case, the strict homogeneity/isotropy for each particle is broken by the presence of the other particles. Therefore, an extra term may appear in the Lagrangian that depends on the positions and velocities of all the particles. We call it the *potential*, V:

$$L = \sum_i \frac{1}{2} m_i \boldsymbol{v}_i^2 - V(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{v}_1, \boldsymbol{v}_2, \ldots). \tag{1.22}$$

For simplicity, we'll directly consider the case for two particles; the arguments will hold generally. The equations of motion follow in the same way. The equation for the ith particle is

$$m_i \boldsymbol{a}_i = -\frac{\partial V}{\partial \boldsymbol{x}_i}. \tag{1.23}$$

While the homogeneity for each particle is broken, the overall homogeneity and isotropy assumptions still hold for the system on a whole. Therefore, as before, there may be no explicit time dependence in the new potential term (it may depend implicitly on time through the time-dependence of the positions). Further, for the overall homogeneity/isotropy assumptions to hold, there can be no explicit position dependence in the potential term, besides the *relative* positions between particles. That is, the only position dependence allowed in V comes in the form of *differences*:

$$V(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots) = V(\boldsymbol{x}_1 - \boldsymbol{x}_2, \ldots). \tag{1.24}$$

From this, without knowing anything else about about the form of V, we immediately see that (in the case of a pair of particles), we have

$$\frac{\partial V}{\partial \boldsymbol{x}_1} = -\frac{\partial V}{\partial \boldsymbol{x}_2} \quad \Longrightarrow \quad m_1 \boldsymbol{a}_1 = -m_2 \boldsymbol{a}_2. \tag{1.25}$$

This states that forces between particles come in pairs, which are equal in magnitude and opposite in direction; you will recognise this as *Newton's third law*, which we see is a direct consequence of the homogeneity of space. We can rest assured that our formulation of mechanics encodes Newton's laws of motion.

As a matter of definition, if a Lagrangian can be broken into terms which are proportional to $\dot{q}^2$, we'll call those terms the *kinetic energy*, and remaining terms the (negative of the) *potential energy*. The choice of these terms is clear from the link to Newton's equations of motion, even if we haven't formally defined *energy* in the context of Lagrangian mechanics yet. In such cases, we write:

$$L = T - V, \tag{1.26}$$

Figure 1.3: Left: A mass, free to slide in a spherical bowl. Right: a driven pendulum.

where the kinetic energy is $T = (1/2)\sum_i m_i \dot{q}_i^2$, and V is the potential energy.

As a final remark, you might wonder if we can continue making such arguments to work out an explicit form the V in the same way as we did for a system of free particles. The answer is we cannot, at least not without introducing new assumptions. The form that V takes will be called a (classical) physical theory; any physical theory that obeys the above assumptions is equally valid, and it is up to experiment to determine which is the correct description of nature.

1.3.3 Elementary examples

As a simple example showing the power of generalised coordinates, consider a small particle of mass m, which is free to slide without friction inside a spherical bowl of radius R as shown in Fig. 1.3 (left). The bowl is centred at $(x, y, z) = (0, 0, 0)$, and the z-direction is taken directly upwards. Since the mass is confined to sit on the curved plane of the bowl, we can use its two-dimensional position on this plane to specify its location. Specifically, instead of the three coordinates x, y and z, we can use two generalised coordinates, θ (the angle defined such that $z = R - R\cos\theta$), and ϕ (the angular position in the x-y plane). Further, under the assumption that there is no initial velocity in the y direction, under the rotational symmetry, we can further confine the particle to lie in the x-z plane, meaning its position may be completely specified only by θ, with $x = R\sin\theta$. The kinetic energy of the particle is

$$T = \frac{1}{2}m(\dot{x}^2 + \dot{z}^2) = \frac{mR^2}{2}\dot{\theta}^2, \tag{1.27}$$

and if the potential energy is due to gravity, it is:

$$V = mgz = mgR(1 - \cos\theta). \tag{1.28}$$

The Euler-Lagrange equation is thus

$$\ddot{\theta} = -\frac{g}{R}\sin\theta. \tag{1.29}$$

As you can see, it is much simpler to solve these equations using the generalised coordinates than using Cartesian coordinates.

As an aside, if we make the further assumption that the angle be kept small so that $\sin\theta \approx \theta$, you may recognise the result as the equation for simple harmonic motion. The solution in that case is sinusoidal oscillations in the displacement. Substituting for x, the equation becomes $\ddot{x} = -(g/R)x$, with solution

$$x = x_0 \cos(\sqrt{g/R}\, t). \tag{1.30}$$

This gives the famous formula for the period of oscillation of an idealised pendulum: $T = 2\pi\sqrt{R/g}$.

***Problem* 1.3.1:** Consider a driven pendulum: a bob of mass m sits on a rigid pendulum of length l, which swings only in the x direction. The pendulum support is driven by an external force, and moves up and down in the y-direction with $y_s(t)$, as shown in Fig. 1.3 (right). Write the Lagrangian in terms of the generalised coordinate θ, the angle from the equilibrium.

