1 Classical Mechanics

[April 19, 2023]

This set of notes should best be thought of as a *companion* to a good textbook, and is by no means complete. Some prior knowledge is assumed, including a basic knowledge of calculus, and some familiarity with elementary concepts of mechanics (meaning of forces, energy, coordinates etc.). I recommend the books:

- H. Goldstein, C. Poole, and J. Safko, *Classical Mechanics*, 3rd ed. (Addison Wesley, 2001). One of the standards, for good reason; very solid and thorough coverage.
- L. Susskind and G. Hrabovsky, *Classical Mechanics: The Theoretical Minimum* (Penguin, 2014). Not exactly a textbook, not exactly a popular science book; somewhere in between. An enjoyable read at an introductory level.
- L. D. Landau and E. M. Lifshitz, *Mechanics (L&L Vol. 1)* (Pergamon, Oxford, 1976). An older and technical book. A very elegant coverage of classical mechanics, that heavily influenced these notes. More appropriate at a slightly more advanced level.

1.1 Principles of classical mechanics

The aim of classical mechanics is to describe physical systems, and determine their evolution through time. In other words, it is to predict the future. At the core of classical mechanics are three basic assumptions:

1. Mechanical systems are deterministic,
2. The universe is homogeneous and isotropic,
3. Laws of physics are the same in all (Galilean) 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 the 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*; we shall soon see that this follows from the assumptions of determinism and homogeneity. The assumption that the universe is homogeneous means that the laws of nature do not depend explicitly on *position*, or in other words, there is no preferred location in space. We also assume the universe is homogeneous in time, meaning the laws of nature do not depend explicitly on time; this is usually wrapped into the homogeneity assumption. The assumption that the universe is *isotropic* means that the laws of nature do not depend explicitly on *orientation*, or that there is no preferred direction in space. The third assumption is more subtle, we will return to it later.

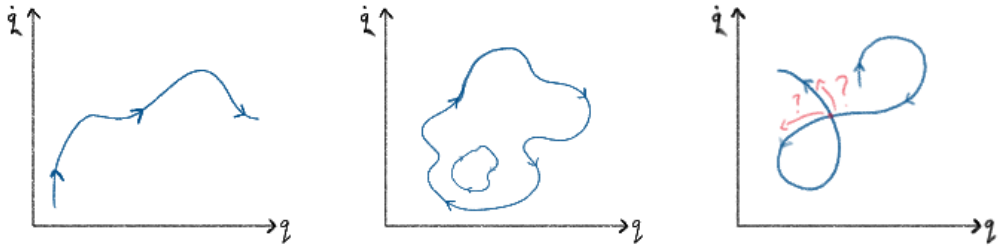

Figure 1: Two valid (deterministic and reversible) configuration paths, and one invalid path. (These are not meant to be realistic, just illustrative.)

We will focus most of our discussion on the dynamics of *particles*. By particle, we mean a body whose shape and dimension can be neglected in describing its motion.[1] 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* $V(\{\boldsymbol{x}_i\}, \{\dot{\boldsymbol{x}}_i\})$, 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, from the reversibility assumption, 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 may 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 $M \neq 3N$, 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 derivatives, $\{\dot{q}_i\}$, are the *generalised velocities*.

We may consider the *phase space* of generalised positions and velocities.[2] We can then imagine the physical system tracing some path through this phase space 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 is not allowed, see Fig. 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 (generally, some function of qs and $\dot{q}$s) that remains constant throughout the evolution of the system.

[1] More complicated systems can always be considered as complex arrangements of particles.

[2] Usually, the term phase space actually refers to the space of generalised positions q and *momentums* p, rather than velocities. For our arguments here, it makes no difference. We will return the concept of generalised momentum in the coming sections.

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: $\boldsymbol{x} \to \boldsymbol{x}'$. We may also make more general coordinate transforms that depend on time. For example, consider 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 is:

$$\begin{aligned} \boldsymbol{x} \to \boldsymbol{x}' &= \boldsymbol{x} - \boldsymbol{V}t \\ t \to t' &= t. \end{aligned} \tag{1}$$

These equations define what is called a *Galilean transformation*. Reference frames that are linked by Galilean transformations are considered *inertial* frames.[3] 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 (relativity is not considered in these notes).

