Integrable systems

S. K. Devalapurkar

Lecture 1: The Lagrangian formalism

In a physics class, one tends to figure out the equations of motion for a particle with some forces acting on it by figuring out the degrees of freedom, and then setting up constraint equations coming from the conservation of energy. This often requires careful analysis of precise geometric aspects of the situation at hand. In this lecture, we will introduce the *Lagrangian* method, which was introduced by Lagrange (when he was 18 years old!) in a famous book called "Mécanique analytique".

We will consider the motion of a particle on a (smooth, always) manifold M. The manifold M can be taken to be $\mathbf{R}^n$, in which case the motion is "unconstrained". But one could also consider, say, motion on S^1, which might correspond to the motion of a pendulum (the position of the mass at the end of the pendulum is constrained to move on a circle). The sorts of constraints we will consider are called *holonomic*. (We will not dwell on this; all it means is that the particle cannot suddenly slip off of the manifold. For example, the pendulum's string is not allowed to snap so that the mass just flies off of the S^1 we've constrained it to live on.)

Let's say that $q : \mathbf{R} \to \mathbf{R}$ is the path of a particle moving on $\mathbf{R}$. Its derivative $\dot{q}$ can be used to construct the kinetic energy $K := \dot{q}^2/2$, and Newton's second law says that the force acting on this particle is given by $\ddot{q} = \dot{K}$. Let's say the particle moves from initial position $q(t_0)$ to final position $q(t_1)$. Then the work done by this particle is given by $\int_{t_0}^{t_1} F\dot{q}dt = \int_{q(t_0)}^{q(t_1)} Fdq$, and one asks that the net work done by this particle if the motion is a *closed* loop is zero. This means that $F = -\partial_q U(q)$ for some function $U(q)$, called the potential energy. In other words, $\dot{K} = -\partial_q U(q)$; this describes the equation of motion of q in the potential U. This can be rewritten using the function $L(q, \dot{q}) = K(\dot{q}) - U(q)$ as the equation $\frac{d}{dt}\partial_{\dot{q}} L = \partial_q U$; this is the Euler-Lagrange equation, and as we will now show, it can be obtained purely from a study of L.

To generalize this to an arbitrary smooth manifold M, let us view L as a function $L : TM \to \mathbf{R}$, where TM is the tangent bundle of M. It will actually be convenient to allow a little more leeway, where we allow L to depend on a time variable. That is, we allow L to be a function $TM \times \mathbf{R} \to \mathbf{R}$. Given a curve $q : \mathbf{R} \to M$, one obtains a tangent vector $\dot{q}(t)$ at the point $q(t) \in M$, which allows

Part of this work was done when the author was supported by NSF DGE-2140743.

us to define the *action functional*

$$S := \int_{t_0}^{t_1} L(q, \dot{q}, t) dt.$$

We will fix a metric $\langle -, - \rangle$ on M. Say that a path $q : \mathbf{R} \to M$ *extremizes* S if

$$\frac{d}{ds} S[q_s, \dot{q}_s] \Big|_{s=0} = 0$$

for all infinitesimal variations $q : [0, \epsilon) \times \mathbf{R} \to M$ of q.

Theorem 1 (Euler-Lagrange)**.** *The curve $q : \mathbf{R} \to M$ is an extremizer of S if and only if the **Euler-Lagrange** equations are satisfied:*

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0$$

for all tangent vectors $(q, \dot{q}) \in TM$.

Proof. Let us just sketch the argument when $M = \mathbf{R}$. If we vary q by $q \mapsto \delta q$, the variation of S is given by

$$\begin{aligned} \delta S &= \int_{t_0}^{t_1} \left(\frac{\partial L}{\partial \dot{q}} \delta \dot{q} + \frac{\partial L}{\partial q} \delta q \right) dt \\ &= \int_{t_0}^{t_1} \left(-\frac{d}{dt} \frac{\partial L}{\partial \dot{q}} + \frac{\partial L}{\partial q} \right) \delta q \; dt \end{aligned}$$

by integrating by parts. Now the "fundamental lemma of calculus of variations", which says that if a continuous function $f(x)$ satisfies $\int_a^b f(x) g(x) dx = 0$ for all compactly supported smooth functions $g(x)$ on (a, b), then $f(x) = 0$, implies the desired statement. (Try to prove this lemma by yourself.) $\square$

In a coordinate system on TM, this is just saying that

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} = \frac{\partial L}{\partial q^i}.$$

Let us illustrate this in an example.

