221A Lecture Notes
Notes on Classica Mechanics I

1 Precursor: Fermat's Principle in Geometric Optics

In geometric optics, you talk about how light rays go. In homogeneous mediums, the light rays go straight. At the interace between different mediums, light rays get reflected or refracted. Fermat formulated the problem in a very elegant way: the light rays choose their paths to minimize the optical length. The optical length, as you know, is defined by the actual length of the path times the index of refraction of the medium. Allowing the index of refraction to vary as a function of the position $n(\vec{x})$, the optical length is given by

$$l = \int n(\vec{x})dl, \tag{1}$$

where $dl = \sqrt{(dx)^2 + (dy)^2 + (dz)^2}$ is the line segment along the path. The Fermat's principle is stated mathematically as

$$\delta l = 0 \tag{2}$$

along the actual path of a light ray.

In a homogeneous medium, the index of refraction is a constant, and the problem is just to minimize the path length, correctly choosing the straight lines. If there are two mediums attached at $z = 0$, Fermat's principle tells us that the light ray gets refracted at the interface. Choosing the initial point as $(x_i, 0, z_i)$ ($z_i < 0$, the index of refraction n_1) and the final point $(x_f, 0, z_f)$ ($z_f > 0$, the index of refraction n_2), the only parameter to be determined is the value of x at $z = 0$ where the light ray gets refracted. The optical length is

$$l = n_1\sqrt{(x - x_i)^2 + z_i^2} + n_2\sqrt{(x_f - x)^2 + z_f^2}. \tag{3}$$

Minimizing it with respect to x,

$$0 = \frac{dl}{dx} = n_1\frac{x - x_i}{\sqrt{(x - x_i)^2 + z_i^2}} - n_2\frac{x_f - x}{\sqrt{(x_f - x)^2 + z_f^2}} = n_1\sin\theta_1 - n_2\sin\theta_2. \tag{4}$$

This is nothing but Snell's law of refraction. On the other hand, if the light ray is reflected at the surface, $z_f < 0$ and the optical length is

$$l = n_1\sqrt{(x - x_i)^2 + z_i^2} + n_1\sqrt{(x_f - x)^2 + z_f^2}. \tag{5}$$

Minimizing it with respect to x,

$$0 = \frac{dl}{dx} = n_1\frac{x - x_i}{\sqrt{(x - x_i)^2 + z_i^2}} - n_1\frac{x_f - x}{\sqrt{(x_f - x)^2 + z_f^2}} = n_1(\sin\theta_1 - \sin\theta_2). \tag{6}$$

Therefore the incident and reflection angles are the same, as we know well.

If the index of refraction depends continuosly on the position, you have to derive a differential equation from Fermat's principle and solve it. In order to do so, you can introduce a parameter along the path s so that the positions on the path are given by functions of s: $\vec{x}(s) = (x(s), y(s), z(s))$. Then the optical length is written as an integral over the parameter s as

$$l[\vec{x}(s)] = \int n(\vec{x}(s))\sqrt{\dot{\vec{x}}^2}\,ds, \tag{7}$$

where the dot refers to the s-derivative $\dot{x} = dx/ds$ etc. Now we take the variation of l with respect to the path, and require that the variation vanishes. This will choose the paths of extremal legnth, which can be either minimum or maximum. Fermat's principle actually allows the maximum as well. Plugging in $\vec{x}(s) + \delta\vec{x}(s)$ in the optical length and keeping only the first order in $\delta\vec{x}$, we find

$$l[\vec{x}(s) + \delta\vec{x}(s)] - l[\vec{x}(s)] = \int\left[\left(\delta\vec{x}\cdot\vec{\nabla}n\right)\sqrt{\dot{\vec{x}}^2} + n\frac{\delta\dot{\vec{x}}\cdot\dot{\vec{x}}}{\sqrt{\dot{\vec{x}}^2}}\right]ds. \tag{8}$$

Now we require that this expression vanishes for any variation $\delta\vec{x}(s)$. By integrating the second term in parts, find

$$0 = \int \delta\vec{x}\cdot\left[\vec{\nabla}n\sqrt{\dot{\vec{x}}^2} - \frac{d}{ds}n\frac{\dot{\vec{x}}}{\sqrt{\dot{\vec{x}}^2}}\right]ds \tag{9}$$

for any variation, and hence

$$\left(\vec{\nabla}n\right)\sqrt{\dot{\vec{x}}^2} - \frac{d}{ds}n\frac{\dot{\vec{x}}}{\sqrt{\dot{\vec{x}}^2}} = 0. \tag{10}$$