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.

It seems reasonable to assume that the path the system will take will be *optimal*, with respect to some quantity. 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).

We call the function that should be minimised (or more generally, optimised), the *action*, denoted S. The action is itself typically written as the integral of another function, called the *Lagrangian*, L:

$$S = \int_{t_0}^{t_f} L(q, \dot{q}, t)\,\mathrm{d}t. \tag{2}$$

We may consider this simply convention.[4] The task to completely specify the dynamics of the system, then, is to find the set of functions $q_i(t)$, which *minimise* the action[5]. This is called the *principle of least action*, or Hamilton's principle.

We shall now show how to derive the set of differential equations that will determine these paths. For simplicity, we will work in the case of a single particle, so there is only one function, $q(t)$, that we must find, 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

$$\tilde{q}(t) = q(t) + \delta q(t), \tag{3}$$

where δq (called a *variation*) is a small shift in the path, subject to the constraint $\tilde{q}(t_0) = q(t_0)$, and $\tilde{q}(t_f) = q(t_f)$. In other words,

$$\delta q(t_0) = \delta q(t_f) = 0. \tag{4}$$

[3] More concretely, an inertial frame is one in which free particles do not accelerate.

[4] This can be linked to locality arguments; the action S, is not a *local* function, in that it depends on the entirety of the path taken. The Lagrangian, L, on the other hand, is a local function of time.

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

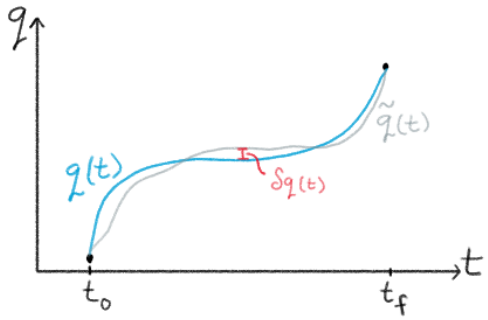

Figure 2: Path that minimises the action, $q(t)$, and its variation, $\delta q(t)$.

The resulting variation in S is

$$\delta S \equiv S(q+\delta q) - S(q) \tag{5}$$

$$= \int_{t_0}^{t_f} L(q+\delta q, \dot{q}+\delta\dot{q}, t)\,\mathrm{d}t - \int_{t_0}^{t_f} L(q,\dot{q},t)\,\mathrm{d}t, \tag{6}$$

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.[6] For S to be a minimum, this variation must vanish:

$$\delta S = \int_{t_0}^{t_f} \left(\frac{\partial L}{\partial q}\delta q + \frac{\partial L}{\partial \dot{q}}\delta\dot{q}\right)\,\mathrm{d}t = 0. \tag{7}$$

Using integration by parts for the second term, we have

$$\delta S = \int_{t_0}^{t_f} \left(\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}}\right)\delta q(t)\,\mathrm{d}t + \frac{\partial L}{\partial \dot{q}}\delta\dot{q}\Big|_{t_0}^{t_f} = 0. \tag{8}$$

From Eq. (4), the integrated term is zero. The remaining term must be zero for *all* functions δq, implying the integrand must be zero:

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}} = \frac{\partial L}{\partial q}. \tag{9}$$

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

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_i} = \frac{\partial L}{\partial q_i} \quad (i = 1, 2, \ldots, N), \tag{10}$$

where the Lagrangian is, in general, 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:

$$L(q,\dot{q},t) \to L'(q,\dot{q},t) = L(q,\dot{q},t) + \frac{\mathrm{d}}{\mathrm{d}t}f(q,t). \tag{11}$$

[6] 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. 1.9.

Notice that the action (2) simply changes by a constant term:

$$S' = S + \int_{t_0}^{t_f} \frac{\mathrm{d}f}{\mathrm{d}t}\,\mathrm{d}t = S + f\Big|_{t_0}^{t_f}. \tag{12}$$

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 $f(q,t)$.

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: that the universe is *homogeneous* and *isotropic*.

