Chapter 16
Least action principle in mechanics

December 3, 2019

The least action principle is introduced in mechanics, in preparation for application to electrodynamics. The principle leads to an elegant derivation of the momentum and energy of a relativistic free particle.

Contents

Note
A good reference is the *Feynman Lectures* [1]. Students may also wish to view a video and read more extensive notes on this topic [2].

1 General principles

1.1 Different formulations of dynamics

There are several ways of writing down the dynamics of a system. Consider for example a charged particle in an EM field.

(1) Component by component
One can write the equation of motion component by component, e.g.,

$$\frac{dp_x}{dt} = F_x = q(E_x + v_y B_z - v_z B_y) \qquad (1)$$

(2) In vector form
Any physics undergraduate would know how to write the above (3 equations) as a single vector equation:

$$\frac{d\mathbf{p}}{dt} = \mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \qquad (2)$$

This exhibits covariance under rotations, but does not (readily) exhibit covariance under Lorentz transformations.

(3) Covariant form
The following is even better:

$$\frac{dp^\mu}{d\tau} = K^\mu = qF^{\mu\nu}u_\nu \qquad (3)$$

This is explicitly covariant under Lorentz transformations. However, it may still look slightly "unnatural" or at least "unexpected".

(4) Action formalism
Now we come to the fourth way of doing dynamics — the least action principle (LAP), which is schematically

$$\boxed{\delta\mathcal{S} = 0} \qquad (4)$$

This formulation (a) is explicitly covariant and therefore well suited to relativity; (b) is particularly simple, so that the Lorentz force law emerges

"naturally"; and (c) ties together the Lorentz force law and Maxwell's equations by revealing that they come from the same interaction term.

This Chapter starts by reviewing LAP in mechanics. The next Chapter extends these ideas to EM.

1.2 Newtonian mechanics

Forget about EM for the moment, and go back to Newtonian mechanics in 1D. The variable is $x(t)$. Suppose it is given that

$$x(t_1) = x_1 \quad , \quad x(t_2) = x_2 \tag{5}$$

What is $x(t)$ in between? Graphically (**Figure 1**), this means determining the correct path (solid line) among all possible paths (say the broken line) in a t–x diagram. To conform with the convention in spacetime diagrams, the t-axis is drawn vertically, even though it is usually the independent variable. In the equation of motion approach, the correct path is the one which satisfies the differential equation of motion, such as (2), or a Newtonian equation, e.g.,

$$m\frac{d^2x(t)}{dt^2} \quad = \quad F(x) = -\frac{\partial V(x)}{\partial x} \tag{6}$$

together with the conditions (5). The system is specified by $F(x)$ or $V(x)$, e.g., for a spring, $F = -kx$, or $V = kx^2/2$.

The LAP takes a completely different point of view.

1.3 Least action principle

Prescription

The main idea of LAP is as follows.

- Consider any *path* P between (t_1, x_1) and (t_2, x_2). It need not be the correct path. Each path P corresponds to a function $[x(t)]$.
- Give a way of calculating a number $\mathcal{S}[P] = \mathcal{S}[x(t)]$ for each path $P = [x(t)]$; $\mathcal{S}$ is called the *action*.
- $\mathcal{S}$ is made up of contributions $\Delta\mathcal{S}$ for each segment of path (additivity assumption).
- Consider all possible paths, and select the one whose action is minimum. This is the correct path.

Advantages

A path P is independent of the coordinates used to describe it. **Figure 2** shows the same path P in three different coordinate systems (whether differing by a rotation or a Lorentz transformation).[1]

[1] It is conventional to represent the latter by the t and x axes turning in opposite directions.

If the action $\mathcal{S}(P)$ does not depend on the coordinate system, then LAP will select the same path in any coordinates. It guarantees that the physics is invariant. Thus, LAP is specially convenient for relativity, where invariance is a central issue, and also in Newtonian physics when using generalized coordinates — otherwise it is difficult to verify that the equations in different generalized coordinates give the same physics.

We shall continue with Newtonian physics, and discuss (a) how we can guess the form of $\mathcal{S}$; and (b) how the equation of motion is obtained from $\mathcal{S}$.

2 Non-relativistic physics

2.1 Choice of action

Consider a single particle of mass m, moving in 1D under a potential $V(x)$. Generalization to higher dimensions and/or more variables is straightforward.

Because of the additivity assumption, we need only consider a small segment of path (**Figure 3**), centered at (t, x) and of length $(\Delta t, \Delta x)$. The action $\Delta\mathcal{S}$ must be proportional to the size of the interval, which we represent through Δt:

$$\boxed{\Delta\mathcal{S} = L\,\Delta t} \tag{7}$$

and L is called the *Lagrangian*. By the way, going from $\mathcal{S}$ to L is the conventional approach when these ideas are first encountered, but (as we shall see) is not the optimal approach in relativity — because in defining L, we have singled out one particular direction in spacetime and therefore ruined explicit invariance.

The Lagrangian L can only depend on t, x and $\Delta x/\Delta t = \dot{x}$

$$L \quad = \quad L(x, \dot{x}, t) \tag{8}$$

If we assume the system has no explicit time dependence, then t does not appear.