2

Finally, it is useful to note that s is an arbitrary parameter, and we can always choose the parameter such that $\sqrt{\dot{\vec{x}}^2} = n$ (or you can choose it to be a constant, too, if you want, but that would be less convenient). Then the equation simplifies drastically to

$$\ddot{\vec{x}} = n\vec{\nabla}n, \qquad \dot{\vec{x}}^2 = n^2. \tag{11}$$

In other words, the equation is the same for a particle of unit mass moving in the "potential energy" $V = -n^2(\vec{x})/2$ as a function of "time" s with the "total energy" $E = \frac{1}{2}\dot{\vec{x}}^2 + V = \frac{1}{2}\dot{\vec{x}}^2 - \frac{1}{2}n^2(\vec{x}) = 0$. With this idea, the Snell's law is "explained" because the "particle" is "pulled" into the region with higher index of refraction where it moves "faster" and moves away quicker from the interface in the perpendicular direction. This is indeed the original Newton's Corpuscular Theory of Light; this interpretation of course turned out to be wrong, but the equation we have here is nonetheless the right one for geometric optics.[1]

For example, in fiber optics, you set up a "potential well" that has a minimum at the center of the cable and goes up away from the center. If you approximate the "potential well" by a quadratic function around the minimum $V = -\frac{1}{2}n_1^2 + \frac{1}{2}\omega^2 r^2$, the "particle" oscillates around the center with the frequency ω just as a harmonic oscillator, and does not go away too far from the center.

But in this case you have to be careful that the "time" s has nothing to do with the actualy time t. In fact, the paths "faster" in s are often slower in t.

2 Lagrangian

Lagrangian formalism for mechanics uses the "action" whose minimum (or more precisely, extremum) determines the equation of motion. Clearly, it was inspired by Fermat's principle for optics, and I frankly don't see any motivations to formulate mechanics this way other than aesthetics. It turns out, however, that the symmetry of the system and change of variables are much clearer with Lagrangian than with Newton's equation of motion $\vec{F} =$

---

[1]It is ironical that one of the earliest evidence for wave nature of light is Newton's rings. See an interest account at `http://web.clas.ufl.edu/users/rhatch/pages/01-Courses/current-courses/08sr-newton.htm`.

$m\vec{a}$. Moreover, it appears very naturally in the path-integral formulation of quantum mechanics.

A Lagrangian[2] depends on the (generalized) coordinates $q_i(t)$ and their first time-derivative $\dot{q}_i(t)$. The action is defined by integrating the Lagrangian over time,

$$S[q_i(t)] = \int_{t_i}^{t_f} L(q_i(t), \dot{q}_i(t))dt. \tag{12}$$

The action is said to be a "functional", a function of function. The trajectory $q(t)$ is a function of time, and the action is a function of the trajectory $S[q(t)]$, a functional. The principle of the least action states that the particles follow trajectories that minimize the action. (Again, the trajectory can be extremum, not necessarily a minimum.) By taking the variation of the action with respect to q_i,

$$\delta S = S[q_i + \delta q_i] - S[q_i] = \int_{t_i}^{t_f} \left(\frac{\partial L}{\partial q_i} \delta q_i + \frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i \right). \tag{13}$$

The latter term is integrated by parts, and

$$\delta S = \int_{t_i}^{t_f} \left(\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right) \delta q_i + \frac{\partial L}{\partial \dot{q}_i} \delta q_i \Big|_{t_i}^{t_f}. \tag{14}$$

Here, we are interested in variations where the initial and final positions are fixed $\delta q_i(t_i) = \delta q_i(t_f) = 0$. Then the action is minimized if

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0. \tag{15}$$

This is Euler–Lagrange equation of motion.

For example, the Lagrangian of a point particle in a potential $V(q)$ is

$$L = \frac{m}{2} \dot{q}_i^2 - V(q). \tag{16}$$

The Euler–Lagrange equation is then given by

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = m\ddot{q}_i + \frac{\partial V}{\partial q_i} = 0. \tag{17}$$

---

[2]I've once read a debate article if the correct spelling is "Lagrangian" or "Lagrangean." The name of the person is Lagrange, but he is originally from Italy, with the original name Lagrangi. Apparently most people have decided that Lagrangian is correct. But we talk about Euler–Lagrange equation of motion, Lagrange multiplier, etc. Can anybody explain why to me?

This is nothing but Newton's equation of motion $F = ma$ with the identifications $F_i = -\partial V/\partial q_i$, $a_i = \ddot{q}_i$.