Solution (1.3.1): Arbitrarily measuring y from the top of the pendulum support, so that the gravitational potential is $-mgy$, we have

$$L = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) + mgy.$$

In terms of the generalised coordinates, we have $x = l\cos\theta$, and $y = y_s(t) + l\cos\theta$, so that

$$L = \frac{m}{2}(l^2\dot{\theta}^2 + \dot{y}_s^2 + 2l\sin\theta\dot{y}_s\dot{\theta}) + mg(y_s + l\cos\theta),$$

which depends explicitly on time, through $y_s(t)$. We can simplify the Lagrangian by noticing that the terms involving y_s (that is, the terms that depend explicitly on time), $(m/2)\dot{y}_s^2 + ml\sin\theta\dot{y}_s\dot{\theta} + mgy_s$, are *almost* the total time derivative of:

$$K = \frac{m}{2}\int \dot{y}_s^2 \mathrm{d}t + ml\cos\theta\dot{y}_s + mg\int y_s \mathrm{d}t$$

(remember that y_t is externally driven, and not a dynamical variable). Writing $L \to L - \frac{\mathrm{d}K}{\mathrm{d}t}$, we are left with

$$L = \frac{m}{2}l^2\dot{\theta}^2 + ml\cos\theta(g - \ddot{y}_s(t)).$$

1.4 Symmetries and conservation laws

In the above, we considered *closed* systems of particles. That is, systems where anything (all the particles and interactions) that could impact the motion of the particles was included in the system. The mechanisms of classical mechanics are such that they can be applied in more general situations; we often can consider some subset of a larger system as "the system", and consider the effect of the excluded particles as some "external" force (or interaction more generally). For example, when determining the orbital dynamics of the earth around the sun, we don't need to consider the effect the earth has on the sun (at least to first order). So the gravitational force of the sun on the earth can be taken as an external interaction. It's important to realise that, in such situations, the assumptions of homogeneity and isotropy no longer hold. In reality, such situations are almost always *approximations*, since there will be some back reaction on whatever is

producing the external interactions. If we included *everything* into the system, these symmetries would be restored. The art of classical mechanics is to make reasonable assumptions about what must be considered internal or external to the system.

One concept that is extremely important in physics of all kinds, is that of *conservation laws*. As a mechanical system evolves in time, there may exist some function of the generalised coordinates q and $\dot{q}$ that remains constant throughout the motion, and depend only on the initial conditions. Such functions are known as *constants of the motion*, or integrals of the motion, or simply as conserved quantities. In this section, we will investigate the profound link between symmetries and conserved quantities.

A transformation is any change we can make to a system. They are usually defined through mathematical operations on the system or set of chosen variables. We will focus on continuous transformations, and in particular on coordinate transformations, but most of the logic holds for more general cases. A transformation is considered a *symmetry* if it leaves the dynamics unchanged. To prove this, it suffices to show that it leaves the Lagrangian unchanged (up to the addition of a total time derivative). As we shall see, it is a general rule that for any symmetry in a physical system, there is a corresponding quantity that is conserved. This is Noether's theorem. We will first consider a few important cases, and then show this for the general case.

1.4.1 Translation symmetry: momentum conservation

The first transformation we consider is a *translation*. We make a coordinate change which shifts all the coordinates by some infinitesimal constant:

$$\boldsymbol{q} \to \boldsymbol{q}' = \boldsymbol{q} + \delta\boldsymbol{q}. \tag{1.31}$$

It suffices to consider infinitesimal translations, as any finite translation can be built up from many repeated infinitesimal ones – this is the assertion that translation is a *continuous transformation*. The corresponding change in the Lagrangian

$$\delta L = \sum_i \frac{\partial L}{\partial \boldsymbol{q}_i} \cdot \delta\boldsymbol{q}_i \tag{1.32}$$

If this translation is a symmetry, then $\delta L = 0$. Since $\delta\boldsymbol{q}$ is a arbitrary, we have

$$\sum_i \frac{\partial L}{\partial \boldsymbol{q}_i} = 0,$$

and so the Euler-Lagrange equations become:

$$\frac{\mathrm{d}}{\mathrm{d}t} \sum_i \frac{\partial L}{\partial \dot{\boldsymbol{q}}_i} \equiv \frac{\mathrm{d}}{\mathrm{d}t} \sum_i \boldsymbol{p}_i = 0, \tag{1.33}$$

where we used the definition of canonical momentum (1.18). For the case of regular coordinates $\boldsymbol{q} = \boldsymbol{x}$ (and when the potential does not depend explicitly on velocity), $p_i = mv_i$ is the usual mechanical momentum. Since the time derivative of momentum is zero, it is conserved. As we saw above, this is clearly a symmetry for a closed system, which shows that the total momentum is conserved for a closed system. This also follows from the combination of Newtons second and third laws, which we saw above. It's also clearly true in the case that the potential does not depend on position, $\partial V/\partial x = 0$. We can summarise this as: translation symmetry implies momentum conservation. More generally, if a system is invariant under translations in generalised coordinate q, then $p = \partial L/\partial \dot{q}$, the canonical momentum conjugate to q, is a conserved quantity.

1.4.2 Time-translation symmetry: energy conservation

We now consider a translation in time:

$$t \to t' = t + \delta t. \tag{1.34}$$

The corresponding change in the Lagrangian is

$$\delta L = \frac{\partial L}{\partial t}\delta t \tag{1.35}$$

The condition for this to be a symmetry is $\delta L = 0$. Therefore, time translation is a symmetry if there is no explicit time dependence in the Lagrangian.

