Physics 106a 3 October, 2013

Lecture 2: Variational Approach

Hamilton's Principle

I introduce the *variational approach to classical mechanics*. This is formulated in terms of the *Lagrangian* $L(\{q_k\}, \{\dot{q}_k\}, t)$ which is a function of the particle coordinates $\{q_k\}$, velocities $\{\dot{q}_k\}$, and time t (k runs over the number of coordinates that need to be considered). I'm using the symbol q for the coordinate because often we will not use Cartesian coordinates. To emphasize this, we often talk about these as *generalized coordinates*.

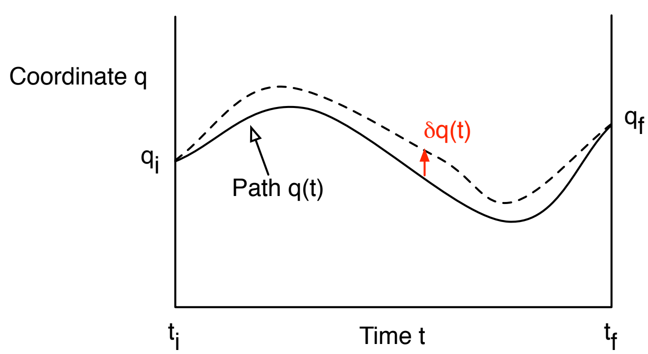
The starting point is *Hamilton's principle*:

The dynamics of the system $\{q_k(t)\}$ from time t_i to time t_f is such that the *action*

$$S = \int_{t_i}^{t_f} L(\{q_k\}, \{\dot{q}_k\}, t)\, dt \tag{1}$$

is stationary over all trial paths with fixed endpoints $\{q_k(t_i)\}$, $\{q_k(t_f)\}$.

Stationary (minimum, maximum, or saddle) means that for an infinitesimal change in path $\delta q_k(t)$ we have $\delta S = 0 + O(\delta q_k^2)$. The following figure for dynamics with a single coordinate may help to explain the procedure.



Calculus of variations

The *calculus of variations* shows us how to find the stationary value of a *functional* such as the action $S[\{q_k(t)\}]$ with respect to the paths (functions) $q_k(t)$. First consider a single particle coordinate $q(t)$. Consider an infinitesimal change in the path $\delta q(t)$. This gives the change in the Lagrangian at each time by the chain rule

$$\delta L = \frac{\partial L}{\partial q}\delta q + \frac{\partial L}{\partial \dot{q}}\delta \dot{q} \tag{2}$$

and so the change in the action

$$\delta S = \int_{t_i}^{t_f} \delta L\, dt = \int_{t_i}^{t_f} \left(\frac{\partial L}{\partial q}\delta q + \frac{\partial L}{\partial \dot{q}}\delta \dot{q}\right) dt\,. \tag{3}$$

Note that the coordinate $q(t)$ over the path determines the velocity $\dot{q}(t)$, and so $\delta q(t)$ and $\delta\dot{q}(t)$ are not independent variables. We therefore integrate the second term by parts, using $\delta q = 0$ at the endpoints, to give

$$\delta S = \int_{t_i}^{t_f} \frac{\delta L}{\delta q} \delta q(t)\ dt, \tag{4}$$

with the *variational derivative* defined as

$$\frac{\delta L}{\delta q} = \frac{\partial L}{\partial q} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right). \tag{5}$$

We also call this the *functional derivative* of S

$$\frac{\delta S[q(t)]}{\delta q} \equiv \frac{\delta L(q, \dot{q}, t)}{\delta q}, \tag{6}$$

(same symbol, different meaning). Hamilton's principle tells us to set $\delta S = 0$. Since $\delta q(t)$ is arbitrary we get the *Euler condition* $\delta L/\delta q = 0$, i.e.

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) - \frac{\partial L}{\partial q} = 0. \tag{7}$$

This gives us the *Euler-Lagrange* equation which is the *equation of motion* for the particle. The generalization to N independent particle coordinates is straightforward, since each δq_k is independent, and we get N equations of motion $\delta L/\delta q_k = 0$, i.e.

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k}\right) - \frac{\partial L}{\partial q_k} = 0. \tag{8}$$

Comments on math

The calculus of variations is used in a much wider class of problems than optimizing the action, for example many problems involve minimizing a number that depends on a function $y(x)$ (x might be space), and derivatives

$$\text{Minimize} \qquad \int_{x_i}^{x_f} F\left(y, \frac{dy}{dx}, \frac{d^2y}{dx^2}, \ldots, x\right) dx, \tag{9}$$

(the approach generalizes to functions of higher derivatives). Examples are

- Fermat's principle of least time in optics;
- the shape of a frictionless wire in gravity connecting two points for the shortest passage time of a particle released from rest (Brachistochrone problem);
- the shortest distance between two points in flat or curved space (geodesic).