We first consider the case of a single free particle. The homogeneity of space and time implies that there can be no explicit position $\boldsymbol{x}$ or time t dependence in the Lagrangian. The Lagrangian must therefore be a function of velocity $\dot{\boldsymbol{x}}$. Further, the isotropy of space means the Lagrangian must also be independent of the direction, and must therefore be only a function of its magnitude, $\dot{\boldsymbol{x}}^2 = v^2$:

$$L = L(v^2). \tag{13}$$

The Euler-Lagrange equations (10) are particularly simple in this case,[7]

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \boldsymbol{v}}\right) = 0, \tag{14}$$

since the Lagrangian is independent of $\boldsymbol{x}$ so we have $\partial L/\partial \boldsymbol{x} = 0$. From this, we see that $\partial L/\partial \boldsymbol{v}$ is a constant. Since L is a function only of $\dot{x}$, this implies the *velocity is constant*:

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{v} = 0. \tag{15}$$

In other words, in the absence of any potential terms, the velocity of a 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.

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; that coordinates change according to Galilean transformations (1) when changing between inertial reference frames[8], and that the equations of motion have the same form in every inertial frame. Consider two frames of reference, moving relative to each other by infinitesimal velocity $\delta\boldsymbol{v}$, such that $\boldsymbol{v}' = \boldsymbol{v} + \delta\boldsymbol{v}$. Under our assumptions, $L' = L(v'^2)$ must differ from $L(v^2)$ by at most a total time derivative. 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

$$L(v'^2) = L(v^2) + \frac{\partial L}{\partial(v^2)} 2\delta\boldsymbol{v}\cdot\boldsymbol{v}. \tag{16}$$

[7] 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 $\mathrm{d}\phi/\mathrm{d}\boldsymbol{x} \equiv \nabla\phi$.

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

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. (12). This term is a total time derivative *only* if it is linear in $\boldsymbol{v}$ (since $\boldsymbol{v} = \dot{\boldsymbol{x}}$), and so $\partial L/\partial(v^2)$ must be independent of velocity. We arbitrarily set $\partial L/\partial(v^2) = m/2$ and integrate to find

$$L = \frac{1}{2}mv^2. \tag{17}$$

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 *mass*. 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).

We note that we did not explicitly invoke the assumption of reversibility. Since the Lagrangian does not depend explicitly on t, it follows that the equations of motion remain the same on substitution $t' = -t$. In this sense, the reversibility follows from homogeneity.

Notice that, in this free particle case, we have

$$\frac{\partial L}{\partial \dot{\boldsymbol{x}}} = m\boldsymbol{v}, \tag{18}$$

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 generalised coordinates $\{q_i\}$, then we can define a *canonical momentum*

$$p_i \equiv \frac{\partial L}{\partial \dot{q}}, \tag{19}$$

which is also called the *generalised momentum*, or the momentum conjugate to q. In general, this quantity is not simply $m\boldsymbol{v}$. With this, the Euler-Lagrange equations may be written:

$$\frac{\mathrm{d}p_i}{\mathrm{d}t} = \frac{\partial L}{\partial q_i}. \tag{20}$$

This is just a definition for now; we will return to this in our discussion of symmetries.

We will now consider a *closed* system of several particles. By closed, we mean that nothing outside the considered system may impact its dynamics. In this case, the strict homogeneity/isotropy for each particle is broken, by the presence of the other particles. Therefore, and extra term may appear in the Lagrangian, that may depend on the positions and velocities of all the particles. We'll call it the *potential*, V:

$$L = \sum_i \frac{1}{2}m_i\boldsymbol{v}_i^2 - V(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{v}_1, \boldsymbol{v}_2, \ldots), \tag{21}$$

where the choice of negative sign is arbitrary. For simplicity, we'll directly consider the case for two particles; the arguments will hold generally. The equations of motion can be determined from an application of Eq. (10). In this case, the equation of motion for the ith particle is

$$m_i\boldsymbol{a}_i = -\frac{\partial V}{\partial \boldsymbol{x}_i}, \tag{22}$$

which you will recognise as Newton's second law. We'll call the derivative term on the right-hand-side the *force* on the ith particle. This justifies our choice for the definition of the mass constant, m.

