Part II Classical Field Theory

William Royce

September 13, 2024

Part II Physics, The University of Cambridge

Preface

Contents

List of Tables

List of Figures

Chapter 1

Basic Lagrangian Mechanics

The initial purpose of Lagrangian mechanics is to express the relevant equations of motion, essentially Newton's laws, in a form involving a set $q_1, q_2, \cdots, q_n$ of generalised position coordinates, and their first time-derivatives $\dot{q}_1, \dot{q}_2, \cdots, \dot{q}_n$. The n-component vector $\{\mathbf{q}\}$ can represent any physical system or process, as long as this set of numbers completely describes the state of the system (n is the number of degrees of freedom). In the most straightforward case, these can be the 3 Cartesian coordinates of a material point (a "particle"), but, say, the spherical polar set (r, θ, ϕ) is just as good. Just as $\{\mathbf{q}\}$ is the generalised coordinate vector, so is $\{\dot{\mathbf{q}}\}$ the generalised velocity, both explicit functions of time. It is assumed that simultaneous knowledge of $\{\mathbf{q}(t)\}$ and $\{\dot{\mathbf{q}}(t)\}$ completely defines the mechanical state of the system. Mathematically, this means that the complete set of $\{\mathbf{q}(t)\}$ and $\{\dot{\mathbf{q}}(t)\}$ also determines the accelerations $\{\ddot{\mathbf{q}}(t)\}$. The mathematical relations that relate accelerations with coordinates and velocities are what one calls the equations of motion.

In many cases, not all n degrees of freedom are completely free. A system may have *constraints*; for example $q_1 = \text{const.}$, $q_2 = \text{const.}, \cdots, q_r = \text{const.}$ could represent the r constraints and $q_{r+1}, \cdots, q_n$ the remaining independent coordinates. Most often the choice of generalised coordinates $\{\mathbf{q}\}$ is dictated by the nature of the constraints. For instance, if a particle is constrained to move on the surface of an expanding balloon of radius $R = a\sqrt{t}$, we might use spherical polar coordinates, scaled such that $q_1 = r/a\sqrt{t}, q_2 = \theta, q_3 = \phi$; in that case the single constraint is expressed as $q_1 = 1$ (it would look a lot more complicated if we tried to express it in Cartesians). The Lagrangian formalism is developed, partially, to enable one to deal efficiently with the sometimes complicated constraints imposed on the evolution of physical systems. Constraints are called **holonomic** if they are of the form $g(q_1, q_2, \cdots, q_n, t) = 0$. We shall shortly return to their treatment, but first, let us revise some basic starting points.

1.1 Hamilton's Principle

A very general formulation of the equations of motion of mechanical (and many other) systems is given by Hamilton's *Principle of Least Action*. It states that every mechanical system can be characterised by a certain function

$$L(q_1, q_2, \cdots, q_n; \dot{q}_1, \dot{q}_2, \cdots, \dot{q}_n; t) \equiv L(q, \dot{q}; t). \tag{1.1}$$

Theorem 1 (Hamilton's Principle)**.** *The actual motion of a system from A to B is that which makes the integral $S = \int_A^B L \, \mathrm{d}t$ a minimum.*

The function $L(q, \dot{q}; t)$ is called the Lagrangian of the given system and the integral S defined in Theorem 1 is called the **action**. Later in this course we shall have some deeper

insights into what this object, the action functional $S[\mathbf{q}(t)]$, represents and why it has to be minimal. For the time being let us take this as an axiom.

The principle of least action implies that, with a sufficient command of mathematics, in particular the calculus of variations, the solution of any mechanical problem is achieved by the following recipe:

'Minimise $S = \int_A^B L\,\mathrm{d}t$ for fixed starting and finishing (representative) points, $A = (\mathbf{q}_A, t_A)$ to $B = (\mathbf{q}_B, t_B)$, taking proper account of all the constraints.'