Symmetry of the system is much more manifest in Lagrangians than in Newton's equation of motion. For example, the case of the electron in the hydrogen-like atoms is given by $V = -Ze^2/r$, a spherically symmetric system. In Newton's equation,

$$m\vec{a} = \frac{Ze^2 \vec{r}}{r^3}, \tag{18}$$

which is *covariant* under the rotation. In other words, there are three equations of motion for x, y, z, and they are mixed with each other under the rotation. But the Lagrangian is *invariant* under the rotation:

$$L = \frac{m}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) + \frac{Ze^2}{r}. \tag{19}$$

Already at the first sight, it is clear that the system is rotationally invariant.

Another advantage of using Lagrangian is the ease in changing the variables. For instance, going to the spherical coordinate is straight-forward for the above example of hydrogen-like atoms,

$$L = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\dot{\phi}^2) + \frac{Ze^2}{r}. \tag{20}$$

The Euler–Lagrange equations follow very easily,

$$m\ddot{r} + \frac{Ze^2}{r^2} = 0, \tag{21}$$

$$m\frac{d}{dt}(r^2\dot{\theta}) - mr^2\sin\theta\cos\theta\dot{\phi}^2 = 0, \tag{22}$$

$$m\frac{d}{dt}(r^2\sin^2\theta\dot{\phi}) = 0. \tag{23}$$

You already see that there is a conserved quantity, $mr^2\sin^2\theta\dot{\phi}$, which is nothing but the angular momentum L_z. But anybody who tried to rewrite the Newton's equation of motion in spherical coordinates from that in the Cartesian coordinates,

$$m\ddot{\vec{r}} + \frac{Ze^2}{\vec{r}}r^3 = 0, \tag{24}$$

would know that it is *much* more tedious that way.

5

In the path integral formulation of quantum mechanics, the quantum mechanical amplitudes are expressed in terms of integral of a phase factor $e^{iS[q(t)]/\hbar}$ for all possible paths $q(t)$. The extremum of the action is nothing but a pass $q(t)$ where the phase factor is stationary. In the limit of $\hbar \to 0$, the stationary phase approximation of the integral tells you that the integral is dominated by the stationary phase configuration, which is the Euler–Lagrange equation. This way, the classical mechnics is reproduced in the $\hbar \to 0$ limit.

3   Hamiltonian

The Hamitonian of the sytem is nothing but the total energy, but it deserves a special name because of its important role in both classical and quantum mechanics.

The first thing you define in Hamiltonian mechanics is the "canonical momentum" p_i "conjugate" to the "canonical coordinates" q_i. Then the Legendre transform from the original variables $(\dot{q}_i, q_i)$ in the Lagrangian to the new variables (p_i, q_i) defines the Hamiltonian. This is very similar to what you do in thermodynamics. For instance, the internal energy U is regarded as a function of the entropy S and the volume V. The temperature of the sytem is defined by

$$T = \left.\frac{\partial U}{\partial S}\right|_V. \tag{25}$$

The Legendre transform from the variables (S, V) to (T, V) defines the free energy $F(T, V)$

$$F(T, V) = U(S, V) - TS \tag{26}$$

and the inverse Legendre transform defines the entropy as

$$S = -\left.\frac{\partial F}{\partial T}\right|_V. \tag{27}$$

The temperature T and the entropy S are "conjugate." What we do now is the same Legendre transform from the variables $(\dot{q}_i, q_i)$ to (p_i, q_i), except that the Hamiltonian has the opposite sign from what you would naively do.

Starting from the Lagrangian $L(\dot{q}_i, q_i)$ that depends on q_i and $\dot{q}_i$, the

canonical momenta are defined by

$$p_i = \left.\frac{\partial L}{\partial \dot{q}_i}\right|_{q_i}. \tag{28}$$

And then you define the Hamitonian (here comes the opposite sign!)

$$H(p_i, q_i) = p_i \dot{q}_i - L(\dot{q}_i, q_i). \tag{29}$$

The point is that you always *eliminate* $\dot{q}_i$ using p_i. The Hamiltonian is a function of p_i and q_i only, without any time derivatives of them. The inverse Legendre transform brings $\dot{q}_i$ back:

$$\dot{q}_i = \left.\frac{\partial H}{\partial p_i}\right|_{q_i}. \tag{30}$$

We normally do not write it explicitly that q_i are held fixed in taking derivatives with respect to p_i, because it is understood that the Hamiltonian is a function of independent variables (p_i, q_i).

The Euler–Lagrange equation of motion is then rewritten in the following way.