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

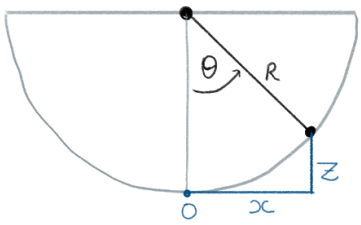

Figure 3: A small mass, free to slide inside a spherical bowl.

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:

$$V(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots) = V(\boldsymbol{x}_1 - \boldsymbol{x}_2, \ldots). \tag{23}$$

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

$$\frac{\partial V}{\partial \boldsymbol{x}_1} = -\frac{\partial V}{\partial \boldsymbol{x}_2}. \tag{24}$$

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:

$$L = T - V, \tag{25}$$

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.

1.3.1 Elementary examples

As a simple example showing to 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. 3. The bowl in 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 $z = R - R\cos\theta$, $x = R\sin\theta$), 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. The kinetic energy of the particle is

$$T = \frac{1}{2}m(\dot{x}^2 + \dot{z}^2) = \frac{mR^2}{2}\dot{\theta}^2, \tag{26}$$

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

$$V = mgz = mgR(1 - \cos\theta). \tag{27}$$

The Euler-Lagrange equation is thus

$$\ddot{\theta} = -\frac{g}{R}\sin\theta. \tag{28}$$

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 such that $\sin\theta \approx \theta$, you will recognise the result at 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

$$x = x_0\cos(\sqrt{g/R}\,t). \tag{29}$$

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

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" field (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 field. 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 constant:

$$\boldsymbol{x} \to \boldsymbol{x}' = \boldsymbol{x} + \boldsymbol{\epsilon}. \tag{30}$$

The *epsilon* is to make explicit that we consider an infinitesimal shift. 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

$$\delta L = \sum_i \frac{\partial L}{\partial \boldsymbol{x}_i} \cdot \delta \boldsymbol{x}_i = \sum_i \frac{\partial L}{\partial \boldsymbol{x}_i} \cdot \boldsymbol{\epsilon}. \tag{31}$$

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

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

and so the Euler-Lagrange equations become:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\sum \frac{\partial L}{\partial \dot{\boldsymbol{x}}_i}\right) \equiv \frac{\mathrm{d}}{\mathrm{d}t}\left(\sum \boldsymbol{p}_i\right) = 0, \tag{32}$$

where we used the definition of canonical momentum (19). For a free particle (or any particle where the potential is independent of 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.

1.4.2 Time-translation symmetry: energy conservation

We now consider a translation in time:

$$t \to t' = t + \epsilon. \tag{33}$$

The corresponding change in the Lagrangian is

$$\delta L = \frac{\partial L}{\partial t}\epsilon \tag{34}$$

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:

$$\frac{\mathrm{d}L}{\mathrm{d}t} = \sum_i \left(\frac{\partial L}{\partial q_i}\dot{q}_i + \frac{\partial L}{\partial \dot{q}_i}\ddot{q}_i \right) + \frac{\partial L}{\partial t} \tag{35}$$

$$= \sum_i \left(\frac{\mathrm{d}p}{\mathrm{d}t}\dot{q}_i + \dot{p}_i\ddot{q}_i \right) + \frac{\partial L}{\partial t} \tag{36}$$

$$= \frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_i p_i\dot{q}_i \right) + \frac{\partial L}{\partial t}, \tag{37}$$

where we used the Euler-Lagrange equations (20) to replace $\partial L/\partial q$, and the canonical momentum (19) to replace $\partial L/\partial \dot{q}$. We define a new quantity, called the *Hamiltonian* (or the *energy*):

$$H \equiv \sum_i p_i\dot{q}_i - L. \tag{38}$$

With this definition, the above equation reads:

$$\frac{\mathrm{d}H}{\mathrm{d}t} = -\frac{\partial L}{\partial t}. \tag{39}$$

We this see that the Hamiltonian (or the energy) is conserved if the Lagrangian has no explicit time dependence, which from Eq. (34), is implied if there is time translation symmetry.