Example 2 (Brachistochrone problem)**.** This is closely related to the original application Lagrange had in mind for his theory. The question is: can one describe the curve for which the time taken by a particle sliding without friction from point $(0, 0)$ to (x_0, y_0) is minimized? If the particle is in position (x, y), the instantaneous velocity of this particle must be $v = \sqrt{2gy}$, where g is the gravitational constant (this comes from conservation of energy). Therefore, since $v = \frac{ds}{dt}$, where $ds^2 = dx^2 + dy^2$ is the infinitesimal arc length, we find that $dt = \frac{ds}{\sqrt{2gy}}$. I will ignore the factor $\sqrt{2g}$. Writing $x = x(y)$ (so that $\dot{x}$ will denote $\frac{\partial x}{\partial y}$, and not the "time derivative" of x), we have $ds = \sqrt{\dot{x}^2 + 1} dy$, and so

$$\text{total time} = \int dt = \int_0^{y_0} \sqrt{\frac{\dot{x}^2 + 1}{y}} dy.$$

Let us take L to be the square root factor above, with $q = x$ and $\dot{q} = \dot{x}$. The Euler-Lagrange equations just say that

$$0 = \frac{\partial L}{\partial x} = \frac{d}{dy} \frac{\partial L}{\partial \dot{x}} = \frac{d}{dy} \frac{\dot{x}}{\sqrt{y} \sqrt{\dot{x}^2 + 1}}.$$

But this means that
$$\frac{\dot{x}}{\sqrt{y}\sqrt{\dot{x}^2+1}} = \frac{1}{\sqrt{y(1+\dot{x}^{-2})}}$$
is a constant, and simplifying (using that $\frac{\partial y}{\partial x} = \left(\frac{\partial x}{\partial y}\right)^{-1}$), we see that $\sqrt{y(1+\dot{y}^2)}$ is a constant. It is a nice exercise to try to solve the resulting differential equation to show that taking y to be a constant multiple of $\sin^2(\theta)$, one has
$$x(\theta) = a(2\theta - \sin(2\theta)),$$
$$y(\theta) = a(1 - \cos(2\theta)).$$
This is called a *cycloid*, and the constant a is essentially the arc-length of the resulting curve.

Example 3 (Geodesics)**.** If $L : TM \to \mathbf{R}$ is the map $(q, \dot{q}) \mapsto \langle \dot{q}, \dot{q}\rangle_q$, the Euler-Lagrange equations precisely describe the *geodesics* on M.

One important observation in Example 2 is that there is a natural conserved quantity (the arc-length), which came simply from the Lagrangian function L being independent of q. Of course, this is not special to the above example: it follows immediately from Theorem 1 that if L is independent of q, then
$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} = \frac{\partial L}{\partial q} = 0,$$
and so $\frac{\partial L}{\partial \dot{q}}$ is conserved. Note that when $L = \dot{q}^2/2$ is the Lagrangian for a free particle, $\frac{\partial L}{\partial \dot{q}} = \dot{q}$ is the *momentum* of the particle. In other words, translation invariance leads to momentum conservation.

This is a special case of a much more general and important result, called *Noether's theorem*, which says that any symmetry of the system (such as translation invariance) leads to a conserved quantity (such as linear momentum). For simplicity, we will assume that L is "time independent", i.e., that it is a map $TM \to \mathbf{R}$ (and not a map $TM \times \mathbf{R} \to \mathbf{R}$). First, what is a "symmetry"? (We will be very lax with analytic subtleties below.)

Definition 4. A 1-*parameter symmetry* of $L : TM \to \mathbf{R}$ is the data of a family $\{f_s : M \to M\}_{s\in\mathbf{R}}$ of smooth maps such that $f_s^*(L) = L$, where f_s^* is the map induced on functions by the map $f_{s,*} : TM \to TM$. Actually, we only need a weaker notion below, namely $\{f_s\}$ only needs to be an *infinitesimal symmetry*, i.e., the variation $\delta L = \frac{d}{ds} f_s^*(L)\big|_{s=0}$ must vanish. Given such a symmetry and any point $q \in M$, we obtain a vector $\delta q \in T_qM$ by the formula $\frac{d}{ds} f_s(q)\big|_{s=0}$.

In order to express the quantity conserved by a 1-parameter symmetry, it will be convenient to introduce the *Legendre transform*.

Definition 5. Let $\mathcal{E} \to M$ be a vector bundle over M with dual vector bundle $\mathcal{E}^*$, and say $L : \mathcal{E} \to \mathbf{R}$ is a smooth function. If $x \in M$, we obtain a function $L_x : \mathcal{E}_x \to \mathbf{R}$, and hence a map $dL_x : T\mathcal{E}_x \to \mathbf{R}$. Note that $\mathcal{E}_x$ is a vector space, so $T\mathcal{E}_x = \mathcal{E}_x \oplus \mathcal{E}_x$. Given a point $e \in \mathcal{E}_x$, we therefore obtain a *linear* map $dL_x(e) : \mathcal{E}_x \to \mathbf{R}$, i.e., a point $dL_x(e) \in \mathcal{E}_x^*$. In other words, the linear map $dL_x(e) : \mathcal{E}_x \to \mathbf{R}$ can be viewed as the map
$$\mathcal{E}_x \ni e' \mapsto \frac{d}{dt} L_x(e + te')\bigg|_{t=0} \in \mathbf{R}.$$

The *Legendre transform* of L is the map $\Phi_L : \mathcal{E} \to \mathcal{E}^*$ sending $(x, e) \mapsto (x, dL_x(e))$. It is convenient to simply think of this as the map $(x, e) \mapsto \frac{\partial L}{\partial e}(x, e)$.