To investigate this, consider the total time derivative of the Lagrangian:

$$\begin{aligned}\frac{\mathrm{d}L}{\mathrm{d}t} &= \sum_i \left(\frac{\partial L}{\partial q_i}\dot{q}_i + \frac{\partial L}{\partial \dot{q}_i}\ddot{q}_i\right) + \frac{\partial L}{\partial t} = \sum_i \left(\dot{p}_i\dot{q}_i + p_i\ddot{q}_i\right) + \frac{\partial L}{\partial t} \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_i p_i\dot{q}_i\right) + \frac{\partial L}{\partial t},\end{aligned} \tag{1.36}$$

where we used the Euler-Lagrange equations (1.19) to replace $\partial L/\partial q$, and the canonical momentum (1.18) to replace $\partial L/\partial \dot{q}$. We define a new quantity, called the *Hamiltonian* (or the *energy*):

$$H \equiv \sum_i p_i\dot{q}_i - L. \tag{1.37}$$

With this definition, the above equation reads:

$$\frac{\mathrm{d}H}{\mathrm{d}t} = -\frac{\partial L}{\partial t}. \tag{1.38}$$

We this see that the Hamiltonian (or the energy) is conserved if the Lagrangian has no explicit time dependence, which from Eq. (1.35), is implied if there is time translation symmetry.

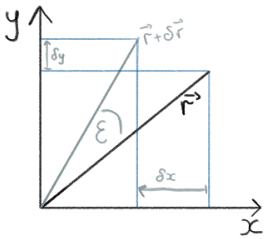

Figure 1.4: Small rotation, ϵ, about the z-axis.

The homogeneity of time implies energy is conserved for all closed systems. Further, the above shows that, even for open systems, the energy is conserved so long as the potential V is independent of time (since these will be invariant under time translations).

In the case where the Lagrangian is simply $L = T - V$, where T is a quadratic function in $\dot{q}$, the Hamiltonian is seen to be $H = T + V$. Therefore, we can recognise T as the kinetic and V as the potential energies. Note that for more complicated Lagrangians, it is not always simple or even possible to separate the energy terms into kinetic and potential contributions; still, Eq. (1.37) defines the link between the Lagrangian and the Hamiltonian.

1.4.3 Rotation symmetry: angular momentum conservation

The final explicit example we shall consider is a rotation about an axis. For now, let's consider an infinitesimal rotation around the z axis. It's a quick geometry exercise to see that the changes in the x and y coordinates are:

$$\delta x = -\delta\theta\, y, \quad \delta y = \delta\theta\, x, \tag{1.39}$$

see Fig. 1.4. The generalisation is the cross-product

$$\delta \boldsymbol{r} = \delta\boldsymbol{\theta} \times \boldsymbol{r}, \tag{1.40}$$

where $\delta\boldsymbol{\theta}$ is a vector with magnitude $\delta\theta$ that points along the axis of rotation, and $\boldsymbol{r}$ is the coordinate vector. Since the velocities are just the time-derivatives of the position vectors, $\boldsymbol{v} = \dot{\boldsymbol{r}}$, and $\delta\boldsymbol{\theta}$ is a constant, the velocity vectors change in the same way:

$$\delta \boldsymbol{v} = \delta\boldsymbol{\theta} \times \boldsymbol{v}. \tag{1.41}$$

The corresponding change in the Lagrangian is

$$\delta L = \sum_i \left(\frac{\partial L}{\partial \boldsymbol{r}_i} \cdot \delta \boldsymbol{r}_i + \frac{\partial L}{\partial \boldsymbol{v}_i} \cdot \delta \boldsymbol{v}_i \right) \tag{1.42}$$

$$= \sum_i \left[\dot{\boldsymbol{p}}_i \cdot (\delta \boldsymbol{\theta} \times \boldsymbol{r}_i) + \boldsymbol{p}_i \cdot (\delta \boldsymbol{\theta} \times \boldsymbol{v}_i) \right]$$

$$= \delta \boldsymbol{\theta} \cdot \frac{\mathrm{d}}{\mathrm{d}t} \sum_i \boldsymbol{r}_i \times \boldsymbol{p}_i, \tag{1.43}$$

where we used the Euler-Lagrange equations and the permutation properties of the cross-product, and the sum extends over each particle in the system. Therefore, the condition that the rotation is a symmetry ($\delta L = 0$), implies that term om the right-hand-side $\boldsymbol{r} \times \boldsymbol{p}$ is a constant. We call this quantity *angular momentum*:

$$\boldsymbol{l} = \boldsymbol{r} \times \boldsymbol{p} \tag{1.44}$$

(some places use $\boldsymbol{L}$ or $\boldsymbol{M}$). Note that each component of the angular momentum is conserved separately; rotational symmetry about an axis j implies the j-component of the angular momentum is conserved.