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

In the case where the Lagrangian is simple $L = T - V$, 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. (38) 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 rotational translation 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:

$$\delta x = -\delta\theta\, y, \quad \delta y = \delta\theta\, x, \tag{40}$$

see Fig. 4. The generalisation is the cross-product

$$\delta\boldsymbol{r} = \delta\boldsymbol{\theta} \times \boldsymbol{r}, \tag{41}$$

where $\delta\boldsymbol{\theta}$ is a vector with magnitude $\delta\theta$ that points along the axis of rotation. Notice that *directions*, not just positions, change with this transformation. We therefore must also update the velocity vectors, which change in the same way[9]:

$$\delta\boldsymbol{v} = \delta\boldsymbol{\theta} \times \boldsymbol{v}. \tag{42}$$

[9]Since $v = \dot{x}$, $\delta v = \delta(\dot{x})$, and $\delta\theta$ is independent of time.

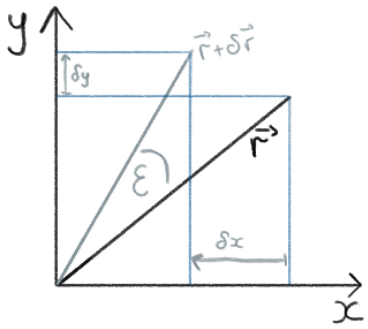

Figure 4: Small rotation, ϵ, about the z-axis.

The corresponding change in the Lagrangian is

$$\delta L = \sum_i \left(\frac{\partial L}{\partial \boldsymbol{r}_i} \cdot \delta \boldsymbol{r}_i + \frac{\partial L}{\partial \boldsymbol{v}_i} \cdot \delta \boldsymbol{v}_i \right) \tag{43}$$

$$= \sum_i \left[\dot{\boldsymbol{p}}_i \cdot (\delta \boldsymbol{\theta} \times \boldsymbol{r}_i) + \boldsymbol{p}_i \cdot (\delta \boldsymbol{\theta} \times \boldsymbol{v}_i) \right] \tag{44}$$

$$= \delta \boldsymbol{\theta} \cdot \frac{\mathrm{d}}{\mathrm{d}t} \sum_i \boldsymbol{r}_i \times \boldsymbol{p}_i, \tag{45}$$

where we used the Euler-Lagrange equations, and the permutation properties of the cross-product. 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*:

$$\boldsymbol{l} = \boldsymbol{r} \times \boldsymbol{p} \tag{46}$$

(some places use $\boldsymbol{L}$ or $\boldsymbol{M}$). Rotational symmetry about an axis i implies the i-component of 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:

$$\delta q_i = f_i(q)\epsilon. \tag{47}$$

In general, the velocities will also change under the transformation:

$$\delta \dot{q}_i = \frac{\mathrm{d}}{\mathrm{d}t}(\delta q_i). \tag{48}$$

The general change in the Lagnrangian is then

$$\delta L = \sum_i \left(\frac{\partial L}{\partial q_i} \delta q_i + \frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i \right) = \sum_i \left(\dot{p}_i \delta q_i + p_i \delta \dot{q}_i \right) \tag{49}$$

$$= \frac{\mathrm{d}}{\mathrm{d}t} \sum_i p_i \delta q_i, \tag{50}$$

where we used Eqs. (19) and (20) 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:

$$\delta L(f) = 0 \quad \Longrightarrow \quad \frac{\mathrm{d}}{\mathrm{d}t}Q = 0, \tag{51}$$

$$\text{where} \quad Q \equiv \sum_i p_i f_i(q), \tag{52}$$

which is a statements of Noether's theorem. Notice the small parameter ϵ is not included in the definition of Q.