Next assume the system knows about the position x only though the potential $V(x)$:

$$L \quad = \quad L(V(x), \dot{x})$$

and moreover assume that L is linear in $V(x)$.

$$L \quad = \quad A(\dot{x}) + B(\dot{x})V(x)$$

where A, B are unknown functions.

Then expand A, B. The lower-order expansion should be adequate for small velocities

$$\begin{aligned} A(\dot{x}) &= a_0 + a_1\dot{x} + a_2\dot{x}^2 + \ldots \\ B(\dot{x}) &= b_0 + b_1\dot{x} + b_2\dot{x}^2 + \ldots \end{aligned}$$

The term a_0 would contribute

$$\mathcal{S} \quad = \quad \int_{t_1}^{t_2} a_0 \, dt = a_0(t_2 - t_1)$$

which is the same for all paths satisfying the boundary conditions. It is irrelevant for picking the minimum, and we set it to zero.

The term $a_1 \dot{x}$ would change sign under $t \mapsto -t$. It is not allowed if the physics is time-reversal invariant; so $a_1 = 0$.

The lowest-order guess is therefore

$$L \quad = \quad a_2 \dot{x}^2 + b_0 V(x)$$

We shall show that the choice

$$a_2 = \frac{1}{2} m \quad , \quad b_0 = -1$$

gives the Newtonian equation of motion. (Actually, only the ratio b_0/a_2 matters, since multiplying S by a constant has no effect.)

2.2 Euler-Lagrange equation

We now start with

$$\boxed{L = \frac{1}{2} m \dot{x}^2 - V(x)} \tag{9}$$

and derive the Newtonian equation of motion. Let $x(t)$ be the correct path, and let $x(t) + \eta(t)$ be a neighboring path; η is considered a small quantity (**Figure 1**). Moreover, since the neighboring paths must satisfy the same initial and final conditions,

$$\eta(t_1) = \eta(t_2) \quad = \quad 0 \tag{10}$$

Now compare the two actions[2]

$$\begin{aligned}
& \delta\mathcal{S} \equiv \mathcal{S}\left[x(t) + \eta(t)\right] - \mathcal{S}\left[x(t)\right] \\
= & \int_{t_1}^{t_2} dt \left\{ \frac{1}{2} m \left[\dot{x}(t) + \dot{\eta}(t)\right]^2 - V(x+\eta) \right\} \\
& - \int_{t_1}^{t_2} dt \left\{ \frac{1}{2} m \left[\dot{x}(t)\right]^2 - V(x) \right\} \\
= & \int_{t_1}^{t_2} dt \left[m\dot{x}(t)\dot{\eta}(t) - \frac{\partial V(x)}{\partial x} \eta \right] \\
= & \int_{t_1}^{t_2} dt \left\{ \frac{d}{dt}[m\dot{x}(t)\eta(t)] \right. \\
& \left. - m\ddot{x}(t)\eta(t) - \frac{\partial V(x)}{\partial x}\eta(t) \right\}
\end{aligned}$$

[2]We use $\delta\mathcal{S}$ to denote the variation of $\mathcal{S}$ between two paths, whereas $\Delta\mathcal{S}$ denotes a small contribution to $\mathcal{S}$ due to a segment of any one path.

Only first-order terms in η have been kept. The first term can be integrated exactly.

$$\int_{t_1}^{t_2} dt \frac{d}{dt} \left[m\dot{x}(t)\eta(t) \right] \quad = \quad m\dot{x}(t)\eta(t) \Big|_{t_1}^{t_2} = 0$$

because of (10). Thus we are left with

$$\delta\mathcal{S} \quad = \quad - \int_{t_1}^{t_2} dt \left(m\ddot{x} + \frac{\partial V}{\partial x} \right) \eta(t) \tag{11}$$

If the original $x(t)$ gives a minimum, then δS must vanish for all first-order changes, i.e., for all $\eta(t)$. This means the bracket above must be zero for every t:

$$m\ddot{x} + \frac{\partial V}{\partial x} = 0 \tag{12}$$

which is the Newtonian equation of motion.

For the benefit of future use, note that the crucial step in getting rid of $\dot{\eta}$ is an integration by parts, which can be denoted schematically as

$$A(t)\,\dot{\eta}(t) \quad \mapsto \quad -\dot{A}(t)\,\eta(t)$$

when this is placed under an integral over t, and subject to boundary conditions that ensure that the integrated term vanishes. Later we shall encounter a generalization: functions of a spacetime variable x can be handled similarly:

$$A(x)\left[\partial^\mu B(x)\right] \quad \mapsto \quad -\left[\partial^\mu A(x)\right] B(x)$$

when placed under a 4-fold integral d^4x.

2.3 More general form of action

It is important to realize that L is not defined by KE − PE; it is just whatever expression that would give the correct equation of motion, and is KE − PE only in Newtonian mechanics. We shall later come across more general forms of action, at two levels:

- still staying with Newtonian mechanics, with L given by KE − PE, but expressed in coordinates that are not necessarily cartesian; and
- going beyond Newtonian mechanics, so that L is not KE − PE.

There are two important examples of the latter, both of which we shall study:

- a relativistic free particle (see Section 3), and
- motion of a charge in a magnetic field (whose effect is certainly not described by any PE, since the magnetic field does no work).