1.4.4 General symmetries

In the general case, we define an infinitesimal transformation that may itself be a function of coordinates and time:

$$\delta q_i = f_i(q, t)\epsilon, \tag{1.45}$$

where the *epsilon* is to make explicit that we consider an infinitesimal shift. In general, the velocities will also change under the transformation:

$$\delta \dot{q}_i = \frac{\mathrm{d}}{\mathrm{d}t}(\delta q_i). \tag{1.46}$$

The general change in the Lagrangian is then

$$\delta L = \sum_i \left(\frac{\partial L}{\partial q_i} \delta q_i + \frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i \right) = \sum_i \left(\dot{p}_i \delta q_i + p_i \delta \dot{q}_i \right)$$

$$= \frac{\mathrm{d}}{\mathrm{d}t} \sum_i p_i \delta q_i = \frac{\mathrm{d}}{\mathrm{d}t} \overbrace{\sum_i p_i f_i(q, t)}^{Q} \epsilon, \tag{1.47}$$

where we used Eqs. (1.18) and (1.19) in the first line, and the product rule for differentiation in the second. If the transformation is a symmetry, i.e., if $\delta L = 0$, then the time derivative of the term on the right-hand-side is zero, and thus

this term is conserved. We can re-state this general result more concretely. If a transformation $\delta q_i = f_i(q)\epsilon$ leaves the Lagrangian unchanged, $\delta L = 0$, then:

$$\delta L(f) = 0 \quad \Longrightarrow \quad \frac{\mathrm{d}}{\mathrm{d}t} Q = 0, \tag{1.48}$$

$$\text{where} \quad Q \equiv \sum_i p_i f_i(q), \tag{1.49}$$

which is a statement of *Noether's theorem*. We shall return to Noether's theorem in our study of relativistic field theory, where it becomes a much stronger condition.

As a final note, consider the situation when this transformation does *not* leave the Lagrangian unchanged, $\delta L \neq 0$. Equation (1.47) still tells us that the change in the Lagrangian is a total time derivative, and therefore that the dynamics on the particles remain unchanged. This is of course not surprising; it tells us that *any* coordinate transformation of the form of Eq. (1.45) leaves the dynamics unchanged, and therefore that any valid system of coordinates is as good as any other. To be in the form of Eq. (1.45), the transformation must be smooth (differentiable), and invertible (i.e., f_i, and thus δq cannot go to zero).

1.5 Hamiltonian formulation

In Eq. (1.37), we defined the *Hamiltonian*, which is a function of qs and ps. As we shall now see, this quantity is very important, and leads to a new formulation of the equations of motion.

To see this, consider a small variation in H from Eq. (1.37), treating the qs and $\dot{q}$s as independent variables:

$$\delta H = \sum_i \left(p_i \delta\dot{q}_i + \delta p_i \dot{q}_i\right) - \delta L \tag{1.50}$$

$$= \sum_i \left(p_i \delta\dot{q}_i + \delta p_i \dot{q}_i - \frac{\partial L}{\partial q_i}\delta q_i - \frac{\partial L}{\partial \dot{q}_i}\delta\dot{q}_i\right)$$

$$= \sum_i \left(\cancel{p_i \delta\dot{q}_i} + \delta p_i \dot{q}_i - \dot{p}_i \delta q_i - \cancel{p_i \delta\dot{q}_i}\right) \tag{1.51}$$

where we used the Euler-Lagrange equations. At the same time, treating the qs and ps as independent variables, we have

$$\delta H = \sum_i \left(\frac{\partial H}{\partial q_i}\delta q_i + \frac{\partial H}{\partial p_i}\delta p_i\right), \tag{1.52}$$

which holds for any general function of qs and ps. Equations (1.51) and (1.52) are equivalent. Equating these, and matching terms, we find

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad \text{and} \quad \dot{p}_i = -\frac{\partial H}{\partial q_i}, \tag{1.53}$$

which are *Hamilton's equations of motion*; it's interesting to note the near symmetry between the equations for q and p.

This way of writing to equations of motion is called Hamilton's formulation of classical mechanics. Compared to the Euler-Lagrange equations of motion, we now have twice as many equations, however, they are each *first-order* differential equations, rather than second-order. In particular, certain problems become much simpler to solve in the Hamiltonian formulation. This formulation also lends itself naturally to thinking about problems in the (p, q) *phase-space*, rather than the coordinate space q, as was hinted at in the introduction. This is often a very powerful and insightful way of treating problems. Transforming the set of coordinates from the Lagrangian to the Hamiltonian formulation $(q, \dot{q}, t) \to (q, p, t)$ is an example of a *Legendre transformation*.

1.5.1 Harmonic oscillator

As a simple example, consider a Lagrangian of the form

$$L = \frac{m\dot{x}^2}{2} - \frac{k}{2}x^2, \tag{1.54}$$

which (as we'll see) corresponds to a classical harmonic oscillator, with x being the displacement from the equilibrium, and k being the spring constant. It's not too difficult to check that if we make the change of variables $q = (mk)^{1/4}x$ and define $\omega = \sqrt{k/m}$, then the form of the Lagrangian becomes simpler:

$$L = \frac{\dot{q}^2}{2\omega} - \frac{\omega}{2}q^2. \tag{1.55}$$

We may solve this by finding the Euler-Lagrange equations of motion:

$$\ddot{q} = -\omega^2 q. \tag{1.56}$$

This differential equation is easy enough to solve; it implies sinusoidal motion about $q = 0$, with angular frequency ω. The amplitude and phase are determined by the initial conditions. Since this is a second-order equation, we require two initial conditions, q_0 and $\dot{q}_0$.