1.5 Hamiltonian formulation

In Eq. (38), 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. (38)

$$\delta H = \sum_i \left(p_i \delta\dot{q}_i + \delta p_i \dot{q}_i\right) - \delta L \tag{53}$$

$$= \sum_i \left(p_i \delta\dot{q}_i + \delta p_i \dot{q}_i - \frac{\partial L}{\partial q_i}\delta q_i - \frac{\partial L}{\partial \dot{q}_i}\delta\dot{q}_i\right) \tag{54}$$

$$= \sum_i \left(\cancel{p_i \delta\dot{q}_i} + \delta p_i \dot{q}_i - \dot{p}_i \delta q_i - \cancel{p_i \delta\dot{q}_i}\right) \tag{55}$$

where we used $L = L(\{q\}, \{\dot{q}\})$, and the Euler-Lagrange equations. At the same time, we have

$$\delta H = \sum_i \left(\frac{\partial H}{\partial q_i}\delta q_i + \frac{\partial H}{\partial p_i}\delta p_i\right), \tag{56}$$

which holds for any general function of qs and ps. Equations (55) and (56) are equivalent. Equating these, and matching terms, we find

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad \text{and} \quad \dot{p}_i = -\frac{\partial H}{\partial q_i}, \tag{57}$$

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.

1.6 Harmonic Oscillator

Consider a Lagrangian of the form

$$L = \frac{m\dot{x}^2}{2} - \frac{k}{2}x^2, \tag{58}$$

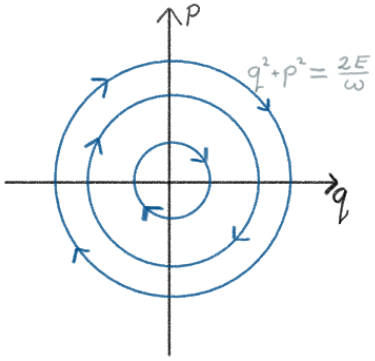

Figure 5: Simple harmonic oscillator tracing closed paths in phase space. Each path is for a constant energy. Each point on the plot corresponds to a possible set of (q_0, p_0) initial conditions; from there, a concentric circle will be traced.

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:

$$L = \frac{\dot{q}^2}{2\omega} - \frac{\omega}{2}q^2. \tag{59}$$

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

$$\ddot{q} = -\omega^2 q. \tag{60}$$

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. (38):

$$H = \frac{\omega}{2}(p^2 + q^2). \tag{61}$$

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:

$$\dot{q} = \omega p, \quad \text{and} \quad \dot{p} = -\omega q. \tag{62}$$

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. (61) 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. 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.

1.7 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

$$\dot{f} = \frac{\partial f}{\partial t} + \sum_i \left(\frac{\partial f}{\partial q_i}\dot{q}_i + \frac{\partial f}{\partial p_i}\dot{p}_i \right), \tag{63}$$

which from Hamilton's equations, can be expressed:

$$\dot{f} = \frac{\partial f}{\partial t} + \sum_i \left(\frac{\partial f}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial g}{\partial p_i}\frac{\partial H}{\partial q_i} \right). \tag{64}$$

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

$$[f, g] \equiv \sum_i \left(\frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i} \right). \tag{65}$$

In this case, time derivatives can be expressed:

$$\dot{f} = \frac{\partial f}{\partial t} + [f, H]. \tag{66}$$

When it is not clear from context which variables are used for the derivatives, these are given as subscripts. For example, the above (65) would be $[f, g]_{qp}$. Generally:

$$[f, g]_{uv} \equiv \sum_i \left(\frac{\partial f}{\partial u_i}\frac{\partial g}{\partial v_i} - \frac{\partial f}{\partial v_i}\frac{\partial g}{\partial u_i} \right). \tag{67}$$

Clearly, the Poisson bracket is asymmetric: $[A, B] = -[B, A]$. It's also fairly clear to show the linearity properties:

$$[\lambda A, B] = \lambda[A, B], \quad \text{and} \quad [A + D, C] = [A, C] + [D, C]. \tag{68}$$

Also, from the product rule, we have:

$$[AB, C] = A[B, C] + [A, C]B. \tag{69}$$

Finally, the Poisson brackets of the canonical variables are

$$[q_i, q_j] = 0, \qquad [p_i, p_j] = 0, \qquad [p_i, q_j] = \delta_{ij}. \tag{70}$$

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

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

$$\dot{q}_i = [q_i, H], \qquad \dot{p}_i = [p_i, H]. \tag{71}$$

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

$$[f, q_i] = -\frac{\partial f}{\partial p_k}, \qquad [f, p_i] = \frac{\partial f}{\partial q_k} \tag{72}$$

These can all be proven without much effort.

