INTRODUCTION TO CLASSICAL MECHANICS

ERIC Y. CHEN

1. Lagrangian Formulation

The central question in classical mechanics is: given some particles moving in a space, possibly with potential U, and given the initial position and momentum, can you predict what happens at a given time t?

There are two main ways to formulate the answer:

- **Newtonian formulation.** Let $x(t) : \mathbb{R}_t \to \mathbb{R}^3$ be the trajectory of our particle, and $U(x)$ be the given potential. Then
$$\operatorname{grad} U = m\ddot{x}$$
- **Lagrangian formulation.** There is a real valued function $\mathcal{L}(x, \dot{x})$ such that the critical points of
$$\begin{aligned} S : \mathrm{Maps}(\mathbb{R}_t, \mathbb{R}^3) &\longrightarrow \mathbb{R} \\ x &\longmapsto \int_{\mathbb{R}} \mathcal{L}(x, \dot{x})\, dt \end{aligned}$$
are the preferred trajectories. We call the equation $dS = 0$ the **equations of motion**.

The two formulations are essentially equivalent by fundamental theorem of calculus. Let's look at a simple example.

Example 1. Consider the Lagrangian consisting of a kinetic term and a potential term
$$\mathcal{L}(x, \dot{x}) = \frac{1}{2} m\dot{x}^2 + U(x)$$
Let's compute the critical paths. By definition, we need to make a small variation $x \mapsto x + \varepsilon y$ and check that the first variation in S vanishes:
$$\begin{aligned} S(x + \varepsilon y) &= \int \frac{1}{2} m(\dot{x} + \varepsilon \dot{y})^2 + U(x + \varepsilon y)\, dt \\ &= S(x) + \varepsilon \int m\dot{x}\dot{y} + U'(x) y\, dt + O(\varepsilon^2) \\ &= S(x) + \varepsilon \int (-m\ddot{x} + U'(x)) y\, dt + O(\varepsilon^2) \end{aligned}$$
In the last line we used integration by parts, which required
$$\dot{x} y = 0 \ \text{ for } |t| \to \infty$$
There are two ways to make it happen: force $y = 0$ at infinity, this is called **Dirichlet boundary condition**, or force $\dot{x} = 0$ at infinity, this is called **Neumann boundary condition**. In any case, we want the first variation to vanish for all y, which gives us
$$U'(x) = m\ddot{x}$$
which is Newton's law, as we expected.

If we start with a general Lagrangian, the resulting "Newton's law" is called **Euler-Lagrange equations**.

Theorem 1.1 (Euler-Lagrange Equations)**.** *Assuming either Neumann or Dirichlet boundary conditions, the equations of motions are satisfied if and only if*

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}} = \frac{\partial \mathcal{L}}{\partial x}$$

Proof. Let $x \mapsto x + \varepsilon y$ be a small variation. Then the Lagrangian varies by

$$\mathcal{L}(x+\varepsilon y, \dot{x}+\varepsilon\dot{y}) = \mathcal{L}(x,\dot{x}) + \varepsilon\left[\frac{\partial \mathcal{L}}{\partial x}y + \frac{\partial \mathcal{L}}{\partial \dot{x}}\dot{y}\right] + O(\varepsilon^2)$$

So the variation in the action is

$$\delta S = \int \frac{\partial \mathcal{L}}{\partial x}y + \frac{\partial \mathcal{L}}{\partial \dot{x}}\dot{y}\,dt = \int \left[\frac{\partial \mathcal{L}}{\partial x} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}}\right] y\,dt$$

where we used the boundary conditions and integration by parts. This vanishes for every y if and only if the Euler-Lagrange equations hold. □

2. Conserved Quantities

There are two main reasons why the Lagrangian formulation is more convenient than the Newtonian one:

- The Euler-Lagrange equations hold in any choice of coordinates, unlike Newton's equations.
- Symmetries are more evident: this will be the main theme in many classical and quantum systems we consider.

Suppose we have a Lie group G "acting" on our classical system; what does this mean? Maybe we mean that G leaves $\mathcal{L}$ invariant, or $\delta\mathcal{L} = 0$ for a variation generated by any element in $\mathfrak{g} = \mathrm{Lie}(G)$. But actually it just has to preserve S, so it is allowed to change $\mathcal{L}$ by a total derivative, i.e., if $\varepsilon y \in \mathfrak{g}$ and $\delta\mathcal{L} = dF/dt$ for some function F vanishing at the boundary, then

$$\delta S = \int \frac{dF}{dt}\,dt = 0$$

where $\mathrm{EL}(x)$ vanishes for a critical path x.