This is also a good example for the simplicity offered by the Hamiltonian formulation. The Hamiltonian can be found simply from Eq. (1.37):

$$H = \frac{\omega}{2}(p^2 + q^2). \tag{1.57}$$

Note that momentum is *not* conserved in this example, though the Hamiltonian formulation makes it plain that energy *is*. Then, the Hamiltonian equations of motion are:

$$\dot{q} = \omega p, \quad \text{and} \quad \dot{p} = -\omega q. \tag{1.58}$$

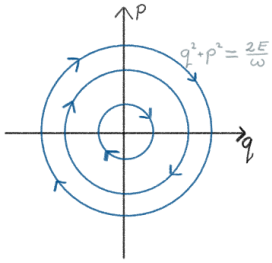

Figure 1.5: Simple harmonic oscillator tracing closed paths in phase space. Each path is for a constant energy. Each point on the plot corresponds to a possible set of (q_0, p_0) initial conditions; from there, a concentric circle will be traced.

It's easy to verify that these are equivalent to the Euler-Lagrange version by taking the time derivative of the first equation. But Hamilton's formulation makes the dynamics for p clear.

Since the energy is constant, Eq. (1.57) immediately tells us that the solutions in (q, p) phase space plot out concentric circles, with the "radius" of the circle corresponding to the energy, as shown in Fig. 1.5. The solution to the differential equations are also sinusoidal, and it can be seen that, not only does q oscillate around $q = 0$, but p oscillates around $p = 0$, with the same angular frequency ω.

1.5.2 Deriving Hamilton's equations from the action

It is instructive to note that we may derive Hamilton's equations of motion directly from the action. First, we need to write the action in terms of the Hamiltonian and its corresponding canonical variables p and q.

In section 1.2, we considered general variations of the action, finding

$$\delta S = \int_{t_0}^{t_f} \left(\frac{\partial L}{\partial q} - \frac{\mathrm{d}p}{\mathrm{d}t} \right) \delta q \, \mathrm{d}t \; + \; p \, \delta q \Big|_{t_0}^{t_f}, \tag{1.59}$$

where we used the definition of the canonical momentum (1.18). Previously, we enforced the variation δq to be zero at the boundaries in order to derive the Euler-Lagrange equations of motion. Here, we will take a slightly different route.

First, note that we may interpret the action as an explicit function of time, for example, by considering $S(t) = \int_{t_0}^{t} L$ for points $t_i < t < t_f$ (for simplicity, we leave the initial point fixed, though in general we could vary this as well). Then, the boundary term for $\delta S(t)$ is not zero: in our choice we have $\delta q(t_i) = 0$, though $\delta q(t)$ is non-zero. Since the equations of motion (1.9) are satisfied, the term in the parenthesis in Eq. (1.59) is zero, and we have

$$\delta S(t) = p \, \delta q(t). \tag{1.60}$$

This means we can identify the conjugate momenta as the partial derivatives of the action with respect to the coordinates:

$$p_i = \frac{\partial S}{\partial q_i}. \tag{1.61}$$

At the same time, considering the action as an explicit function of time and coordinates, we may write its variation as

$$\delta S = \frac{\partial S}{\partial t}\delta t + \frac{\partial S}{\partial q}\delta q = \frac{\partial S}{\partial t}\delta t + p\,\delta q. \tag{1.62}$$

Dividing by δt, and using the definition $\mathrm{d}S/\mathrm{d}t = L$, we see that

$$\frac{\partial S}{\partial t} = L - p\,\dot{q} = -H, \tag{1.63}$$

where we used definition of Hamiltonian (1.37). By integrating Eq. (1.62), we arrive at an expression of the action in terms of the Hamiltonian canonical variables:

$$S = \int \left(\sum p_i\,\mathrm{d}q_i - H\mathrm{d}t \right), \tag{1.64}$$

where we restored the general case of more than one particle/degree of freedom.

Now, in the Hamiltonian formulation, treating the coordinates and momenta as the independent variables, we may once again vary the action:

$$\begin{aligned}\delta S &= \int \left(-\frac{\partial H}{\partial q}\delta q\,\mathrm{d}t - \frac{\partial H}{\partial p}\delta p\,\mathrm{d}t + \delta p\,\mathrm{d}q + p\,\mathrm{d}(\delta q) \right) \\ &= \int \left(-\frac{\partial H}{\partial q}\delta q\,\mathrm{d}t - \frac{\partial H}{\partial p}\delta p\,\mathrm{d}t + \delta p\,\mathrm{d}q - \mathrm{d}p\,\delta q \right) + p\delta q\Big|_{t_0}^{t_f},\end{aligned} \tag{1.65}$$

where we used integration by parts for the last term. Dropping the boundary term as before (the variation at the boundary is zero), we arrive at

$$\delta S = \int \left(\left[\mathrm{d}q - \frac{\partial H}{\partial p}\,\mathrm{d}t \right] \delta p - \left[\mathrm{d}p + \frac{\partial H}{\partial q}\,\mathrm{d}t \right] \delta q \right). \tag{1.66}$$

For this to be zero for general variations in the ps and qs, the terms in the brackets must each be zero independently. Dividing these terms by $\mathrm{d}t$ yields Hamilton's equations of motion (1.53).