Remark. *There is no worldline whose action is a true maximum, that is, for which* $\delta^2 S < 0$ *or more generally for which* $S - S_0 < 0$ *for every adjacent curve.*

Proof. Here follows an intuitive proof by contradiction for the Lagrangian $L = T - V$ with T positive definite. Consider an actual worldline for which it is claimed that S is a true maximum. Now modify this worldline by adding wiggles somewhere in the middle. These wiggles are to be of very high frequency and very small amplitude so that they increase the kinetic energy T compared to that along the original worldline with only a small change in the corresponding potential energy V. The Lagrangian $L = T - V$ for the region of wiggles is larger for the new curve and so is the overall time integral S. The new worldline has greater action than the original worldline, which we claimed to have maximum action. Therefore S cannot be a true maximum for any actual worldline. □

1.2 Derivation of the Equations of Motion

First, let's examine the "standard derivation" based on d'Alembert's principle: consider a particle that is subject to the total force $\mathbf{F}$ and has momentum $\mathbf{p}$. Then if we construct a vector $(\mathbf{F} - \dot{\mathbf{p}})$, this vector will always be perpendicular to the instantaneous line of motion. In other words, the scalar product is zero:

$$\sum_i (F_i - \dot{p}_i)\delta x_i = 0. \tag{1.2}$$

That's almost trivially true for an arbitrary set of coordinate variations δx_i because Newton's second law ($\mathbf{F}^{\text{total}} = m\ddot{\mathbf{r}}$ for each particle) makes each $(F_i - \dot{p}_i) = 0$. However, we shall only be interested in sets of displacements δx_i consistent with the constraints. Constraints exert their own forces on each particle, which we call internal: see the reaction force R exerted by the wire in Fig. **??**. By definition of the constraint, these internal forces are perpendicular to the line of motion, that is $\sum_i F_i^{\text{internal}} \delta x_i = 0$. Therefore, d'Alembert's principle states

$$\sum_i \left(F_i^{\text{internal}} - \dot{p}_i\right)\delta x_i = 0. \tag{1.3}$$

Let us try rewriting this in an arbitrary set of generalised coordinates $\{\mathbf{q}\}$ to which the Cartesians $\{\mathbf{r}\}$ could be transformed via matrices $\partial q_i/\partial x_j$. The aim is to present Eq. (1.3) as a generalised scalar product $\sum_j (\text{something})\delta q_j = 0$, so that we can say this is true for arbitrary sets of variations δq_i of the reduced number $(n - r)$ of generalised coordinate that are not subject to the constraints.

Fig. 1.1: An example of a constraint, restricting the motion of a particle (which may be subject to external forces, such as gravity) along a specific path: the bead on a wire.

The coordinate transformation in the first term, involving the external force, is easy:

$$\sum_i F_i \delta x_i = \sum_{i,j} F_i \frac{\partial x_i}{\partial q_j} \delta q_j \equiv \sum_j Q_j \delta q_j. \tag{1.4}$$

One must take great care over precisely what partial differentials mean. In the following, $\partial/\partial q_j$ means evaluating $(\partial/\partial q_j)$ with the other components $q_{i\neq j}$, all velocities $\dot{q}_i$ and time t held constant.

It is clear that $\partial x_i/\partial q_j$ should mean $(\partial x_i/\partial q_j)_{\text{all other } q,t}$; holding the $\dot{q}_i$ constant only becomes relevant when we differentiate a velocity w.r.t. q_j – a velocity component changes with $\mathbf{q}$ for fixed $\dot{\mathbf{q}}$ because the conversion factors from the $\dot{q}_j$ to the $\dot{r}_i$ change with position. Similarly, $\partial/\partial t$ means $(\partial/\partial t)_{\mathbf{q},\dot{\mathbf{q}}}$, e.g. $\partial x_i/\partial t$ refers to the change in position, for *fixed* $\mathbf{q}$ and $\dot{\mathbf{q}}$, due to the prescribed motion of the q-coordinate system.

Dealing with the second term, involving the rate of change of momentum, is a bit harder – it takes a certain amount of algebra to manipulate it into the required form. First, by definition of momentum in Cartesians:

$$\sum_i \dot{p}_i \delta x_i = \sum_{i,j} m_i \dot{v}_i \frac{\partial x_i}{\partial q_j} \delta q_j. \tag{1.5}$$

We shall need

$$v_i \equiv \dot{x}_i = \sum_j \frac{\partial x_i}{\partial q_j} \dot{q}_j + \frac{\partial x_i}{\partial t}, \quad \text{whence} \quad \frac{\partial v_i}{\partial \dot{q}_j} = \frac{\partial x_i}{\partial q_j}. \tag{1.6}$$

Now we are in a position to start work on the second term. The relevant product is

$$\dot{v}_i \frac{\partial x_i}{\partial q_j} = \frac{\mathrm{d}}{\mathrm{d}t}\left(v_i \frac{\partial x_i}{\partial q_j} \right) - v_i \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial x_i}{\partial q_j} \right). \tag{1.7}$$

