Note on the Legendre transform (for Math 242)

Michael Hutchings

September 2024

The purpose of this note is to explain the Legendre transform as I presented it in class, to reduce confusion. This is standard material presented in slightly different ways in numerous textbooks, including the books by Arnold, Cannas da Silva, Geiges, and McDuff-Salamon (Introduction to Symplectic Topology) recommended for the class.

1 Geodesic flow as a Hamiltonian flow

In this section we state a basic result about geodesics which we would like to understand, and which the Legendre transform will generalize. We begin by reviewing the definition of geodesic.

Let Q be a smooth n-dimensional manifold, and let g be a Riemannian metric on Q. A smooth path $\gamma : [a,b] \to Q$ is a **geodesic** if it is a critical point of the **energy** functional

$$E(\gamma) = \int_a^b \frac{1}{2}\|\gamma'(t)\|^2 dt$$

on the space of paths $[a,b] \to Q$ with the same endpoints. Concretely, for γ to be a critical point means that if

$$\Gamma : (-\varepsilon, \varepsilon) \times [a,b] \longrightarrow Q$$

is a smooth map with $\Gamma(0,\cdot) = \gamma$ and $\Gamma(s,a) = \gamma(a)$ and $\Gamma(s,b) = \gamma(b)$ for all $s \in (-\varepsilon, \varepsilon)$, then

$$\left.\frac{d}{dt}\right|_{s=0} E(\Gamma(s,\cdot)) = 0.$$

Any constant-speed length-minimizing path between two points in Q is a geodesic. Conversely, any geodesic has constant speed, and, although we will not prove this here, is length-minizing if restricted to a sufficiently small segment.

The above definition of geodesic is equivalent to a second order ODE for γ (we will prove a more general statement below). In particular, there is a vector field

X_g on TQ such that a path $\widetilde{\gamma} : [a,b] \to TQ$ is a flow line of X_g if and only if $\widetilde{\gamma}(t) = (\gamma(t), \gamma'(t))$ where $\gamma : [a,b] \to Q$ is a geodesic.

We would now like to understand geodesics symplectically. Let λ denote the canonical 1-form on the cotangent bundle T^*Q, and consider the symplectic form on T^*Q defined by $\omega = -d\lambda$. The metric g induces an isomorphism of vector bundles $\phi : TQ \xrightarrow{\simeq} T^*Q$, and thereby an inner product on T^*Q. Let $Y \subset T^*Q$ denote the unit cotangent bundle:

$$Y = \{(q,p) \mid q \in Q,\ p \in T_q^*Q,\ \|p\|^2 = 1\}.$$

This is a level set of the Hamiltonian $H : T^*Q \to \mathbb{R}$ defined by

$$H(q,p) = \frac{1}{2}\|p\|^2.$$

Let X_H denote the associated Hamiltonian vector field on T^*Q defined by

$$\omega(X_H, \cdot) = dH.$$

Since Y is a level set of H, it follows that X_H is tangent to Y. We would like to prove:

Proposition 1.1. *The Hamiltonian vector field X_H on Y agrees with the geodesic flow X_g on the unit tangent bundle under ϕ.*

In particular, a unit speed geodesic in Q is equivalent to a trajectory of the Hamiltonian vector field X_H on Y.

2 Lagrangian mechanics on a manifold

We will prove Proposition 1.1 as a special case of the relation between Lagrangian mechanics and Hamiltonian mechanics.

The Lagrangian formulation of mechanics is as follows. Let Q be a smooth manifold. One can think of a point in Q as representing the position of a particle or a mechanical system. We begin with a "Lagrangian" function

$$L : TQ \longrightarrow \mathbb{R}.$$

(There is also a time-dependent version in which L is a function on $\mathbb{R} \times TQ$, but we will stick with the time-independent case for simplicity.) Given a Lagrangian L, for a smooth path $\gamma : [a,b] \to \mathbb{R}$, we define the energy

$$E(\gamma) = \int_a^b L(\gamma(t), \gamma'(t))dt.$$

One then postulates that the laws of motion are that the path γ represents the time evolution of a particle or system if and only if γ is a critical point of the energy on the space of paths $[a,b] \to Q$ with the same endpoints.

This statement of the laws of motion is not very transparent, so we now translate it into an ordinary differential equation. Consider a path $\gamma : [a,b] \to \mathbb{R}$, and assume (after shrinking the domain if necessary) that it maps to a single coordinate chart on Q with coordinates $q_1, \ldots, q_n$. We have induced coordinates $q_1, \ldots, q_n, V_1, \ldots, V_n$ on the tangent bundle of the coordinate chart, where we write a tangent vector as $\sum_{i=1}^n V_i \frac{\partial}{\partial q_i}$.