So let us derive the equation of motion for a general L. To save some writing, and to unify notation, $\eta(t)$ is written as $\delta x(t)$.

Start with

$$\begin{aligned}
\mathcal{S} &= \int dt\, L(x,\dot{x},t) \\
\delta\mathcal{S} &= \int dt\, \delta L \\
&= \int dt\, \left[\frac{\partial L}{\partial x}\delta x + \frac{\partial L}{\partial \dot{x}}\delta\dot{x}\right] \\
&= \int dt\, \left\{\left[\frac{\partial L}{\partial x}\delta x - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right)\delta x\right]\right. \\
&\qquad \left. + \frac{d}{dt}\left[\left(\frac{\partial L}{\partial \dot{x}}\right)\delta x\right]\right\} \\
&= \int dt\, \left[\frac{\partial L}{\partial x} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right)\right]\delta x
\end{aligned}$$

Hence

$$\boxed{\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) = \frac{\partial L}{\partial x}} \tag{13}$$

which is called the *Euler–Lagrange* equation of motion. It is readily generalized to several variables x_i.

Nowhere did we say that the variables are cartesian coordinates. To emphasize this point, the coordinates in Lagrangian mechanics are often denoted as q_i.

Conjugate momentum
We call $\partial L/\partial\dot{x}$ the *conjugate momentum* p.

$$\boxed{p = \frac{\partial L}{\partial \dot{x}}} \tag{14}$$

The conjugate momentum p may or may not be the same as the "ordinary" or mechanical momentum mv.

The Euler–Lagrange equation can be written as

$$\boxed{\frac{dp}{dt} = \frac{\partial L}{\partial x}} \tag{15}$$

2.4 Hamiltonian

Definition of Hamiltonian
Define a function of three variables[3]

$$H(x,\dot{x},p) = \dot{x}p - L(x,\dot{x}) \tag{16}$$

This definition is made at each time, so x and $\dot{x}$ should be considered as independent variables. If it helps your understanding, you may like to replace $\dot{x} \mapsto v$ below. An explicit t-dependence in L and H would not change these formulas.

[3]More generally, we may suppose that H also depends explicitly on time t.

First, we note that, regarded as a function of three independent variables $x, \dot{x}, p$, H is actually independent of $\dot{x}$, because

$$\frac{\partial H}{\partial \dot{x}} = p - \frac{\partial L}{\partial \dot{x}} = 0$$

the last step by definition of p in (14). Therefore we can write

$$H = H(x,p) \tag{17}$$

Hamilton equations of motion
Differentiate (16) with respect to p, giving

$$\frac{\partial H}{\partial p} = \dot{x}$$

Secondly, the Euler–Lagrange equation in the form (15) tells us what $\dot{p}$ is:

$$\frac{dp}{dt} = \frac{\partial L}{\partial x} = -\frac{\partial H}{\partial x}$$

Thus, the evolution of the system can be expressed as a pair of first-order equations for the time-evolution of the coordinates (x,p):

$$\boxed{\frac{dx}{dt} = \frac{\partial H}{\partial p} \quad , \quad \frac{dp}{dt} = -\frac{\partial H}{\partial x}} \tag{18}$$

Problem 1
Take the Newtonian case (9) and show that

$$\begin{aligned}
p &= m\dot{x} \\
H &= \frac{p^2}{2m} + V(x)
\end{aligned}$$

However, these are not valid more generally — in particular in the presence of magnetic fields, as we shall see below. §

Conservation law
Suppose H has no explicit dependence on time

$$\frac{\partial H}{\partial t} = 0$$

then all changes in H occur through its variables $x = x(t)$ and $p = p(t)$. In particular, by the chain rule,

$$\begin{aligned}
\frac{dH}{dt} &= \frac{\partial H}{\partial x}\frac{dx}{dt} + \frac{\partial H}{\partial p}\frac{dp}{dt} \\
&= \left(\frac{\partial H}{\partial x}\right)\left(\frac{\partial H}{\partial p}\right) + \frac{\partial H}{\partial p}\left(-\frac{\partial H}{\partial x}\right) \\
&= 0
\end{aligned} \tag{19}$$

showing that H is conserved. Its value is evidently the energy.[4]

Note that H is more than its numerical value; its functional form specifies how the system evolves in time.

2.5 Generalizations

Generalized coordinates
All the above can be generalized to many coordinates (even an infinite number), and to coordinates that are not cartesian. As an example, consider a particle in a plane, described by polar coordinates (r, ϕ) and subject to a radial potential $V = V(r)$. Generalizing the expression in (9),

$$\begin{aligned} L &= \text{KE} - \text{PE} \\ &= \frac{m}{2}\left(\dot{r}^2 + r^2\dot{\phi}^2\right) - V(r) \end{aligned} \qquad (20)$$

The coordinates are $(q_1, q_2) = (r, \phi)$.