See Hand and Finch §2.1-2 and Problems 2-1 to 2-12, and Assignment 1 for a historical discussion and examples. Note that the scheme is set up for path variations with fixed endpoints. Other schemes might sometimes be appropriate, e.g. see Hand and Finch Chapter 2, Appendix A on Maupertuis principle.

Lagrangian for a single particle

Using the idea that the action should reflect the symmetries and invariances of the physical problem, I argue for a specific form for the Lagrangian of a single particle. The Euler-Lagrange equation then turns out to give Newton's first and second laws!

Free space

First consider a particle in free space. I will use Cartesian coordinates, combined into a vector $\vec{r}$ giving the position of the particle. The velocity is then $\vec{v} = \dot{\vec{r}}$. Space is uniform, and so the Lagrangian should not depend on the particle position, only on the velocity. The isotropy of space means that the Lagrangian should only depend on the magnitude of the velocity (the speed), not the its direction, i.e. $L = L(v^2)$.

We now use the idea that the equations of motion should be the same in all inertial frames, i.e. invariant under $\vec{v}(t) \to \vec{v}'(t) = \vec{v}(t) + \vec{v}_T$ with $\vec{v}_T$ a constant (the transformation velocity given by the relative motion between the two inertial frames), to restrict the function L. In fact we find that L must be proportional to v^2! To match to the Newtonian formulation, we write the proportionality constant $\frac{1}{2}m$ and call m the mass of the particle. Thus

$$L = \tfrac{1}{2}mv^2, \qquad \text{free particle.} \tag{10}$$

With this Lagrangian, let's see how the action changes under the Galilean transformation. In terms of the new velocity the action is

$$S = \frac{1}{2}m\int_{t_i}^{t_f} (\vec{v}' - \vec{v}_T)^2\, dt = \frac{1}{2}m\int_{t_i}^{t_f} v'^2\, dt - m\vec{v}_T \cdot (\vec{r}_f - \vec{r}_i) + \tfrac{1}{2}mv_T^2, \tag{11}$$

where the second term in the last expression comes from integrating $\vec{v}'$ over time, giving the separation between final and end points, which is unchanged by the Galilean transformation. Now we see that the action in terms of $\vec{v}'$ does not have exactly the same form as the original one (i.e. the action is not invariant under the Galilean transformation). However the difference is a *constant* which does not change under path variation (remember the endpoints are fixed), and so the equation of motion are unchanged. You can readily check that this does not work for other functions of v^2, for example v^4. Thus we have derived the Lagrangian and action for a particle in free space, from symmetry arguments. Note that the Euler-Lagrange equation in this case is $m\vec{v} =$ constant: the uniformity (translational invariance) of space together with Galilean invariance directly leads to momentum conservation and Newton's first law! This is a simple example of the profound relationship between symmetries and conservation laws that we will study in more generality later in the course.

Particle with forces

If the particle is not in free space, we must include a space dependent function to account for this. The simplest choice is just to add a space-dependent function $f(\vec{r})$ to the Lagrangian (rather than make the mass depend on $\vec{r}$ for example). We must then learn about the physics of what produces $f(\vec{r})$, just as Newton's second law tells us to investigate the physics of "forces". It turns out that to relate our new function to Newton's forces, we need $f = -V$, with $V(\vec{r})$ the potential energy, so that

$$L = \tfrac{1}{2}mv^2 - V(\vec{r}), \qquad \text{particle with forces.} \tag{12}$$

Note that we are restricted to *conservative forces*. Many nonconservative forces, such as friction, cannot easily be treated in the Lagrangian formulation. These are not fundamental, microscopic forces, but are derived from the mechanics of a very large number of particles (atoms), which themselves *are* governed by conservative forces. The derivation of such *dissipative* effects is quite subtle, and takes us into the realm of *statistical mechanics* rather than mechanics. The magnetic force, although nonconservative, *can* be treated in the Lagrangian approach, but via a *velocity dependent* potential: this is described below.