Lemma 2.1. *A path $\gamma : [a,b] \to \mathbb{R}$ as above is a critical point of the energy functional on the space of paths $[a,b] \to Q$ with the same endpoints if and only if*

$$\frac{\partial L}{\partial q_i}(\gamma(t), \gamma'(t)) = \frac{d}{dt}\frac{\partial L}{\partial V_i}(\gamma(t), \gamma'(t)) \tag{2.1}$$

for $i = 1, \ldots, n$.

Proof. Let $\Gamma : (-\varepsilon, \varepsilon) \times [a,b] \to Q$ be a variation of the path γ with fixed endpoints as considered previously. For $t \in [a,b]$, define

$$\delta(t) = \left.\frac{\partial}{\partial s}\right|_{(0,t)} \Gamma(s,t) \in T_{\gamma(t)}Q.$$

In our local coordinates, write $\gamma(t) = (q_1(t), \ldots, q_n(t))$, and let $\delta_i(t)$ denote the V_i coordinate of $\delta(t)$. We now compute

$$\begin{aligned}
\left.\frac{d}{ds}\right|_{s=0} E(\Gamma(s,\cdot)) &= \left.\frac{d}{ds}\right|_{s=0} \int_a^b L\left(\Gamma(s,t), \frac{\partial}{\partial t}\Gamma(s,t)\right) dt \\
&= \int_a^b \left.\frac{\partial}{\partial s}\right|_{(0,t)} L\left(\Gamma(s,t), \frac{\partial}{\partial t}\Gamma(s,t)\right) dt \\
&= \int_a^b \sum_{i=1}^n \left(\frac{\partial L}{\partial q_i}(\gamma(t), \gamma'(t))\delta_i(t) + \frac{\partial L}{\partial V_i}(\gamma(t), \gamma'(t))\frac{d\delta_i(t)}{dt}\right) dt \\
&= \int_a^b \sum_{i=1}^n \left(\frac{\partial L}{\partial q_i}(\gamma(t), \gamma'(t)) - \frac{d}{dt}\frac{\partial L}{\partial V_i}(\gamma(t), \gamma'(t))\right) \delta_i(t)\, dt.
\end{aligned}$$

Here in the last line we have used integration by parts and the assumption that our variation of paths has fixed endpoints. Now γ is a critical point of the energy if and only if the above expression vanishes for any functions $\delta_i(t)$ vanishing at the endpoints, which means that the equations (2.1) hold. □

So to summarize, the equations of motion are that (2.1) holds and that

$$\frac{dq_i(t)}{dt} = V_i(t). \tag{2.2}$$

The equations (2.1) are called the **Euler-Lagrange equations**.

This still is not a very satisfactory statement of the equations of motion, because to get an ODE we want equations for $dV_i(t)/dt$. However under suitable convexity assumptions on the Lagrangian which we will state below, the numbers $\partial L/\partial V_i$ determine the numbers V_i, so that the Euler-Lagrange equations (2.1) do in fact determine $dV_i(t)/dt$.

For example, consider the motion of a unit mass particle on $Q = \mathbb{R}^n$ where the potential energy is $U : \mathbb{R}^n \to \mathbb{R}$. We define the Lagrangian

$$L(q, V) = \frac{1}{2}\|v\|^2 - U(q),$$

i.e. the kinetic energy minus the potential energy. The Euler-Lagrange equations (2.1) in this example are

$$\frac{dV_i}{dt} = -\frac{\partial U}{\partial q_i}.$$

3 The Legendre transform

We now explain the relation between the Lagrangian and Hamiltonian formulations of mechanics.

Recall that if $q_1, \ldots, q_n$ are local coordinates on a neighborhood Q, and if $q_1, \ldots, q_n, p_1, \ldots, p_n$ are the induced coordinates on the cotangent bundle of this neighborhood, then the symplectic form ω that we are using is given by

$$\omega = \sum_{i=1}^{n} dq_i\, dp_i.$$

If $H : T^*Q \to \mathbb{R}$ is a Hamiltonian, then the associated Hamiltonian vector field is

$$X_H = \sum_{i=1}^{n} \left(\frac{\partial H}{\partial p_i} \frac{\partial}{\partial q_i} - \frac{\partial H}{\partial q_i} \frac{\partial}{\partial p_i} \right).$$

Thus a trajectory of X_H locally satisfies the equations

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i}, \tag{3.1}$$

$$\frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i}. \tag{3.2}$$

Now let $L : TQ \to \mathbb{R}$ be a Lagrangian. The fiberwise derivative of L defines a map

$$\phi : TQ \longrightarrow T^*Q. \tag{3.3}$$

To be precise, if $q \in Q$ and $V \in T_qQ$, then we define a cotangent vector $\phi(q,V) \in T_q^*Q$ as follows: If $W \in T_qQ$, then

$$\phi(q,V)(W) = \frac{d}{dt}\bigg|_{t=0} L(q, V + tW).$$

For example, if g is a Riemannian metric on Q and $L(q,V) = \frac{1}{2}\|V\|^2$, then $\phi : TQ \to T^*Q$ is the isomorphism of vector bundles induced by g as before. For a more general Lagrangian, ϕ still maps fibers to fibers, but in general not linearly.