1.5.3 Canonical Transformations

In the Lagrangian formalism, physical systems are described by a set of coordinates $\{q_i\}$, and their derivatives, $\{\dot{q}_i\}$. Clearly, *any* choice of generalised coordinates that uniquely determines the system is as good as any other; no reference to

a specific coordinate system was used in deriving the Euler-Lagrange equations of motion. We may then consider a *point transformation* as a transform that maps one set of coordinates q to another, Q. The new coordinates will be a function of the old, and may depend explicitly on time: $Q = Q(q, t)$. The derivatives in the new coordinates are of course fixed as $\dot{Q} = \frac{\mathrm{d}}{\mathrm{d}t}q$. If the transformation is *invertible* (i.e., there exists a unique $q = q(Q, t)$), then the new coordinates Q uniquely describe the configuration of the system, just as the original coordinates q did. In that case, the transformation defines a valid new set of generalised coordinates.

One of the advantages of the Hamiltonian formalism – where the system is described by the set of coordinates q and momenta p – is that is allows a much wider class of transformations:

$$q \to Q(q, p, t), \quad p \to P(q, p, t). \tag{1.67}$$

However, since Hamilton's equation of motion define a relation between the ps and qs, it is clear that not any such transform will preserve the form of the equations of motion. The particular transformations that *do* preserve the form of Hamilton's equations of motion (1.53) are called *canonical transformations*:

$$\dot{Q}_i = \frac{\partial \widetilde{H}}{\partial P_i}, \quad \text{and} \quad \dot{P}_i = -\frac{\partial \widetilde{H}}{\partial Q_i}, \tag{1.68}$$

where $\widetilde{H}$ is the Hamiltonian in the new coordinates/momenta.

We will now derive the condition of a transformation of the form (1.67) to be canonical. The Lagrangians in the old and new coordinates may be written

$$L = \sum p_i q_i - H(p, q, t), \tag{1.69}$$

$$\widetilde{L} = \sum P_i Q_i - \widetilde{H}(p, q, t). \tag{1.70}$$

For the new and old coordinates to describe the same physical system, these must differ at most by a total time derivative of some function, F. From Eq. (1.64), we can express this condition as

$$\mathrm{d}F = \sum p_i \,\mathrm{d}q_i - \sum P_i \,\mathrm{d}Q_i + (\widetilde{H} - H)\mathrm{d}t, \tag{1.71}$$

where F is called the *generating function* of the canonical transformation.

If we consider F to be a function of the old and new coordinates, $F = F(q, Q, t)$, then we can read off

$$p_i = \frac{\partial F}{\partial q}, \quad P_i = -\frac{\partial F}{\partial Q}, \quad \text{and} \quad \widetilde{H} = H + \frac{\partial F}{\partial t}. \tag{1.72}$$

Of course, we could consider generating functions of various forms; using the notation of Goldstein [2], we can write:

$$F_1(q, Q, t), \quad F_2(q, P, t), \quad F_3(p, Q, t), \quad \text{or} \quad F_4(p, P, t). \tag{1.73}$$

1.6 Poisson brackets

Let $f(p, q, t)$ be any function of coordinates, momentum, and time. The total time derivative of f can be written as

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial t} + \sum_i \left(\frac{\partial f}{\partial q_i}\dot{q}_i + \frac{\partial f}{\partial p_i}\dot{p}_i \right), \tag{1.74}$$

which from Hamilton's equations, can be expressed:

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial t} + \sum_i \left(\frac{\partial f}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial g}{\partial p_i}\frac{\partial H}{\partial q_i} \right). \tag{1.75}$$

The construction on the right-hand-side proves to be useful so we designate it as a *Poisson bracket*, defined generally:[8]

$$\{f, g\} \equiv \sum_i \left(\frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i} \right). \tag{1.76}$$

In this case, time derivatives can be expressed:

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \{f, H\} + \frac{\partial f}{\partial t}. \tag{1.77}$$

When it is not clear from context which variables are used for the derivatives, these are given as subscripts. For example, the above (1.76) would be $\{f, g\}_{qp}$. Generally:

$$\{f, g\}_{uv} \equiv \sum_i \left(\frac{\partial f}{\partial u_i}\frac{\partial g}{\partial v_i} - \frac{\partial f}{\partial v_i}\frac{\partial g}{\partial u_i} \right). \tag{1.78}$$

Clearly, the Poisson bracket is anti-symmetric: $\{A, B\} = -\{B, A\}$. It's also fairly clear to show the linearity properties:

$$\{\lambda A, B\} = \lambda\{A, B\}, \quad \text{and} \quad \{A + D, C\} = \{A, C\} + \{D, C\}. \tag{1.79}$$

Also, from the product rule, we have:

$$\{AB, C\} = A\{B, C\} + \{A, C\}B. \tag{1.80}$$

Finally, the Poisson brackets of the canonical variables are

$$\{q_i, q_j\} = 0, \qquad \{p_i, p_j\} = 0, \qquad \{q_i, p_j\} = \delta_{ij}. \tag{1.81}$$

The above set of equations are in fact enough to define the Poisson bracket, and can be used as a set of axioms.