The following theorem is one of the most used "principles" in physics, in classical/quantum/mechanics/field theory, as it lets us translate between the notions of symmetry and **conserved quantities**, i.e., an observable quantity that does not change in time.

Theorem 2.1 (Noether's Theorem)**.** *Suppose $x \mapsto x + \varepsilon y$ is an infinitesimal symmetry in the above sense, i.e.,*

$$\delta\mathcal{L} = \frac{dF}{dt}$$

Then the quantity

$$J := \frac{\partial \mathcal{L}}{\partial \dot{x}}y - F$$

is conserved in time.

Proof. The first variation of the Lagrangian is

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial x}y + \frac{\partial\mathcal{L}}{\partial\dot{x}}\dot{y}$$

which is equal to dF/dt by assumption. Now we can "complete the product rule" and write

$$0 = \frac{\partial\mathcal{L}}{\partial x}y + \frac{\partial\mathcal{L}}{\partial\dot{x}}\dot{y} - \frac{dF}{dt} = \frac{\partial\mathcal{L}}{\partial x}y + \left[-\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{x}}y + \frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{x}}y\right] + \frac{\partial\mathcal{L}}{\partial\dot{x}}\dot{y} - \frac{dF}{dt} \tag{1}$$

$$= \mathrm{EL}(x)y + \frac{d}{dt}\left[\frac{\partial\mathcal{L}}{\partial\dot{x}}y - F\right] \tag{2}$$

If x is critical, then $\mathrm{EL}(x) = 0$, and we see that $J := \frac{\partial\mathcal{L}}{\partial\dot{x}}y - F$ is conserved in time. $\square$

Example 2. Let's take the very simple example of harmonic oscillator, whose Lagrangian looks like

$$\mathcal{L} = \frac{1}{2}\dot{x}^2 - \frac{1}{2}x^2$$

The Euler-Lagrange equations can be calculated as

$$\ddot{x} = -x$$

which is also known as Hook's law. This differential equation is easy enough to solve (like $x = \cos(t)$), but conserved quantities are not so easy to guess in the outset.

Let's consider a symmetry that will always work: time translation, i.e., $x \mapsto x(t + \varepsilon\tau)$. Then $\delta\mathcal{L} = d\mathcal{L}/dt$ by definition, which is a total derivative. The corresponding conserved current is

$$J = \frac{\partial\mathcal{L}}{\partial\dot{x}}\dot{x} - \mathcal{L} = \frac{1}{2}\dot{x}^2 + \frac{1}{2}x^2$$

The level sets of J are circles in the **phase space** where we plot $(x, \dot{x})$. In retrospect this is just $\cos^2 + \sin^2 = 1$ once we solve Newton's equations, but clearly we can't always hope to solve the equations of motion.

Let's look at this example a bit more artificially.

Example 3. For any Lagrangian at all, we can consider time translation symmetry. Since $\delta\mathcal{L} = d\mathcal{L}/dt$ by definition of time translation, the conserved current is easy to write:

$$H := \frac{\partial\mathcal{L}}{\partial\dot{x}}\dot{x} - \mathcal{L}$$

This conserved quantity gets a special name called the **Hamiltonian**, and it represents energy. We will see that classical systems given by a Hamiltonian is essentially one that privileges time translation symmetry.

Remark 2.2. It is often said, but we won't justify here, that H is the Legendre transform of $\mathcal{L}$, and the Legendre transform is involutive, so a mechanical system could be equivalently be described by giving Hamiltonian instead.

We're gonna now let q be the position variable now because it's more customary, and define its **conjugate momentum** as

$$p := \frac{\partial\mathcal{L}}{\partial\dot{q}}$$

We won't really use this today, but it looks really nice, which is one of the pros of the Hamiltonian perspective.

Proposition 2.3 (Hamilton's Equations)**.** *Let $\mathcal{L}$ be a Lagrangian, and $H = p\dot{q} - \mathcal{L}$ the corresponding Hamiltonian. Then the equations of motion are*

$$\frac{\partial H}{\partial p} = \dot{q} \quad \textit{and} \quad \frac{\partial H}{\partial q} = -\dot{p}$$

Proof. Taking the total dervative of the Hamiltonian, we get

$$dH = pd\dot{q} + \dot{q}dp - \frac{\partial \mathcal{L}}{\partial q}dq - \frac{\partial \mathcal{L}}{\partial \dot{q}}d\dot{q} = \dot{q}dp - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}}\, dq = \dot{q}dp - \dot{p}dq$$

by applying Euler-Lagrange and recognizing that the first and last terms cancel. These are Hamilton's equations. □

Formally speaking, the data involved here is the following: the path space Ω has a canonical 1-form α, which is given by the variation of the Lagrangian. Its differential $\omega = d\alpha$ endows Ω with the structure of a symplectic manifold. If X is a vector field on Ω, then we can take its ω-dual 1-form $\iota_X\omega$, and ask if it is integrable to a function f_X. This function is conserved in time if the conditions of Noether's theorem are satisfied.