Further transforming the second term in (1.7):

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial x_i}{\partial q_j} \right) &= \frac{\partial}{\partial q_k}\left(\frac{\partial x_i}{\partial q_j} \right) \dot{q}_k + \frac{\partial}{\partial t}\left(\frac{\partial x_i}{\partial q_j} \right) \quad \text{(summed over } k\text{)} \\ &= \frac{\partial}{\partial q_j}\left[\left(\frac{\partial x_i}{\partial q_k} \right) \dot{q}_k + \frac{\partial x_i}{\partial t} \right] \\ &= \left(\frac{\partial v_i}{\partial q_j} \right)_{\text{other } q,\dot{q},t} \end{aligned} \tag{1.8}$$

Using (1.6) on the first term and (1.8) on the second term of (1.7) we finally get

$$\begin{aligned} \sum_i \dot{p}_i \delta x_i &= \sum_{i,j} \left[\frac{\mathrm{d}}{\mathrm{d}t}\left(m_i v_i \frac{\partial v_i}{\partial \dot{q}_j} \right) - m_i v_i \frac{\partial v_i}{\partial q_j} \right] \delta q_j \\ &= \sum_j \left[\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{q}_j} \right) - \frac{\partial T}{\partial q_j} \right] \delta q_j, \end{aligned} \tag{1.9}$$

which is equal to $\sum_j Q_j \delta q_j$, from Eq. (1.4). Here the total kinetic energy of the system has been defined from the Cartesian representation $T = \sum_i \frac{1}{2} m_i v_i^2$. The last equation is a consequence of D'Alembert's principle. Since the components δq_j allowed by the constraints are all independent, it follows that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} = Q_j. \tag{1.10}$$

In many systems the external forces are gradients of a scalar potential; in Cartesians:

$$F_i = -\frac{\partial V}{\partial x_i}, \quad \text{where} \quad V = V(\mathbf{x}, t), \tag{1.11}$$

(i.e. V is independent of the particle velocities), so that

$$Q_j = -\sum_i \frac{\partial V}{\partial x_i}\frac{\partial x_i}{\partial q_j} = -\frac{\partial V}{\partial q_j}. \tag{1.12}$$

Therefore, substituting this vector $\mathbf{Q}$ into the r.h.s. of Eq. (1.10) and using the fact that it does not depend on $\dot{\mathbf{q}}$, we can write

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} = 0, \tag{1.13}$$

where $L \equiv T - V$, but by construction it is a function of the generalised coordinates q, $\dot{q}$ and not the original Cartesians. We discover that this function is exactly the Lagrangian of the system, the one used in the definition of the action in Hamilton's principle. Indeed, (1.13) is the differential equation that one obtains by the calculus of variations from the condition $\delta S = 0$ for the minimum of the action.

Now let's reverse the argument. Define a function $L = T - V$. In Cartesian coordinates, an N-particle system moving in a potential $V(x_1, x_2, \cdots, x_{3N})$ has

$$p_i \equiv m_i v_i = \frac{\partial}{\partial v_i}\left(\frac{1}{2} m_i v_i^2\right) = \frac{\partial T}{\partial v_i} \equiv \frac{\partial L}{\partial v_i}, \tag{1.14}$$

and

$$F_i = -\frac{\partial V}{\partial x_i} = \frac{\partial L}{\partial x_i}, \tag{1.15}$$

provided that V is independent of velocities, so the equations of motion $\dot{\mathbf{p}} = \mathbf{F}$ are

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{x}_i}\right) - \frac{\partial L}{\partial x_i} = 0. \tag{1.16}$$

Therefore the motion obeys Hamilton's principle of least action. But Hamilton's principle is a statement independent of any particular coordinate system; it is true in any coordinate system. Therefore write down the Euler-Lagrange equations for Hamilton's principle in our new q coordinate system (in which constraints are of the kind $q_j = \text{const.}$):

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0, \tag{1.17}$$

a much more memorable way to derive Lagrange's equations.

But what about the constraints? We must invent some local potentials near the actual path of the system, whose gradients are perpendicular to the actual path and just right to keep the system on the constrained trajectory, see Fig. 1.1. The important thing is that they will have no gradient along the local 'directions' of the q_is, allowed by the constraints, and will therefore not affect the dynamics for those q_is.