[8] Both curly braces $\{,\}$ and square brackets $[,]$ are commonly used to denote Poisson brackets.

The Poisson bracket proves a useful tool in many situations. Hamilton's equations can be readily expressed in this form:

$$\dot{q}_i = \{q_i, H\}, \qquad \dot{p}_i = \{p_i, H\}. \tag{1.82}$$

When one of the functions is one of the canonical variables, the Poisson bracket becomes a partial derivative with respect to the other:

$$\{f, q_i\} = -\frac{\partial f}{\partial p_i}, \qquad \{f, p_i\} = \frac{\partial f}{\partial q_i}, \tag{1.83}$$

which follow immediately from the definition Eq. (1.76), noting that p and q are independent variables.

There are many useful properties of Poisson brackets. One of particular importance is *Jacobi's identity*:

$$\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0. \tag{1.84}$$

An important property of the Poisson bracket is that, if f and g are conserved quantities (i.e., they are constants of the motion), then so is $\{f, g\}$. This is called *Poisson's theorem*, which can be proven by setting $h = H$ in Jacobi's identity.

As an aside, if you have studied quantum mechanics, you may have noticed that the classical Hamilton's equations (1.82) look identical to Heisenberg's quantum equations of motion. This is not simply a coincidence. It will be the case that the quantum version of Poisson brackets become the commutation relations.

Poisson brackets allow us to encode symmetries and conserved quantities algebraically. In the next section, we'll see how conserved quantities act as generators of transformations.

***Problem* 1.6.1:** Prove the identities in Eq. (1.83).

***Problem* 1.6.2:** Prove Jacobi's identity, Eq. (1.84).

***Problem* 1.6.3:** Prove Poisson's theorem.

1.6.1 Generators of transformations

We now explore how conserved quantities, like momentum and angular momentum, not only remain constant but actively generate transformations – such as translations or rotations – via the Poisson bracket.

In Sec. (1.4.3), we considered a small rotation of $\delta\theta$ about the z-axis, and saw the corresponding changes in the coordinates were $\delta x = -\delta\theta y$, $\delta y = \delta\theta x$, $\delta z = 0$. We saw that, if the system was symmetric with respect to rotation around z, then the z-component of angular momentum, $l_z = xp_y - yp_x$, was conserved. It's instructive to take the Poisson bracket of the coordinates with respect to the conserved quantity, l_z. Without too much difficulty, one finds:

$$\{x, l_z\} = -y, \quad \{y, l_z\} = x, \quad \{z, l_z\} = 0. \tag{1.85}$$

Notice that this can be written:

$$\delta\theta\{x, l_z\} = \delta x, \quad \delta\theta\{y, l_z\} = \delta y, \quad \delta\theta\{z, l_z\} = \delta z. \tag{1.86}$$

In other words, the Poisson bracket of coordinates with l_z give the expressions for the change in those coordinates due to a rotation around the z axis (up to the factor ϵ). This of course generalises to rotations about an arbitrary axis. If we instead rotate with $\delta\boldsymbol{r} = \delta\boldsymbol{\theta} \times \boldsymbol{r}$ as in Eq. (1.40), we have[9]

$$\{q_i, l_j\} = \sum_k \epsilon_{ijk} q_k = \delta q_i/\delta\theta. \tag{1.87}$$

In this sense, we can call the angular momentum vector the *generator* of rotations. Also, for the momentum, we similarly have:

$$\{p_i, l_j\} = \sum_k \epsilon_{ijk} p_k = \delta p_i/\delta\theta. \tag{1.88}$$

This is the same, since momentum vectors transform under rotations the same way as positions.

This can be extended to arbitrary functions of the canonical variables, $f(q, p)$. Under a rotation about the j-axis, we have

$$\delta f = \sum_i \left(\frac{\partial f}{\partial q_i}\delta q_i + \frac{\partial f}{\partial p_i}\delta p_i \right) = \delta\theta \sum_i \left(\frac{\partial f}{\partial q_i}\{q_i, l_j\} + \frac{\partial f}{\partial p_i}\{p_i, l_j\} \right), \tag{1.89}$$

where we used Eqs.(1.87) and (1.88). Finally, by using Eq. (1.83), we have

$$\delta f = \delta\theta \sum_i \left(\frac{\partial f}{\partial q_i}\frac{\partial l_j}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial l_j}{\partial q_i} \right) = \delta\theta\{f, l_j\}. \tag{1.90}$$

We can show the similar result for spatial and temporal translations. Above, we saw that invariance under spatial translations implied momentum conservation. Consider, then, the Poisson bracket of any function of position with momentum,

$$\{f(q), p\} = \frac{\mathrm{d}f}{\mathrm{d}q}. \tag{1.91}$$

Since the change in f under the transformation $q \to q + \delta q$ is $\frac{\mathrm{d}f}{\mathrm{d}q}\delta q$, we have

$$\delta f = \delta q\{f, p\}. \tag{1.92}$$

We may thus recognise momentum as the generator of spatial translations.

[9] ϵ_{ijk} is the entirely anti-symmetric Levi-Civita symbol. It is equal to zero if any of the indices repeat, $+1$ for any even permutation of (123), and -1 for any odd permutation. The cross-product $\boldsymbol{a} \times \boldsymbol{b} = \boldsymbol{c}$ can be expressed $\sum_{jk} \epsilon_{ijk} a_j b_k = c_i$.