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = \frac{d}{dt}p_i - \frac{\partial L}{\partial q_i} = 0. \tag{31}$$

The last term $\partial L/\partial q_i$ is calculated with $\dot{q}_i$ held fixed, because it was derived from the variation of the action. It turns out to be the same as $-\partial H/\partial q_i$ where p_i are held fixed. This can be seen as follows:

$$
\begin{aligned}
\left.\frac{\partial L}{\partial q_i}\right|_{\dot{q}_i} &= \left.\frac{\partial (p_j \dot{q}_j - H)}{\partial q_i}\right|_{\dot{q}_i} \\
&= \left.\frac{\partial p_j}{\partial q_i}\right|_{\dot{q}_i} \dot{q}_j + p_j \left.\frac{\partial \dot{q}_j}{\partial q_i}\right|_{\dot{q}_i} - \left.\frac{\partial H}{\partial q_i}\right|_{q_i} - \frac{\partial H}{\partial p_j} \left.\frac{\partial p_j}{\partial q_i}\right|_{\dot{q}_i} \\
&= -\frac{\partial H}{\partial q_i}.
\end{aligned} \tag{32}
$$

At the last step, we used Eq. (30) to cancel the first and the last terms, and the trivial point that

$$\left.\frac{\partial \dot{q}_i}{\partial q_i}\right|_{\dot{q}_i} = 0 \tag{33}$$

because $\dot{q}_i$ are held fixed. Therefore, Eq. (31) becomes

$$\frac{d}{dt}p_i + \frac{\partial H}{\partial q_i} = 0. \tag{34}$$

Putting Eqs. (30) and (34) together,

$$\frac{d}{dt}p_i = -\frac{\partial H}{\partial q_i}, \tag{35}$$

$$\frac{d}{dt}q_i = \frac{\partial H}{\partial p_i}. \tag{36}$$

These are called Hamilton equations of motion.

One of the most important properties of Hamilton equations of motion is that they are first order differential equations, while the Euler–Lagrange equation of motion is second order. This simple point makes the mathematical structure much more tractable. In studies of chaos, non-linear dynamics and beam dynamics, Hamilton equations of motion are indispensable.

Another important property of the Hamiltonian is that it is conserved; it is nothing but the total energy! The fact that the Hamitonian is conserved $dH/dt = 0$ can be checked explicitly using Euler–Lagrange equation of motion,

$$\begin{aligned}
\frac{dH}{dt} &= \frac{d}{dt}(p_i\dot{q}_i - L) \\
&= \frac{d}{dt}\left(\dot{q}_i\frac{\partial L}{\partial \dot{q}_i} - L\right) \\
&= \dot{q}_i\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} + \ddot{q}_i\frac{\partial L}{\partial \dot{q}_i} - \dot{q}_i\frac{\partial L}{\partial q_i} - \ddot{q}_i\frac{\partial L}{\partial \dot{q}_i} = 0,
\end{aligned} \tag{37}$$

where the first and third terms cancel because of the Euler–Lagrange equation, while the second and fourth cancel trivially. Note that we used the fact the only time dependence of L comes through the time-dependence of q_i and $\dot{q}_i$, not explicitly $\partial L/\partial t|_{q,\dot{q}} = 0$. Namely, we assumed the invariance of the Lagrangian under time translation.

Much more important is the fact that the conservation of the Hamiltonian can also be shown as a consequence of the invariance of the system under the time translation. Under the time translation $t \to t + \delta t$, the action changes as

$$S = \int_{t_i}^{t_f} L dt \to \int_{t_i+\delta t}^{t_f+\delta t} L dt = S + L(t_f)\delta t - L(t_i)\delta t. \tag{38}$$

8

and hence $\delta S = L(t_f)\delta t - L(t_i)\delta t$. On the other hand, it can also be worked out as

$$
\begin{aligned}
\delta S &= \delta \int_{t_i}^{t_f} L dt \\
&= \int_{t_i}^{t_f} \delta L dt \\
&= \int_{t_i}^{t_f} \left(\frac{\partial L}{\partial q_i} \dot{q}_i \delta t + \frac{\partial L}{\partial \dot{q}_i} \ddot{q}_i \delta t \right) dt \\
&= \int_{t_i}^{t_f} \left(\left(\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right) \dot{q}_i + \frac{\partial L}{\partial \dot{q}_i} \ddot{q}_i \right) \delta t dt \\
&= \left. \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i \right|_{t_i}^{t_f} \delta t.
\end{aligned}
\tag{39}
$$

Comparing above two equations, we find

$$
L(t_f) - L(t_i) = \left. \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i \right|_{t_i}^{t_f}
\tag{40}
$$

or

$$
(p_i \dot{q}_i - L)(t_f) = (p_i \dot{q}_i - L)(t_i).
\tag{41}
$$

Therefore, systems invariant under the time translation lead to the conserved Hamiltonian.