1.2.0.1 What is Lagrangian Mechanics Good For?

Lagrangian mechanics will do nothing that Newtonian mechanics won't do. It's just a reformulation of the same physics. In fact, it will do slightly less, because some problems (notoriously, the motion of a bicycle) have what is called non-holonomic constraints. One thing that one can say for the Lagrangian formulation is that it involves scalars (T and V) instead of vectors (forces, couples) which makes it less confusing to use in messy problems. Several examples will be treated in the lectures and, more particularly, in the examples classes. Figure 1.2 gives a few simple examples of dynamical systems with constraints. Use them to practise: in each case first write the full potential (in all cases it's gravity) and kinetic energies (don't forget the moment of inertia for the rotating cylinder), then implement the constraint (for this you need to have the appropriate choice of coordinates) and write the resulting Lagrangian, as well as the dynamical equation(s).

Fig. 1.2: Examples of dynamical systems with holonomic constraints (in each case their expression is in the frame). In each case the number of coordinates is reduced (for the pulley and rolling cylinder Cartesians are sufficient, the circular wire requires plane polar coordinates and rolling on the inside of a cone, spherical polar coordinates; the last case has two free generalised coordinates left).

1.2.0.2 A Few More Remarks

So far we have assumed forces depend only on position: $F_i = -\,\partial V(\mathbf{x},t)/\partial x_i$. For *some* **velocity-dependent** forces one can still use Lagrange's equations; it's possible if you can find a $V(\mathbf{x},\dot{\mathbf{x}},t)$ such that

$$F_i = -\frac{\partial V(\mathbf{x},\dot{\mathbf{x}},t)}{\partial x_i} + \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial V(\mathbf{x},\dot{\mathbf{x}},t)}{\partial \dot{x}_i}\right). \tag{1.18}$$

To derive this condition you should repeat the steps between Eqs.(1.4) and (1.13), only now allowing a contribution from V in the $\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\right)$ terms of the Euler-Lagrange equations. The hard bit is to find a V that satisfies such a condition. A bit later we shall examine the case of magnetic forces, which is a good example of such a potential.

There is also an issue of uniqueness of the Lagrangian function. Clearly the condition of zero variation $\delta[S] = 0$ can be maintained if we:

- add any constant to V;
- multiply L by any constant;
- add a total time derivative, $f = \frac{\mathrm{d}}{\mathrm{d}t}g(q,\dot{q},t)$;

etc. Usually the convention is to take $L = T - V$, with no "additives", which then corresponds to the "classical action S".

1.3 Symmetry and Conservation Laws; Canonical Momenta

Symmetry is one of the most powerful tools used in theoretical physics. In this section we will show how symmetries of L correspond to important conservation laws. This theme will be taken further later in this course, and in subsequent solid state and particle physics courses.

When L does not depend explicitly on one of the q_i then Lagrange's equations show that the corresponding $\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\right)$ is zero, directly from Eq. (1.13). Hence we can define an object, $p_i \equiv \frac{\partial L}{\partial \dot{q}_i}$, which is conserved for such a system and q_i is called an ignorable coordinate. The object p_i is generally referred to as the *canonical momentum* corresponding to the coordinate q_i. For a particle moving in Cartesian coordinates this would be the "ordinary" momentum component: for each x_i this is just $p_i = m\dot{x}_i$. However, in generalised coordinates, the physical meaning of each component of the canonical momentum may be very different. In particular, since a generalised coordinate can have any dimensions, the dimensions of the corresponding canonical momentum need not be those of ordinary momentum.

1.3.1 Translational Invariance $\leftrightarrow$ Conservation of Linear Momentum

Suppose L does not depend on the position of the system as a whole, i.e. we can move the position of every particle by the *same* vector ϵ without changing L. Suppose we move the whole system by δx in the x direction (in Cartesian coordinates!). Then

$$L \to L + \sum_n \frac{\partial L}{\partial x_n}\delta x. \tag{1.19}$$

(N.B. In this section and the next we sum over particles, assuming only 1-dimensional motion for simplicity: n is an index of summation over the particles. Also, V in $L = T - V$ now includes the mutual potential energies that impose e.g. the constraints that keep rigid bodies rigid.) Therefore if L does not change

$$\sum_n \frac{\partial L}{\partial x_n} = 0, \quad \text{and so, by Lagrange's equations,} \quad \sum_n \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{x}_n}\right) = 0, \tag{1.20}$$

i.e. if L is independent of the position of the system $\sum \partial L/\partial \dot{x}_n$ is constant.