Problem 2
Write out the Euler–Lagrange equations for r and ϕ, and check that the centrifugal pseudo-force is reproduced correctly. §

Problem 3
Show, in general, that if a coordinate q_i does not appear in L:

$$\frac{\partial L}{\partial q_i} = 0$$

then the corresponding conjugate momentum is conserved:

$$\frac{dp_i}{dt} = 0$$

This is known as Noether's theorem. What is the conjugate momentum corresponding to ϕ, which is conserved if V does not depend on ϕ? §

Statistical mechanics
Statistical mechanics is usually formulated in Hamiltonian phase space, i.e., the space described by

$$(q_1, \ldots, q_n; p_1, \ldots p_n)$$

with probabilities calculated using the measure

$$d\mu = \prod_j dq_j dp_j$$

The point is that the "volume" element is invariant with respect to a change of variables — a key property of Hamiltonian mechanics that we do not prove here.

Quantum mechanics
There are two routes to go from classical mechanics to quantum mechanics (QM). The familiar route would be via the Hamiltonian — find the classical Hamiltonian H, and convert the conjugate momenta to operators:

$$\boxed{p_i \mapsto -i\hbar\frac{\partial}{\partial q_i}}$$

We shall briefly touch upon the QM of charged particles using this method.

The second route is actually even better conceptually (though computationally not as neat for non-relativistic QM). The method is to sum

$$\exp\{i\mathcal{S}[q(t)]/\hbar\}$$

over all possible paths linking the initial point to the final point, and this would give the propagator. This method of *path integrals* is better because it does not single out the time in spacetime — which the Hamiltonian does.

3 Relativistic free particle

We now consider relativity, but starting with a free particle.

3.1 Choice of action

Consider a single particle of mass m, moving in 3D. So think of **Figure 1** to **Figure 3** as spacetime diagrams, and $x \mapsto \mathbf{x}$. A point is now denoted as x^μ and the segment in **Figure 3** is $\Delta\vec{x} = (c\Delta t, \Delta\mathbf{x})$. Since the segment represents a possible path, it must be time-like, i.e., $c|\Delta t| > |\Delta\mathbf{x}|$, $\Delta x^\mu \Delta x_\mu < 0$.

We now repeat the arguments in Section 2.1, but with two differences. (a) Because of Lorentz invariance, the choice of $\mathcal{S}$ is much more limited. (b) It is no longer so natural to use L. The reason is that $\Delta\mathcal{S}$ is invariant, but Δt is not. So $L = \Delta\mathcal{S}/\Delta t$ is not invariant. Conceptually, it is better to deal directly with the invariant quantity $\Delta\mathcal{S}$.

First, because of the additivity property, we can focus on a small segment Δx^μ (**Figure 3**). There is no dependence on t, and if the particle is free, every position is equivalent, so there is no dependence on $\mathbf{x}$. Hence $\Delta\mathcal{S}$ depends only on Δx^μ.

By Lorentz invariance, $\Delta\mathcal{S}$ can only depend on the product $\Delta x^\mu \Delta x_\mu$. However, additivity implies linearity: if we double a small interval, $\Delta\mathcal{S}$ must be doubled. Therefore,

$$\Delta\mathcal{S} \propto \sqrt{-\Delta x^\mu \Delta x_\mu}$$

[4] Some dissipative systems can be described by a Hamiltonian, but in those cases, the Hamiltonian H may not be the same as the energy. I am grateful to D Kiang for drawing my attention to this point.

The minus sign must be inserted because $\Delta x^\mu \Delta x_\mu < 0$. The proportionality constant does not matter. For a free particle, this is the only term, and multiplying $\mathcal{S}$ by an overall constant does not change which path gives the least action. So this constant is a matter of convention, and we take

$$\Delta\mathcal{S} \;=\; -mc\sqrt{-\Delta x^\mu \Delta x_\mu}$$

Adding these up and passing to the continuous form,

$$\boxed{\mathcal{S} = -mc\int \sqrt{-dx^\mu dx_\mu}} \qquad (21)$$

We shall next discuss how to deal with such an object, in which the differential appears in such a "funny" way.

3.2 Equation of motion: non-covariant approach

Lagrangian
Divide and multiply by dt:

$$\mathcal{S} \;=\; -mc\int \sqrt{-\frac{dx^\mu}{dt}\frac{dx_\mu}{dt}}\; dt \qquad (22)$$

which allows the Lagrangian to be identified:

$$L \;=\; -mc\sqrt{-\frac{dx^\mu}{dt}\frac{dx_\mu}{dt}} \qquad (23)$$

More explicitly,

$$\frac{dx^\mu}{dt} \;=\; \left(\frac{dx^0}{dt}, \frac{d\mathbf{x}}{dt}\right) = (c, \mathbf{v})$$

where $\mathbf{v} = \dot{\mathbf{x}}$. Thus

$$\frac{dx^\mu}{dt}\frac{dx_\mu}{dt} \;=\; c^2(-1 + \mathbf{v}^2/c^2)$$

and

$$\boxed{L = -mc^2\sqrt{1 - \mathbf{v}^2/c^2}} \qquad (24)$$

We know that $\mathcal{S}$ is invariant, because of (21). But now we have written it in a way which is not *explicitly* covariant.

Note that the square root is in the numerator, and this is *not* the kinetic energy.