A corollary of this statement is that systems not invariant under the time translation do not conserve energy. For example, ever-expanding Universe does not have conserved energy. The energy can decrease (adiabatic expansion of thermal gas) or can increase (the mysterious negative-pressure Dark Energy found recently).

In quantum mechanics, conserved quantities associated with an invariance of the system actually *generate* the invariance. I'll make this point clear in later sections.

4   Poisson Bracket

Poisson Bracket between two physical quantities $A(p_i, q_i)$ and $B(p_i, q_i)$ is defined as

$$
\{A, B\} = \frac{\partial A}{\partial q_i} \frac{\partial B}{\partial p_i} - \frac{\partial A}{\partial p_i} \frac{\partial B}{\partial q_i}.
\tag{42}
$$

Why do we define such a thing? Because it is useful.

First of all, the Hamilton equations of motion Eq. (36) can be written as

$$\frac{d}{dt}p_i = \{p_i, H\}, \tag{43}$$

$$\frac{d}{dt}q_i = \{q_i, H\}. \tag{44}$$

In fact, it can easily be shown that the time derivative for *any* physical quantity can be written as

$$\frac{d}{dt}A = \{A, H\}. \tag{45}$$

In this sense, the Hamiltonian is said to *generate* the time translation, because Poisson brackets with the Hamiltonian pushes the time forward.

In quantum mechanics, this point is much more manifest, because Schrödinger equation is nothing but the Hamiltonian pushing the time forward:

$$i\hbar \frac{\partial}{\partial t}|\psi\rangle = H|\psi\rangle. \tag{46}$$

To mathematicians, Poisson brackets define the symplectic structure on the manifold (phase space): a closed two-form $\omega = \sum_{i=1}^{n} dp_i \wedge dq_i$ which gives the phase-space volume element $dV = \omega^n$. The time evolution of the system is described by the Hamiltonian vector

$$\frac{\partial}{\partial t} = (\dot{p}_i \frac{\partial}{\partial p_i}, \dot{q}_i \frac{\partial}{\partial q_i}) = (-\frac{\partial H}{\partial q_i}\frac{\partial}{\partial p_i}, \frac{\partial H}{\partial p_i}\frac{\partial}{\partial q_i}) = \frac{\partial H}{\partial x_i}\omega_{ji}\frac{\partial}{\partial x_i},$$

where $x_i = (q_i, p_i)$ is the phase space coordinate.

Similarly, the total momentum is conserved because of the invariance under the spatial translation. For example, the total momentum is nothing but $P_x = \sum_i p_{i,x}$ for one-dimensional problems. Then it is easy to see that

$$\sum_i \frac{d}{q_{i,x}}A = \{A, P_x\}. \tag{47}$$

Indeed, the total momentum *generates* the overall translation for all positions q_i.

The connection is also very interesting for angular momenta. Starting with the orbital angular momenta $\vec{L} = \vec{q} \times \vec{p}$ for a point particle, we can calcualte the Poisson brackets among them.

$$\{L_x, L_y\} = \{q_y p_z - q_z p_y, q_z p_x - q_x p_z\} = -q_y p_x + q_x p_y = L_z, \tag{48}$$

$$\{L_y, L_z\} = \{q_z p_x - q_x p_z, q_x p_y - q_y p_x\} = -q_z p_y + q_y p_z = L_x, \tag{49}$$

$$\{L_z, L_x\} = \{q_x p_y - q_y p_x, q_y p_z - q_z p_y\} = -q_x p_z + q_z p_x = L_y. \tag{50}$$

This is exactly the same as the commutation relation among the spin operators in quantum mechanics

$$[S_x, S_y] = i\hbar S_z, \qquad [S_y, S_z] = i\hbar S_x, \qquad [S_z, S_x] = i\hbar S_y. \qquad (51)$$

It turns out that the angular momentum is conserved because of the rotational invariance of the system, and the angular momentum *generate* the rotation. Because rotation around x, y, z axes are not independent and are related in a particular way, the Poisson brackets or commutation relation among angular momenta cannot be anything else.

Poisson brackets are promoted to commutation relations among operators in quantum mechanics. I'm truly amazed that the classical mechanis before the dawn of quantum mechanics came up with its formulation that so closely resembles the commutation relation among operators!