We can also go backwards: suppose J is a conserved quantity, then the ω-dual of dJ is a vector field which we call X_J. Viewing it as a derivation on functions, we write the corresponding derivation as $\{J, \cdot\}$, which is called the **Poisson bracket**.

The following basic facts about Poisson structure is easily verified with the help of Cartan's magic formula:

Proposition 2.4. *Let M be a symplectic manifold, and $\mathrm{Ham}(M) \subset \mathrm{Vect}(M)$ be the subalgebra generated by Hamiltonian vector fields. Then we have the following:*

(1) $\{f_X, f_Y\} = \omega(X, Y)$*; in particular, the Poisson bracket is antisymmetric.*
(2) *The maps*

$$\mathrm{Ham}(M) \longrightarrow \mathcal{O}_M \longrightarrow \mathrm{Vect}(M)$$

are homomorphisms of Lie algebras, and the composition is the identity map.

Example 4. The Hamiltonian is the conserved quantity corresponding to time translation, so $\{H, f\} = df/dt$. In particular, f is conserved if and only if $\omega(X_H, X_f) = 0$.

Another direct consequence is the following Corollary, which we also won't be using but feels worth mentioning.

Corollary 2.5 (Arnold-Liouville)**.** *Let M be a $2n$-dimensional symplectic manifold, and $f_1, \dots, f_n$ Poisson commuting functions on M. Suppose $(f_1, \dots, f_n) : M \to \mathbb{R}^n$ is a submersion onto an open subset. Suppose the map has smooth comapct fibers. Then M is a Lagrangian torus fibration over its image.*

Proof. The fact that the functions $f_1, \dots, f_n$ Poisson commute implies that the corresponding flows $X_1, \dots, X_n$ lie in a Lagrangian. So we have an action of $\mathbb{R}^n$ on M along fibers, which we assumed are compact, so they must be Lagrangian tori. □

3. The Moment Map

Consider a Lie group G acting on a symplectic manifold M. Assume the action is **Hamiltonian**, meaning that the Lie algebra $\mathfrak{g} = \mathrm{Lie}(G)$ acts via Hamiltonian vector fields.

Definition 3.1. *The moment map is a map* $\mu : M \longrightarrow \mathfrak{g}^*$ *defined by*

$$\mu(p)(X) = f_X(p)$$

It's easy to see that μ is G-equivariant: for $X, Y \in \mathfrak{g}$,

$$\mu(p + \varepsilon Y)(X) = f_X(p + \varepsilon Y) = f_X(p) + \varepsilon df_X(Y)(p) = f_X(p) + \varepsilon f_{\mathrm{ad}_Y(x)}(p)$$

So infinitesimally translating μ by Y amounts to ad_Y^* on the functional $\mu(p)$.

Definition 3.2. *With the above setup, the symplectic reduction of* M *by* G *is*

$$M /\!\!/ G := \mu^{-1}(0)/G$$

Note that this procedure is expected to reduce the dimension of M by $2\dim(G)$, since taking the moment map fiber kills $\dim(G)$ and the quotient kills another $\dim(G)$. There are many many examples of these symplectic reductions in algebraic/symplectic geometry, but in the spirit of the seminar we are going to use these concepts in a physical example.

Consider n particles moving in $\mathbb{R}^3$ with pairwise interacting potential that depends only on the pairwise distance. Then the Lagrangian looks like

$$\mathcal{L}(r, \dot{r}) = \frac{1}{2}\sum_{i=1}^{n} m_i \dot{r}_i^2 - \sum_{i,j} V(|r_i - r_j|)$$

It's obvious that **translation in space** leaves $\mathcal{L}$ literally invariant. For $v \in \mathbb{R}^3$ the conserved charge corresponding to v-translation is

$$\frac{\partial \mathcal{L}}{\partial \dot{r}} \cdot v = \left(\sum_{i=1}^{n} m_i \dot{r}_i\right) \cdot v$$

The parenthesized quantity is the **total momentum** of the system, and is the image of the moment map $\mu : \{(r, \dot{r})\} \to (\mathbb{R}^3)^*$.