Problem 4
Expand (24) for small velocities, and show that except for an additive constant, $L = (1/2)mv^2$, proving that the choice of multiplicative constant in (21) is consistent with the non-relativistic Lagrangian. §

Momentum
Next apply the standard formulas to find the momentum:

$$\begin{aligned} p_i &= \frac{\partial L}{\partial v_i} \\ &= -mc^2(\frac{1}{2})(1 - \mathbf{v}^2/c^2)^{-1/2}(-2v_i)/c^2 \\ &= \frac{mv_i}{\sqrt{1 - \mathbf{v}^2/c^2}} = m\gamma v_i \end{aligned}$$

So the conjugate momentum p is exactly the same as the mechanical momentum. From the Euler–Lagrange equation, we then get

$$\frac{dp_i}{dt} \;=\; 0$$

as is natural for a free particle. We see that the relativistic momentum emerges very naturally.

Problem 5
Use the definition of H to obtain an expression for the energy of a free particle. Express the answer in terms of p, without v. §

3.3 Equation of motion: covariant approach

In the non-covariant approach, we use t as the independent variable, and express $\mathbf{x} = \mathbf{x}(t), \mathbf{v} = \mathbf{v}(t)$ etc. However, the four components of x^μ have the same status, so such a different treatment of one component is not elegant. A better approach is to use an arbitrary *path parameter* s as the independent variable, and let $x^\mu = x^\mu(s)$. For example we can have

$$\begin{aligned} t &= s + 0.3s^2 \;\;,\;\; 0 < s < 1 \\ x &= \sin \pi s \\ y &= \cos \pi s \\ z &= 0 \end{aligned}$$

The path parameter has no physical meaning. For example, we can let $s = s'^2$:

$$\begin{aligned} t &= (s')^2 + 0.3(s')^4 \;\;,\;\; 0 < s' < 1 \\ x &= \sin \pi (s')^2 \\ y &= \cos \pi (s')^2 \\ z &= 0 \end{aligned}$$

It describes exactly the same path, but simply with different labelling. In using a path parameter, there are two types of invariance that needs to be preserved: (a) Lorentz invariance (plus any other invariance of the dynamics); and (b) invariance under re-labelling such as $s \mapsto s'$. For simplicity, we shall not deal with the latter.

With a path-parameter s, we then write (22) as

$$\mathcal{S} \;=\; -mc\int_{s_1}^{s_2}\sqrt{-\frac{dx^\mu}{ds}\frac{dx_\mu}{ds}}\,ds$$

We assume all paths are labelled on the interval $[s_1, s_2]$. Because of the freedom to relabel, this does not impose any real restrictions. The boundary conditions are

$$\begin{aligned} x^\mu(s_1) &= x_1^\mu = (ct_1, \mathbf{x}_1) \\ x^\mu(s_2) &= x_2^\mu = (ct_2, \mathbf{x}_2) \end{aligned}$$

Now consider a change

$$x^\mu(s) \;\mapsto\; x^\mu(s) + \delta x^\mu(s)$$

under which

$$\begin{aligned} & \delta\left(-\frac{dx^\mu}{ds}\frac{dx_\mu}{ds}\right)^{1/2} \\ =\;& \frac{1}{2}\left(\ldots\right)^{-1/2}\left(-2\frac{dx^\mu}{ds}\cdot\frac{d}{ds}\delta x_\mu\right) \\ =\;& -\left(\ldots\right)^{-1/2}\frac{dx^\mu}{ds}\cdot\frac{d}{ds}\delta x_\mu \\ =\;& \frac{d}{ds}\left[\left(\ldots\right)^{-1/2}\frac{dx^\mu}{ds}\right]\delta x_\mu \end{aligned}$$

where in the last step we have imagined placing the expression under an integral and integrated by parts. Thus

$$\delta\mathcal{S} \;=\; -mc\int\frac{d}{ds}\left[\left(\ldots\right)^{-1/2}\frac{dx^\mu}{ds}\right]\delta x_\mu\,ds$$

and the equation of motion is

$$\boxed{\frac{d}{ds}\left[\left(\ldots\right)^{-1/2}m\frac{dx^\mu}{ds}\right]=0} \qquad (25)$$

valid for *any* path parameter s.

The equation of motion becomes simpler if we choose s to be essentially the proper time interval:

$$(\Delta s)^2 \;=\; (\Delta\tau)^2 = -c^{-2}\Delta x^\nu\Delta x_\nu$$

In other words, label each point on the path by the proper time τ elapsed along that path. Then

$$\begin{aligned} -\frac{dx^\nu}{ds}\frac{dx_\nu}{ds} &= c^2 \\ \frac{d}{ds} &\mapsto \frac{d}{d\tau} \end{aligned}$$

and (25) simplifies to

$$\boxed{\frac{d}{d\tau}\left(m\frac{dx^\mu}{d\tau}\right)=0} \qquad (26)$$

in which τ is the proper time.

A technical note
We put s (a general path parameter) $\mapsto \tau$ (proper time) only *after* the variation. Can we set $s \mapsto \tau$ from the beginning? The answer is *no*, since we have assumed that the range of s is the same for all paths, e.g., the different paths in **Figure 1**. This is possible if they are labelled by s — since s is general, we can always re-scale the parameter for one of the paths. But this is not possible if all paths are labelled by the proper time to start with.