The Euler-Lagrange equation is now easy to write down. (Or, as an exercise, go through the explicit calculus of variations manipulations for this simple example to see how things work out.) We have for the x coordinate, $\partial L/\partial \dot{x} = m\dot{x}$ so the Euler-Lagrange equation is

$$m\ddot{x} + \frac{dV}{dx} = 0, \tag{13}$$

with similar equations for the y, z coordinates. Identifying $-dV/dx$ with force, this is Newton's second law of motion.

General Lagrangian

It is straightforward to generalize the preceding discussion to many particles. This leads us to propose the general form of the Lagrangian for mechanics

$$L = T - V, \tag{14}$$

with T the total kinetic energy and V the total potential energy (including interparticle interactions and interactions with external sources).

Advantages of the Lagrangian formulation

Hamilton's principle (leading to the Lagrangian equation of motion) is an alternative *fundamental statement* of the laws of classical mechanics with a number of advantages over the Newtonian formulation. Some advantages are

- the action is a simple scalar quantity that reflects the symmetries of the physical system, and this is often enough for us to guess the form it must take, as in the single particle problem;
- the principle is easy to derive as the $\hbar \to 0$ limit of quantum mechanics;
- the path variation may be expressed in terms of any convenient coordinates defining the particle positions: we are not restricted to Cartesian coordinates in an inertial frame, or even variables with the dimensions of length;
- for dynamics problems with constraints, such as the sliding block and wedge, we can eliminate the constraints from the formulation by using coordinates whose variations are consistent with the constraints;
- in practice, it is often easy to evaluate the action, since the kinetic energy (scalar function of the scalar speed) is usually easy to evaluate even in complicated situations (unlike the acceleration cf. Coriolis and centrifugal "forces" in noninertial frames — see later);
- the approach generalizes quite naturally to relativistic mechanics, quantum mechanics, quantum field theory ....

Most of these ideas will be illustrated in later lectures.

Relativistic Mechanics

It is interesting to try to derive relativistic mechanics using the idea that the action should respect the symmetries of the physics. To derive the equations of relativistic mechanics I make the hypothesis that the action S is a Lorentz invariant (i.e., unchanged on transforming to a different frame of reference). Since time is *not* the same in different frames, we first write the action of a particle as

$$S = \int \mathcal{L}\, d\tau$$

with τ the *proper time* – the time measured in an inertial frame instantaneously co-moving with the particle – that is Lorentz invariant. Now we look for $\mathcal{L}$, a function of particle velocity and position, that is Lorentz invariant and satisfies other symmetries of the problem. For a free particle, the only possible function is a constant! See if you can show this. Since overall scale factors in the action do not affect the equations of motion we could choose $\mathcal{L} = 1$. To connect with the Newtonian Lagrangian in the small velocity limit we instead use

$$\mathcal{L} = -mc^2,$$

with m the mass of the particle and c the speed of light (both Lorentz invariants!).

To find the equations of motion we work in a particular inertial frame, and then in terms of time t in our frame we have $d\tau = dt/\gamma$ with $\gamma = 1/\sqrt{1 - v^2/c^2}$ (time dilation). Here v is the speed of the particle measured in our frame. Thus, we have for a free particle

$$S = \int L\, dt \quad \text{with} \quad L = -mc^2\sqrt{1 - v^2/c^2}.$$

Now let's add an electromagnetic field. The only Lorentz invariant we can add to $\mathcal{L}$ that has the right symmetry properties is a constant times $\mathbf{v} \cdot \mathbf{A}$ where $\mathbf{v}$ is the *velocity 4-vector* and $\mathbf{A}$ is the *electromagnetic potential 4-vector* (the dot product of two 4-vector is a Lorentz invariant). Thus for a particle in an electromagnetic field

$$\mathcal{L} = -mc^2 - q\mathbf{v} \cdot \mathbf{A}$$

where again the constant is chosen to match onto Newtonian physics at small velocities: q is then the charge of the particle. In our inertial frame, the components of the 4-vectors are $\mathbf{v} = \gamma(c, \vec{v})$, $\mathbf{A} = (\Phi/c, \vec{A})$ with Φ the electric scalar potential and $\vec{A}$ the magnetic vector potential. Thus the Lagrangian in our frame of reference is

$$L = -mc^2\sqrt{1 - v^2/c^2} - q\Phi(\vec{x}, t) + q\vec{v} \cdot \vec{A}(\vec{x}, t).$$

With a little effort you can derive the Euler-Lagrange equation, and leading in the Newtonian limit to the equation of motion for a particle with the Lorentz force $\vec{f} = q(\vec{E} + \vec{v} \times \vec{B})$ (see Assignment 1, Problem 4).

Michael Cross, October 2, 2013