Theorem 6 (Noether)**.** *If $\{f_s\}$ is a 1-parameter (infinitesimal) symmetry of $L : TM \to \mathbf{R}$, the function*

$$\mathcal{N} : TM \to \mathbf{R},\ (q, \dot{q}) \mapsto \left\langle \frac{\partial L}{\partial \dot{q}}, \delta q \right\rangle = \langle \Phi_{L_q}(v), \delta q \rangle$$

is conserved, in the sense that for any path $q : \mathbf{R} \to M$, one has $\frac{d}{dt}\mathcal{N}(q(t)) = 0$.

Proof. Again, let us just explicate what's going on when $M = \mathbf{R}$ (the general case is just notation-heavier). We are claiming that if $q : \mathbf{R} \to \mathbf{R}$ is a smooth map, the quantity

$$\mathcal{N}(q, \dot{q}) = \frac{\partial L}{\partial \dot{q}} \frac{d}{ds} f_s(q)\Big|_{s=0}$$

is conserved. This is easy to see. Write $\delta q = \frac{d}{ds} f_s(q)\big|_{s=0}$ for simplicity. Taking time derivatives, one gets

$$\begin{aligned} \frac{d}{dt}\mathcal{N} = \frac{d}{dt}\left(\frac{\partial L}{\partial q}\delta q\right) &= \frac{\partial L}{\partial \dot{q}}\delta \dot{q} + \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right)\delta q \\ &= \frac{\partial L}{\partial \dot{q}}\delta \dot{q} + \frac{\partial L}{\partial q}\delta q = \delta L, \end{aligned}$$

by the Theorem 1. But this vanishes by assumption on $\{f_s\}$. $\square$

Let us see some examples.

Example 7 (Linear momentum)**.** Say M is a vector space V, and fix $v \in V$. Suppose that the family $f_s : V \to V$ given by $q \mapsto q + sv$ is a 1-parameter symmetry of L. Then Theorem 6 says that $\langle \frac{\partial L}{\partial \dot{q}}, v \rangle$ is conserved. If one takes $L = \langle \dot{q}, \dot{q} \rangle / 2$ for some nondegenerate inner product $\langle -, - \rangle$ on V (so that we can identify $V = V^*$), then the conserved quantity is precisely the component $\langle \dot{q}, v \rangle$ of the momentum $\frac{\partial L}{\partial \dot{q}}$ in the v-direction.

Example 8 (Angular momentum)**.** Fix a quadratic vector space V, i.e., a vector space V equipped with a quadratic form. Suppose $A \in \mathfrak{so}(V)$, so that A defines a 1-parameter family $f_s : V \to V$ of rotations. For example, for the standard quadratic vector space $(\mathbf{R}^n, x_1^2 + \cdots, x_n^2)$, the element A is a skew-symmetric $n \times n$-matrix, and $f_s(q) = e^{sA}q$. Suppose that f_s is a symmetry of the Lagrangian $L : TV \to \mathbf{R}$ given by $L(q, \dot{q}) = \langle \dot{q}, \dot{q} \rangle / 2$. Then

$$\delta L = \frac{\partial L}{\partial \dot{q}}\delta \dot{q} = \langle \dot{q}, \delta \dot{q} \rangle.$$

But since $\frac{d}{ds}e^{sA}\big|_{s=0} = A$, we have

$$\delta \dot{q} = \frac{d}{ds} f_s(\dot{q})\Big|_{s=0} = \frac{d}{dt}Aq = A\dot{q}.$$

This means that

$$\delta L = \langle \dot{q}, A\dot{q} \rangle.$$

This is precisely the angular momentum! To be clear, if A is the skew-symmetric matrix with a 1 in the (i, j) entry and -1 in the (j, i) entry, then

$$\delta L = \langle \dot{q}, A\dot{q} \rangle = \dot{q}_i q^j - \dot{q}_j q^i,$$

which is how one computes the angular momentum. For instance, when $\dim(V) = 3$, we may identify $\mathfrak{so}_3 \cong \mathbf{R}^3$, and the above calculation shows that the map

$$(q, p) \mapsto (q, \dot{q}) \mapsto [A \mapsto \langle \dot{q}, A\dot{q} \rangle]$$

can be viewed as the cross product map

$$T^*\mathbf{R}^3 \cong T\mathbf{R}^3 \cong \mathbf{R}^3 \times \mathbf{R}^3 \to \mathbf{R}^3.$$

This is a special case of the *moment map* construction that we will describe later.

References

1 Oxford St, Cambridge, MA 02139
Email address: sdevalapurkar@math.harvard.edu, February 2, 2024