4 Summary

To summarize, LAP is an alternate — and in many ways better — way to deal with mechanics. The Lagrangian and Hamiltonian treatments emerge naturally.

For a relativistic particle, the action is determined by the requirement of invariance. For future reference, denote the action of matter as $\mathcal{S}_M$, then schematically

$$\begin{aligned} \mathcal{S}_M &= -mc^2\int d\tau \\ \delta\mathcal{S}_M &= -\int\frac{d}{d\tau}\left(m\frac{dx^\mu}{d\tau}\right)\delta x_\mu\,d\tau \qquad (27) \end{aligned}$$

understood to be summed over all the particles. The above is schematic and a shorthand in that in $\mathcal{S}_M$, i.e., before doing the variation, we are not really allowed to set the path parameter to be τ.

A Supplement: Lagrangian, Hamiltonian and the phase of a wave

Each supplement will consist of one or more problems, usually more challenging and of an open nature, to stretch the better students.

Students should be familiar with both of the following properties (see boxed equations), which for the present purpose will be discussed only for one degree of freedom, though generalization is straightforward.

Lagrangian and Hamiltonian
The Lagrangian and the Hamiltonian are related by

$$\boxed{L = pv - H} \qquad (28)$$

where $v = \dot{x}$.

Phase of a wave
The phase of a plane wave is, in obvious notation

$$\Phi \;=\; kx - \omega t \tag{29}$$

For situations where the wave number and the frequency may change (i.e., the wave is not "plane"), then the formula at least holds for propagation in a small segment of spacetime, and the incremental phase is

$$\boxed{\Delta\Phi = k\,\Delta x - \omega\,\Delta\,t} \tag{30}$$

Relationship
The insight stressed here is that the two boxed equations are deeply related. Prove this. Hints:

- The phase of the wavefunction is $\Phi = S/\hbar$.
- For a small segment of path $\Delta S = L\,\Delta t$.
- The Hamiltonian is just the energy: $H = E$.
- Wave-particle duality implies $E = \hbar\omega$, $p = \hbar k$.

B Supplement: Hamiltonian phase space

Each supplement will consist of one or more problems, usually more challenging and of an open nature, to stretch the better students.

Statistical mechanics is formulated using the phase-space measure

$$d\mu \;=\; \prod_j dq_j\,dp_j \tag{31}$$

There is more than meets the eye in this simple expression.

Example
To appreciate the power and elegance, consider polar coordinates in 2D, i.e.,

$$\begin{aligned}(q_1, q_2) &= (r, \phi)\\ (p_1, p_2) &= (p_r, p_\phi)\end{aligned} \tag{32}$$

where the conjugate momenta are defined as

$$p_j \;=\; \frac{\partial L}{\partial \dot{q}_j} \tag{33}$$

It is important to realize that the conjugate momenta are not what undergraduates would normally describe as the radial and tangential components of momentum. The latter, denoted as $\tilde{p}_r, \tilde{p}_\phi$ are defined as follows:

$$\begin{aligned}\alpha &= \arctan\left(p_y/p_x\right)\\ \tilde{p}_r &= p_x\,\cos\alpha + p_y\,\sin\alpha\\ \tilde{p}_\phi &= -p_x\,\sin\alpha + p_y\cos\alpha\end{aligned} \tag{34}$$

In other words, these are just p_x, p_y rotated by an angle α. Find the relationship between p_r, p_ϕ and $\tilde{p}_r, \tilde{p}_\phi$.

In the measure (31),

$$\begin{aligned}\prod_j dq_j &= dr\,d\phi\\ &\neq r\,dr\,d\phi = dx\,dy\\ \prod_j dp_j &= dp_r\,dp_\phi\\ &\neq d\tilde{p}_r\,d\tilde{p}_\phi = dp_x\,dp_y\end{aligned} \tag{35}$$

Yet the product of the two is the same as the corresponding expression in cartesian coordinates:

$$(dr\,d\phi)\,(dp_r\,dp_\phi) \;=\; (dx\,dy)\,(dp_x\,dp_y) \tag{36}$$

Check this condition explicitly in the present example.

General theorem
The example above illustrates the general theorem that $d\mu$ is invariant under coordinate transformations. To see this, change from q_j to Q_j, where[5]

$$Q_j \;=\; Q_j(q_1, \ldots, q_N)$$

Then

$$\prod_j dQ_j \;=\; (\det\,A)\prod_i dq_i \tag{37}$$

where A is the Jacobian of the transformation:

$$A_{ji} \;=\; \frac{\partial Q_j}{\partial q_i} \tag{38}$$

But the velocities are related by

$$\dot{q}_i \;=\; \frac{\partial q_i}{\partial Q_j}\,\dot{Q}_j = B_{ij}\,\dot{Q}_j \tag{39}$$

where $B = A^{-1}$. Then

$$P_i \;=\; \frac{\partial L}{\partial \dot{Q}_i} = \frac{\partial L}{\partial \dot{q}_j}\,\frac{\partial \dot{q}_j}{\partial \dot{Q}_i} = p_j B_{ji} \tag{40}$$

Hence

$$\prod_j dP_j \;=\; (\det\,B)\prod_i dp_i \tag{41}$$

so the momentum factors are scaled by the opposite factor and the measure in phase space is invariant.