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

$$[f, [g, h]] + [g, [h, f]] + [h, [f, g]] = 0. \tag{73}$$

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 (71) 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.

1.8 Generators of transformations

Back in Eq. (40), we considered a small rotation of $\delta\theta$ about the z-axis, and saw the corresponding change in x, y, z was $\delta x = -\delta\theta y$, $\delta y = \delta\theta x$, $\delta z = 0$. We saw that, if the system was symmetric with respect to rotation around z, then the z-component of angular momentum, $l_z = xp_y - y_p x$, was conserved. It's instructive to take the Poisson bracket of the coordinates with respect to the conserved l_z

$$[x, l_z] = -y, \quad [y, l_z] = x, \quad [z, l_z] = 0. \tag{74}$$

Notice that this can be written:

$$\delta\theta[x, l_z] = \delta x, \quad \delta\theta[y, l_z] = \delta y, \quad \delta\theta[z, l_z] = \delta z. \tag{75}$$

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. (41),[10]

$$[x_i, l_j] = \sum_k \epsilon_{ijk} x_k. \tag{76}$$

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

$$[p_i, l_j] = \sum_k \epsilon_{ijk} p_k. \tag{77}$$

This is the same, since momentum vectors transform under rotations the same way as positions. The change in any quantity, f, about the i-axis may be written:

$$\delta f = [F, L_i]. \tag{78}$$

This holds rather generally. Above, we saw that invariance under spatial translations implied momentum conservation. Consider, then, the Poisson bracket of any function position with p

$$[f(q), p] = \frac{\mathrm{d}f}{\mathrm{d}q}. \tag{79}$$

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

$$\delta f = \epsilon[f, p]. \tag{80}$$

We may recognise momentum as the generator of spatial translations. Doing the same with time translation will show that the Hamiltonian is the generator of time translations.

We may generalise this. Let $G(q, p)$ be a general function of generalised coordinates and momentum, defined such that it gives small changes to coordinates

$$\delta q_i = [q_i, G], \quad \text{and} \quad \delta p_i = [p_i, G]. \tag{81}$$

This transformation 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, the condition that G represents a symmetry is

$$[H, G] = 0. \tag{82}$$

We could also write this the other way:

$$[G, H] = 0, \tag{83}$$

which, since H is the generator of time translations, tells us that G is constant.

[10] ϵ_{ijk} is the entirely asymmetric Levi-Civita symbol. It is zero if any of the indices repeat, is +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 $c_i = \sum_{jk} a_j b_k \epsilon_{ijk}$.

1.9 Appendix: Calculus of variations

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

$$\frac{\mathrm{d}f}{\mathrm{d}y} \equiv \lim_{\Delta y \to 0} \frac{\Delta f}{\Delta y}. \tag{84}$$

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:

$$\delta f \equiv \frac{\mathrm{d}f}{\mathrm{d}y} \delta y. \tag{85}$$

If f is function of multiple variables x_1, x_2, etc., each of which may depend on y, i.e.: $f(y) = f(x_1(y), x_2(y), \ldots)$, then, by the chain rule, we have

$$\frac{\mathrm{d}f}{\mathrm{d}y} = \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial y} + \frac{\partial f}{\partial x_2}\frac{\partial x_2}{\partial y} + \ldots \tag{86}$$

Combining Eqs. (85) and (86), we have

$$\delta f = \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial y}\delta y + \frac{\partial f}{\partial x_2}\frac{\partial x_2}{\partial y}\delta y + \ldots \tag{87}$$

$$= \frac{\partial f}{\partial x_1}\delta x_1 + \frac{\partial f}{\partial x_2}\delta x_2 + \ldots, \tag{88}$$

which is the general formula for the *variation* in f.

References

[1]H. Goldstein, C. Poole, and J. Safko, *Classical Mechanics*, 3rd ed. (Addison Wesley, 2001).

[2]L. Susskind and G. Hrabovsky, *Classical Mechanics: The Theoretical Minimum* (Penguin, 2014).

[3]L. D. Landau and E. M. Lifshitz, *Mechanics (L&L Vol. 1)* (Pergamon, Oxford, 1976).