Now if V depends on particle positions only, then

$$\sum \frac{\partial L}{\partial \dot{x}_n} = \sum m_n \dot{x}_n \tag{1.21}$$

is just the x-component of the total momentum of the system, and we conclude

$$\text{Homogeneity of space} \implies \text{conservation of linear momentum.} \tag{1.22}$$

N.B.: This is clearly not true for velocity-dependent potentials, e.g. for charged ions moving in magnetic fields.

1.3.2 Rotational Invariance ↔ Conservation of Angular Momentum

Suppose L is independent of the orientation of the system. In particular, suppose L is invariant under rotation of the whole system about the z-axis; then, proceeding as before, we can assume a rotation by $\delta\theta$ and use cylindrical coordinates (r, θ, z) to obtain

$$\sum_n \frac{\partial L}{\partial \theta_n}\delta\theta = 0 \implies \sum_n \frac{\partial L}{\partial \theta_n} = 0 \tag{1.23}$$

$$\implies \sum_n \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{\theta}_n}\right) = 0 \implies \sum_n \frac{\partial L}{\partial \dot{\theta}_n} = \text{const.} \tag{1.24}$$

Again, when V depends on particle positions only,

$$\begin{aligned}\frac{\partial L}{\partial \dot{\theta}_n} &= \frac{\partial}{\partial \dot{\theta}_n}\left(\frac{1}{2}m_n\dot{r}_n^2 + \frac{1}{2}m_n r_n^2 + \dot{\theta}_n^2 + \frac{1}{2}m_n\dot{z}_n^2\right)\\ &= m_n r_n^2\dot{\theta}_n \qquad (1.25)\\ &= \text{ angular momentum of } n^{\text{th}} \text{ particle about } z\text{-axis} \qquad (1.26)\end{aligned}$$

Thus the total angular momentum about the z axis is conserved.

$$\text{Isotropy of space} \implies \text{conservation of angular momentum.} \tag{1.27}$$

1.3.3 Time Invariance ↔ Conservation of Energy

This is a bit more complicated and it also gives us a chance to explore a very useful mathematical result called Euler's homogeneous function theorem. If L does not depend explicitly on time, i.e.

$$L = L(q_i\dot{q}_i) \implies \frac{\partial L}{\partial t} = 0, \tag{1.28}$$

then the total time derivative of the Lagrangian is

$$\begin{aligned}\frac{\mathrm{d}L}{\mathrm{d}t} &= \sum_i \frac{\partial L}{\partial q_i}\dot{q}_i + \sum_i \frac{\partial L}{\partial \dot{q}_n}\ddot{q}_n \qquad (1.29)\\ &= \sum_i \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial q_i}\right)\dot{q}_i + \sum_i \frac{\partial L}{\partial \dot{q}_i}\ddot{q}_n \quad \text{(using the E-L equation)}\\ &= \sum_i \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\dot{q}_i\right) \quad \text{(assembling the total derivative from both terms),} \qquad (1.30)\end{aligned}$$

or, combining the total time derivatives from l.h.s. and r.h.s. into one expression, we have

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}H, \quad \text{where} \quad H = \sum_i \frac{\partial L}{\partial \dot{q}_i}\dot{q}_i - L = \text{a constant, } E \tag{1.31}$$

We shall now identify the new function H, the **Hamiltonian**, as the total energy of the system. Assume the generalised kinetic energy, T, is given by:

$$T = \sum_{i,j}\frac{1}{2}c_{ij}\dot{q}_i\dot{q}_j, \tag{1.32}$$

where we can generically choose $c_{ij} = c_{ji}$ (by symmetry of dummy indices of summation). The coefficients c_{ij} might be functions of $q_1, \cdots, q_n$ but not $\dot{q}_1, \cdots, \dot{q}_n$ or time. This quadratic form is also called a homogeneous second-order polynomial function of $\dot{\mathbf{q}}$. Assume also that the potential energy, V, is given by:

$$V = V(q_1, \cdots, q_n), \tag{1.33}$$

(velocity-independent) and $L = T - V$ as usual (with no explicit time-dependence as we agreed). Then:

$$\begin{aligned}
\sum_i \frac{\partial L}{\partial \dot{q}_i}\dot{q}_i &= \sum_i \dot{q}_i \frac{\partial}{\partial \dot{q}_i}\left(\sum_{j,k}\frac{1}{2}c_{jk}\dot{q}_j\dot{q}_k\right) \quad \text{since} \quad \frac{\partial V}{\partial \dot{q}_i} = 0 \qquad (1.34)\\
&= \sum_i \dot{q}_i \left(\sum_{j,k}\left[\frac{1}{2}c_{jk}\delta_{jk}\dot{q}_k + \frac{1}{2}c_{jk}\dot{q}_j\delta_{ik}\right]\right) \quad \text{since} \quad \frac{\partial \dot{q}_j}{\partial \dot{q}_i} = \delta_{ij}\\
&= \sum_i \dot{q}_i \left(\sum_k \frac{1}{2}c_{ik}\dot{q}_k + \sum_j \frac{1}{2}c_{ji}\dot{q}_j\right)\\
&= \sum_i \dot{q}_i \left(\sum_j c_{ij}\dot{q}_j\right) \quad \text{renaming the dummy index } k \to j\\
&= \sum_{i,j} c_{ij}\dot{q}_i\dot{q}_j\\
&= 2T. \qquad (1.35)
\end{aligned}$$

In effect, what we have just proven is that for any homogeneous quadratic function $T = T(q_i)$, the following property holds:

$$\sum_i q_i \frac{\partial T}{\partial q_i} = 2T \tag{1.36}$$

(an aspect of Euler's more general theorem; guess how this would change for linear, or cubic functions). Returning to our Hamiltonian, we have $H = 2T - L = 2T - (T - V) = T + V \equiv$ total energy $= E$, a constant from Eq. (1.31).

Chapter 2

Hamilton's Equations of Motion

We have already defined the i^{th} component of generalised (canonical) momentum as

$$p_i \equiv \left(\frac{\partial L}{\partial \dot{q}_i}\right)_{\text{other } \dot{q},\mathbf{q},t} \tag{2.1}$$

In Cartesians and for $V = V(\mathbf{r})$, $p_i = m_i v_i$.

Now try to re-write the equations of motion in terms of $\mathbf{q}$ and $\mathbf{p}$ instead of $\mathbf{q}$ and $\dot{\mathbf{q}}$. This operation is fully analogous to what is called the Legendre transformation in thermodynamics, when we change from one potential depending on a given variable to another, depending on its conjugate (like $T\,\mathrm{d}S \to -S\,\mathrm{d}T$ or $-p\,\mathrm{d}V \to V\,\mathrm{d}p$).

The Euler-Lagrange equations say

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\right) \equiv \dot{p}_i = \left(\frac{\partial L}{\partial q_i}\right)_{\text{other } q,\dot{\mathbf{q}},t}, \tag{2.2}$$

but that's not quite what we want, for it refers to $L(\mathbf{q}, \dot{\mathbf{q}}, t)$, not $L(\mathbf{q}, \mathbf{p}, t)$. We proceed thus:

$$\begin{aligned} \delta L &= \frac{\partial L}{\partial t}\delta t + \sum_i \frac{\partial L}{\partial q_i}\delta q_i + \sum_i \frac{\partial L}{\partial \dot{q}_i}\delta \dot{q}_i \\ &= \frac{\partial L}{\partial t}\delta t + \sum_i \dot{p}_i \delta q_i + \sum_i p_i \delta \dot{q}_i. \end{aligned} \tag{2.3}$$

Hence

$$\delta\left(L - \sum_i p_i \dot{q}_i\right) = \frac{\partial L}{\partial t}\delta t + \sum_i \dot{p}_i \delta q_i - \sum_i \dot{q}_i \delta p_i. \tag{2.4}$$

This last equation suggests that the expression in brackets on the l.h.s. is a function of (t, q, p). It is clear that every pair of a generalised velocity $\dot{q}_i$ and its canonical momentum p_i have the same status as conjugate variables in thermodynamics. (Of course, historically, people developed the underlying maths behind Lagrangian and Hamiltonian dynamics first; the physical concepts of thermodynamics were then easy to formalise.)