Similarly, we saw that the Hamiltonian was conserved for systems with time-translation invariance. Forming the Poisson bracket with the Hamiltonian, and using Eq. (1.77), shows that the Hamiltonian is the generator of time translations:

$$\delta f = \delta t\{f, H\}. \tag{1.93}$$

We may now extend this important result to general symmetries. Let $G(q, p)$ be any function of generalised coordinates and momentum, defined such that it produces infinitesimal transformations of the coordinates:

$$\delta q_i \equiv \{q_i, G\}, \quad \text{and} \quad \delta p_i \equiv \{p_i, G\}. \tag{1.94}$$

Notice that it follows that G produces infinitesimal transformations of any arbitrary functions of the coordinates. To see this note that

$$\delta f = \frac{\partial f}{\partial q}\delta q + \frac{\partial f}{\partial p}\delta p = \frac{\partial f}{\partial q}\{q_i, G\} + \frac{\partial f}{\partial p}\{p_i, G\}.$$

Then, in the exact same way as for (1.90), we find

$$\delta f(q, p) \equiv \{f(q, p), G\}. \tag{1.95}$$

The transformation produced by G may or may not be a symmetry. If it is a symmetry, then, by definition, it cannot change the energy of the system, so $\delta H = 0$. In other words, from Eq. (1.95), the condition that G represents a symmetry is

$$\{H, G\} = 0. \tag{1.96}$$

From the anti-symmetry property of the Poisson bracket, we could equivalently write this the other way:

$$\{G, H\} = 0, \tag{1.97}$$

which, since H is the generator of time translations [Eq. (1.93)], tells us that G is conserved: $\frac{\mathrm{d}G}{\mathrm{d}t} = 0$.

1.7 Appendix

1.7.1 Principle of locality

Simply stated, the principle of locality says that systems are influenced only by their immediate (i.e., *local*) environment. As an example, consider Newton's second law:

$$a(t) = F(t)/m. \tag{1.98}$$

It states that the acceleration of a particle at some time t is proportional to the force on the particle at the same time, t. We could just as well have an equation

where the acceleration depends on the force at some other time, t', but this in non-local:

$$\cancel{a(t) = F(t')/m.}$$

In the same vein, the force on a particle at position $\boldsymbol{x}$ and time t is proportional to the gradient of the potential, V, at the *same* position $\boldsymbol{x}$ and time t,

$$F(\boldsymbol{x}, t) = -\nabla V(\boldsymbol{x}, t), \tag{1.99}$$

while locality forbids equations of the form

$$\cancel{F(\boldsymbol{x}, t) = -\nabla V(\boldsymbol{x}', t').}$$

In classical mechanics, the concept of locality can be a little confused, and becomes more strongly defined in relativistic mechanics. This is since the potential, V, is typically due to some source (e.g., gravitational mass, or an electric charge) which acts at a distance. The basic equations of classical mechanics are local, though the theory of how the potential is formed often breaks locality. For example, in Newtonian gravity, the gravitational potential instantaneously follows the mass which produces it up to infinite distances away, and thus breaks locality. This issue is removed within the framework of Einstein's theory of relativity. Similarly, Coulomb's law of electrostatic forces (appears to) violate locality. This issue is removed within the Maxwell's classical theory of electrodynamics, which we shall return to in our study of field theory.

1.7.2 Calculus of variations

The definition of the derivative for some function $f(y)$ may be stated:

$$\frac{\mathrm{d}f}{\mathrm{d}y} \equiv \lim_{\Delta y \to 0} \frac{\Delta f}{\Delta y}. \tag{1.100}$$

where $\Delta f \equiv f(y + \Delta y) - f(y)$. For infinitesimal changes, we write $\Delta f \to \delta f$, and $\Delta y \to \delta y$, and we can write:

$$\delta f \equiv \frac{\mathrm{d}f}{\mathrm{d}y} \delta y, \tag{1.101}$$

where δf is called the infinitesimal *variation* in f. If f is function of multiple variables x_1, x_2, ..., each of which may depend on y, such that $f(y) = f(x_1(y), x_2(y), \ldots)$, then, by the chain rule, we have

$$\frac{\mathrm{d}f}{\mathrm{d}y} = \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial y} + \frac{\partial f}{\partial x_2}\frac{\partial x_2}{\partial y} + \ldots. \tag{1.102}$$

Combining Eqs. (1.101) and (1.102), we have

$$\begin{aligned}\delta f &= \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial y}\delta y + \frac{\partial f}{\partial x_2}\frac{\partial x_2}{\partial y}\delta y + \dots \\ &= \frac{\partial f}{\partial x_1}\delta x_1 + \frac{\partial f}{\partial x_2}\delta x_2 + \dots,\end{aligned} \tag{1.103}$$

which is the general formula for the variation in f.

References

[1]L. D. Landau and E. M. Lifshitz, *Mechanics* (1976).

[2]H. Goldstein, C. Poole, and J. Safko, *Classical Mechanics* (2001).

[3]L. Susskind and G. Hrabovsky, *Classical Mechanics* (2014).