Time evolution
The phase-space volume has another important

[5] We do not consider more general transformations that mix coordinates and momenta.

property: it is invariant with respect to time evolution. For simplicity, consider one degree of freedom; generalization is straightforward. If a collection of particles occupy a volume $dq\,dp$, that volume stays the same with time, even though the particles move to a different part of phase space.

To see this, consider an infinitesimal time t, then a point q moves to

$$q' \;=\; q + \dot{q}t = q + \frac{\partial H}{\partial p}\,t$$

so an interval dq goes to

$$\begin{aligned} dq' &= dq + \frac{\partial^2 H}{\partial p\,\partial q}\,dq\,t \\ &= dq\;(1 + H_{pq}\,t) \end{aligned} \tag{42}$$

where we adopt the shorthand

$$H_{pq} \;=\; \frac{\partial}{\partial p}\frac{\partial}{\partial q}\,H$$

Likewise, for the momentum

$$p' \;=\; p + \dot{p}t = p - \frac{\partial H}{\partial q}\,t$$

so an interval dp goes to

$$\begin{aligned} dp' &= dp - \frac{\partial^2 H}{\partial q\,\partial p}\,dp\,t \\ &= dp\;(1 - H_{qp}\,t) \end{aligned} \tag{43}$$

Thus, to first order in t, we get

$$dq'\,dp' \;=\; dq\,dp \tag{44}$$

This property is important for the following reason. If we postulate a distribution such as

$$d\mu\,\exp(-\beta H) \tag{45}$$

at one time, and let the particles evolve according to Hamiltonian dynamics, then the distribution remains the same in time — in other words, it is a *stationary distribution*.

References

[1] RP Feynman, R Leighton and M Sands, *The Feynman Lectures on Physics*, the New Millennium Edition, California Institute of Technology (2013).
`www.feynmanlectures.info/`
Read especially Vol. 2, Chapter 19, "The principle of least action".

[2] K. Young, *Lecture on Least Action Principle* (2014). Both the video and a pdf file are available from
`http://www.hk-phy.org/talk/pla/`