Defining the Hamiltonian as in Eq. (1.31):

$$H \equiv \left(\sum_i p_i \dot{q}_i - L\right)\Bigg|_{\dot{q}_i = \dot{q}_i(\mathbf{q},\mathbf{p},t)}, \tag{2.5}$$

then (2.4) implies

$$\dot{p}_i = -\left(\frac{\partial H}{\partial q_i}\right)_{\text{other } q,\mathbf{p},t} \qquad \dot{q}_i = \left(\frac{\partial H}{\partial p_i}\right)_{\text{other } p,\mathbf{q},t}, \tag{2.6}$$

which are **Hamilton's equations of motion**. Notice the importance of completing the transformation in Eq. (2.5) and writing H as a function of $\mathbf{q}, \mathbf{p}, t$ only. None of the $\dot{q}_i$

should appear explicitly; rather, they should be replaced by inverting Eq.(2.1) that defines p_i.

We have already shown that, provided L does not depend explicitly on time, T has the form (1.32) and V does not depend on velocities, then $H = T + V$, the total energy, and it is a constant of the motion. This does not imply that the right-hand sides of Hamilton's equations are zero! They are determined by the functional form of the dependence of H on the p_i and q_i.

Note also that p_i and q_i are now on an equal footing. In Hamilton's equations q_i can be anything, not necessarily a position coordinate. For example we could interchange the physical meaning of what we regard as p_i with q_i, and q_i with $-p_i$, and Hamilton's equations would still work.

2.1 Liouville's theorem

Phase space is defined as the space spanned by the canonical coordinates and conjugate momenta, e.g. (x, y, z, p_x, p_y, p_z) for a single particle (a 6D space) and $\{q_1, q_2, \cdots, q_n, p_1, p_2, \cdots, p_n\}$ for a system of n particles. This defines phase space to be a $6n$-dimensional space. A single point in phase space represents the state of the whole system, i.e. the positions and velocities of all its particles. It is called a **representative point**. If there are constraints acting on the system, the representative points are confined to some lower dimensional subspace. The representative points move with velocities $\mathbf{v}$ where:

$$\mathbf{v} = \{\dot{q}_1, \dot{q}_2, \cdots, \dot{q}_n, \dot{p}_1, \dot{p}_2, \cdots, \dot{p}_n\} = \left\{\frac{\partial H}{\partial p_1}, \frac{\partial H}{\partial p_2}, \cdots, \frac{\partial H}{\partial p_n}, -\frac{\partial H}{\partial q_1}, -\frac{\partial H}{\partial q_2}, \cdots, -\frac{\partial H}{\partial q_n}\right\}. \tag{2.7}$$

Liouville's theorem is a very powerful result concerning the evolution in time of ensembles of systems. We can regard the initial state of the ensemble of systems as corresponding to a distribution or density of representative points in phase space. Then Liouville's theorem states that:

Theorem 2 (Liouville's theorem)**.** *The density in phase space evolves as an incompressible fluid.*

Proof. The proof is a simple application of Hamilton's equations and the n-dimensional divergence theorem. The n-dimensional divergence theorem states that for an n-dimensional vector function of n variables $\mathbf{V}(x_1, \cdots, x_n)$:

$$\int_{\mathcal{V}} \sum_i \frac{\partial V_i}{\partial x_i} \, \mathrm{d}\tau = \int_{\mathcal{S}} \sum_i V_i \, \mathrm{d}S_i \,, \tag{2.8}$$

where $\mathrm{d}\tau$ is an n-dimensional volume element and $\mathrm{d}\mathbf{S}$ is an $(n-1)$-dimensional element of surface area. This theorem and its proof are the obvious generalisation the divergence theorem (the Gauss theorem) in 3D:

$$\int_{\mathcal{V}} \boldsymbol{\nabla} \cdot \mathbf{V} \, \mathrm{d}\tau = \int_{\mathcal{S}} \mathbf{V} \cdot \mathrm{d}\mathbf{S} \,. \tag{2.9}$$

To prove Liouville's theorem, suppose that the representative points are initially confined to some (n-dimensional) volume $\mathcal{V}$ with surface $\mathcal{S}$. The points move with velocity $\mathbf{v}$ given by Eq. (2.7). Therefore, at the surface, the volume occupied by the points is changing at a rate $\mathrm{d}V = \mathbf{v} \cdot \mathrm{d}\mathbf{S}$. Hence:

$$\begin{aligned} \Delta V &= \int_{\mathcal{S}} \mathbf{v} \cdot \mathrm{d}\mathbf{S} \\ &= \int_{\mathcal{V}} \boldsymbol{\nabla} \cdot \mathbf{v} \, \mathrm{d}\tau \,, \end{aligned} \tag{2.10}$$

by (2.9). However,

$$\begin{aligned} \boldsymbol{\nabla} \cdot \mathbf{v} &= \sum_i \frac{\partial}{\partial q_i} \dot{q}_i + \frac{\partial}{\partial p_i} \dot{p}_i \\ &= \sum_i \frac{\partial^2 H}{\partial q_i \partial p_i} - \frac{\partial^2 H}{\partial p_i \partial q_i} = 0. \end{aligned} \tag{2.11}$$