Next, we consider **rotations**. Let $w \in \mathbb{R}^3$ be the axis with respect to which we do an infinitesimal rotation. The r_i's transform as

$$r_i \longmapsto r_i + \varepsilon w \times r_i$$

So its time derivative transforms as

$$\dot{r}_i \longmapsto \dot{r}_i + \varepsilon w \times \dot{r}_i$$

The variation in the Lagrangian only happens in the kinetic term:

$$\frac{1}{2} m_i \dot{r}_i^2 \longmapsto \frac{1}{2} m_i (\dot{r}_i + \varepsilon w \times \dot{r}_i)^2 = \frac{1}{2} m_i \dot{r}_i^2 + 2\varepsilon \dot{r}_i \cdot (w \times \dot{r}_i) = \frac{1}{2} m_i \dot{r}_i^2$$

So the rotation also leaves $\mathcal{L}$ invariant. The conserved quantity here is

$$\frac{\partial \mathcal{L}}{\partial \dot{r}} \cdot (w \times r) = \sum_{i=1}^{n} (m_i \dot{r}_i) \cdot (w \times r_i) = w \cdot \left(\sum_{i=1}^{n} r_i \times m_i \dot{r}_i\right)$$

The parenthesized quantity is the **total angular momentum** of the system, and is the image of the moment map $\mu : \{(r, \dot{r}\} \to \mathfrak{so}_3^*$.

Example 5. Let's apply the above to the specific example of two particles, and let's set the masses $m_1 = m_2 = 1$. The Lagrangian is

$$\mathcal{L} = \frac{1}{2}\dot{r}_1^2 + \frac{1}{2}\dot{r}_2^2 - V(|r_1 - r_2|)$$

Let's apply symplectic reduction by spatial translations. In other words, we set the total momentum $\dot{r}_1 + \dot{r}_2 = 0$. We can massage the Lagrangian so that this expression occurs:

$$\mathcal{L} = \frac{1}{4}\left[(\dot{r}_1 + \dot{r}_2)^2 + (\dot{r}_1 - \dot{r}_2)^2\right] - V(|r_1 - r_2|) = \frac{1}{4}\dot{r}^2 - V(|r|)$$

where we set $r = r_1 - r_2$ to be a new variable; this is a choice of new coordinates on the symplectic quotient. We see that the original system had 6×2 degrees of freedom (6 spatial and 6 momenta), and we have reduced to $6 \times 2 - 3 \times 2 = 3 \times 2$ degrees of freedom.

Let's apply symplectic reduction by a rotation. We set $r \times \dot{r} = 0$. This expression again does not occur literally in the Lagrangian, but it tells us that r stays on a plane, and we can pick polar coordinates (ρ, ϕ) in that plane. Then the Lagrangian can be rewritten as

$$\mathcal{L} = \frac{1}{4}(\dot{\rho}^2 + \rho^2\dot{\phi}^2) - V(|\rho|)$$

We are reduced to $3 \times 2 - 1 \times 2 = 2 \times 2$ degrees of freedom.

Now we observe that $\mathcal{L}$ does not depend on ϕ at all! So translation in ϕ is a symmetry, and the conserved quantity is

$$\frac{\partial \mathcal{L}}{\partial \dot{\phi}} = \frac{1}{2}\rho^2\dot{\phi} =: J$$

which is harder to interpret physically, but we see that it has something to do with angular momentum in the (ρ, ϕ)-plane. In any case, we set this quantity to be some constant J (so we're taking the moment map fiber above J, but the Lie group here is abelian so any moment map fiber admits action by the translation). At this point we have $2 \times 2 - 1 \times 2 = 1 \times 2$ degrees of freedom left, namely, we can eliminate $\dot{\phi}$ from the Lagrangian by rearranging for

$$\dot{\phi} = \frac{2J}{\rho^2}$$

and the Lagrangian becomes

$$\mathcal{L} = \frac{1}{4}\dot{\rho}^2 + \frac{J^2}{\rho^2} - V(|\rho|)$$

This is simple enough that we should calculate the equations of motion:

$$\ddot{\rho} = \frac{\partial}{\partial r}\left(\frac{2J^2}{\rho^2} - V\right)$$

Finally an exercise for the reader: **we haven't used time translation symmetry yet**. If we symplectically reduce by time symmetry then we should have 0 degrees of freedom left. That is, if I specify the energy, the number J, the plane of our trajectory, and the center of mass, there is exactly one trajectory that solves the equations of motion. Does this make sense, and how would you compute it?