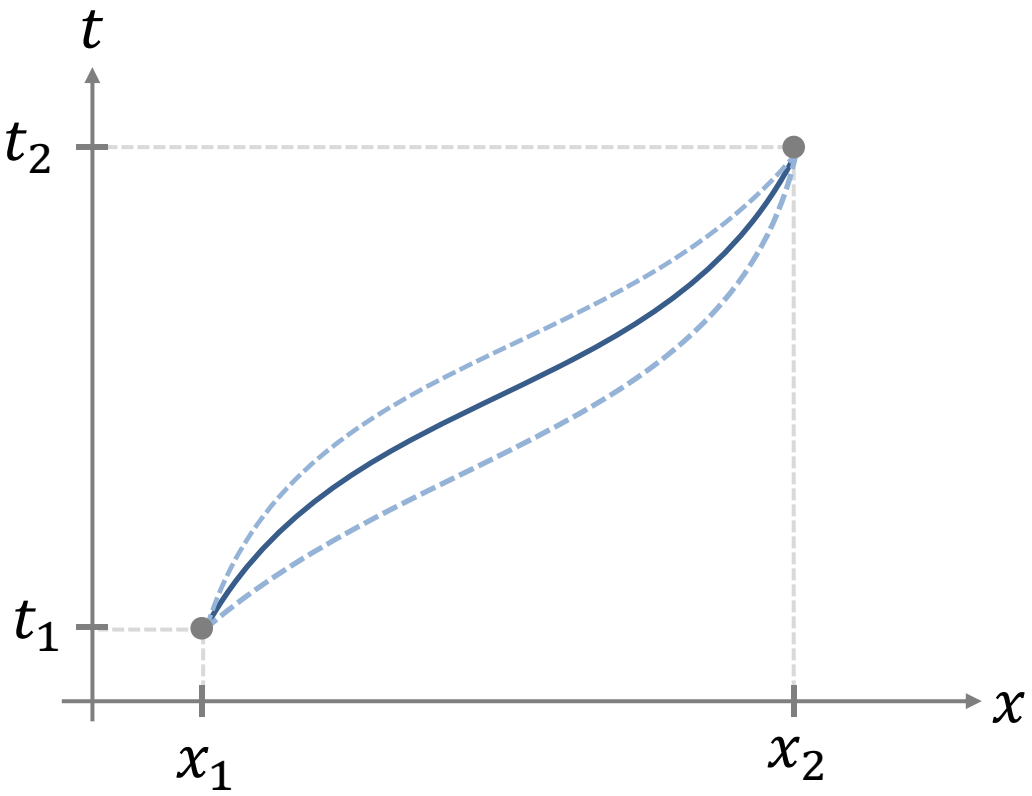
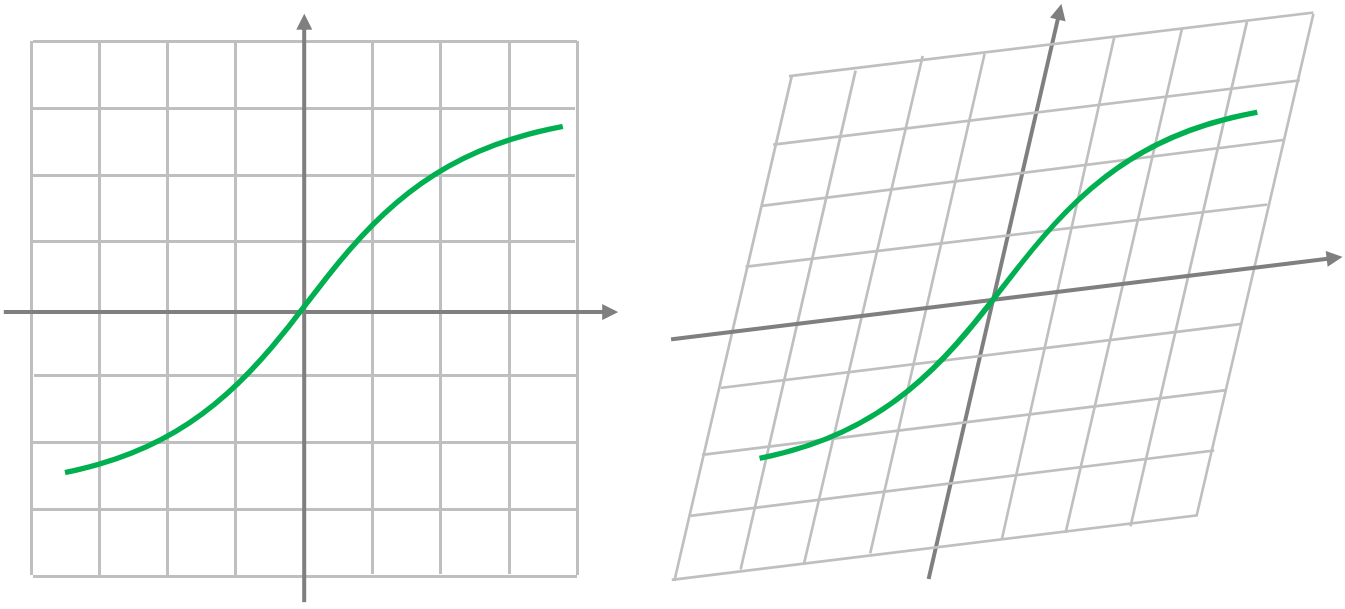
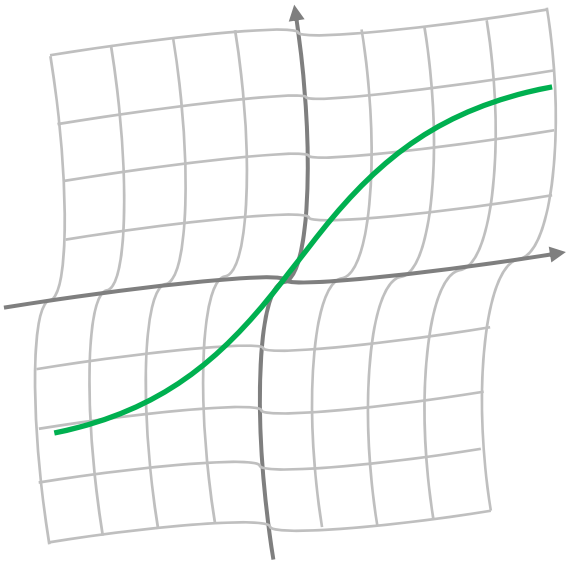


Figure 1 Which is the correct path?



(a) (b)



(c)

Figure 2 The same path in spacetime described by different coordinates.

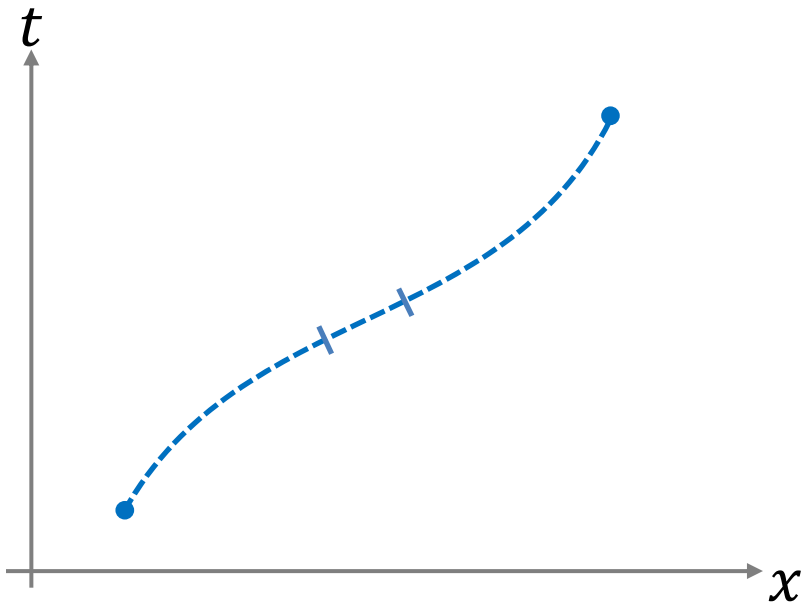


Figure 3 $\Delta\mathcal{S} = L\,\Delta t$ for a small path