So the volume occupied by the ensemble's representative points does not change. □

2.2 Poisson Brackets and the Analogy with Quantum Commutators

Suppose $f = f(q_i, p_i, t)$, i.e. f is a function of the dynamical variables $\mathbf{p}$ and $\mathbf{q}$. Then

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\partial f}{\partial q_i} \dot{q}_i + \frac{\partial f}{\partial p_i} \dot{p}_i + \frac{\partial f}{\partial t} = \frac{\partial f}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial H}{\partial q_i} + \frac{\partial f}{\partial t}, \tag{2.12}$$

which we can rewrite for notational convenience as

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \{f, H\} + \frac{\partial f}{\partial t}, \tag{2.13}$$

where

$$\{f, g\} \equiv \frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i}, \tag{2.14}$$

is called the *Poisson Bracket* of the functions f and g and was first introduced into mechanics by Simon Poisson in 1809 (Note the use of summation convention in Eq. (2.14)).

Eq. (**??**) is remarkably similar to the Ehrenfest theorem in Quantum Mechanics

$$\frac{\mathrm{d}}{\mathrm{d}t} \left\langle \hat{O} \right\rangle = \frac{1}{i\hbar} \left\langle \left[\hat{O}, \hat{H} \right] \right\rangle + \left\langle \frac{\partial \hat{O}}{\partial t} \right\rangle, \tag{2.15}$$

for the variation of expectation values of (Hermitian) operators $\hat{O}$, where $\hat{H}$ is of course the quantum mechanical Hamiltonian operator and $\left[\hat{O}, \hat{H}\right]$ represents the commutator $\hat{O}\hat{H} - \hat{H}\hat{O}$.

It is easy to check that $\{f, g\}$ has many of the properties of the commutator:

$$\{f, g\} = -\{g, f\}, \quad \{f, f\} = 0 \quad \text{etc.} \tag{2.16}$$

Also, if in Eq. (2.12) $\partial f/\partial t = 0$ (no explicit t dependence) and $\{f, H\} = 0$ then $\mathrm{d}f/\mathrm{d}t = 0$, i.e. f is a constant of the motion.

This suggests we can relate classical and quantum mechanics by formulating classical mechanics in terms of Poisson Brackets and then associating these with the corresponding Quantum Mechanical commutator

$$\begin{gathered} \{A, B\} \leftrightarrow \frac{1}{i\hbar}\left[\hat{A}, \hat{B}\right] \\ \text{classical} \leftrightarrow \text{quantum} \\ H \leftrightarrow \hat{H}. \end{gathered} \tag{2.17}$$

Classically

$$\{q, p\} = \frac{\partial q}{\partial p}\frac{\partial p}{\partial p} - \frac{\partial q}{\partial p}\frac{\partial p}{\partial q} = 1. \tag{2.18}$$

Quantum Mechanically,

$$\left[q, -i\hbar\frac{\partial}{\partial q}\right]\Psi = -qi\hbar\frac{\partial}{\partial q}\Psi + i\hbar\frac{\partial}{\partial q}(q\Psi) = i\hbar\Psi, \tag{2.19}$$

i.e.

$$\frac{1}{i\hbar}\left[q, -i\hbar\frac{\partial}{\partial q}\right] = 1 = \{q, p\}. \tag{2.20}$$

Therefore we confirm the identification of $-i\hbar\,\partial/\partial q$ with the canonical **momentum operator** $\hat{p}$ corresponding to q.

2.2.1 Quantum Variational Principle

In QM we associate a wave vector $\mathbf{k} = \mathbf{p}/\hbar$ with a particle of momentum p (de Broglie relation), and frequency $\omega = H/\hbar$ with its total energy $E = H$ (Einstein relation). Thus we can write Hamilton's principle for the classical motion of the particle as

Appendix A

Appendix