To continue, we now assume that the Lagrangian L has the property that:

(*) The fiberwise derivative map (3.3) is a diffeomorphism.

This holds for example if for each $q \in Q$, if $V_1, \ldots, V_n$ are linear coordinates on the tangent space T_qQ, then the matrix

$$\left(\frac{\partial^2 L}{\partial V_i \partial V_j}\right)_{i,j=1,\ldots,n}$$

is positive definite, and

$$\lim_{\|V\| \to \infty} \frac{L(q,V)}{\|V\|} = +\infty.$$

Given a Lagrangian L satisfying (*), we define its **Legendre transform** to be the function

$$H : T^*Q \longrightarrow \mathbb{R}$$

defined as follows. If $q \in Q$ and $p \in T_q^*Q$, then

$$H(q,p) = p(\phi^{-1}(p)) - L(q, \phi^{-1}(p)). \tag{3.4}$$

In a coordinate chart on Q with coordinates $q_1, \ldots, q_n$, if $V_1, \ldots, V_n$ are the induced coordinates for tangent vectors in the chart, and $p_1, \ldots, p_n$ are the induced coordinates for cotangent vectors in the chart, then we have

$$H(q_1, \ldots, q_n, p_1, \ldots, p_n) = \sum_{i=1}^n p_i V_i - L(q_1, \ldots, q_n, V_1, \ldots, V_n), \tag{3.5}$$

where we are using the assumption (*) to regard $V_1, \ldots, V_n$ as functions of $q_1, \ldots, q_n$ and $p_1, \ldots, p_n$.

Theorem 3.1. *Assume that the Lagrangian $L : TQ \to \mathbb{R}$ satisfies (*), and define $H : T^*Q \to \mathbb{R}$ by* (3.4)*. Then a path $\widetilde{\gamma} : [a,b] \to TQ$ satisfies the Lagrangian equations of motion* (2.1) *and* (2.2) *if and only if the path $\phi \circ \widetilde{\gamma} : [a,b] \to T^*Q$ satisfies the Hamiltonian equations of motion* (3.1) *and* (3.2)*.*

Remark 3.2. Before proving the theorem, and to clarify its statement, we need to clear up a notational confusion. Namely, $\partial/\partial q_i$ has different meanings in equations (2.1) and (3.2)! Recall that a partial derivative with respect to a coordinate is only defined relative to all of the other coordinates. Our convention is that $\partial L/\partial q_i$ denotes the change in L as we vary q_i while fixing q_j for $j \neq i$ and $V_1, \ldots, V_n$, while $\partial H/\partial q_i$ denotes the change in H as we vary q_i while fixing q_j for $j \neq i$ and $p_1, \ldots, p_n$.

Example 3.3. If g is a Riemannian metric on Q and $L(q, V) = \frac{1}{2}\|V\|^2$, then $H(q, p) = \frac{1}{2}\|p\|^2$, and Theorem 3.1 recovers Proposition 1.1.

Proof of Theorem 3.1. In local coordinates as in (3.5), the fiberwise derivative map ϕ defines p_i as a function of $q_1, \ldots, q_n, V_1, \ldots, V_n$ by the equation

$$p_i = \frac{\partial L}{\partial V_i}. \tag{3.6}$$

We now relate the partial derivatives of H and L. To reduce confusion as in Remark 3.2, it is easier to first take the total derivative of both sides of (3.5). This gives

$$\begin{aligned}
\sum_{i=1}^{n}\left(\frac{\partial H}{\partial q_i}dq_i + \frac{\partial H}{\partial p_i}dp_i\right) &= \sum_{i=1}^{n}\left(p_i dV_i + V_i dp_i - \frac{\partial L}{\partial q_i}dq_i - \frac{\partial L}{\partial V_i}dV_i\right) \\
&= \sum_{i=1}^{n}\left(V_i dp_i - \frac{\partial L}{\partial q_i}dq_i\right)
\end{aligned}$$

where in the second line we have used (3.6). Thus we have

$$\frac{\partial H}{\partial q_i} = -\frac{\partial L}{\partial q_i}, \tag{3.7}$$

$$\frac{\partial H}{\partial p_i} = V_i. \tag{3.8}$$

By equations (3.6) and (3.7), the Euler-Lagrange equation (2.1) is equivalent to the Hamiltonian equation (3.2). By equation (3.8), the other equation in the Lagrangian equations of motion, namely (2.2), is equivalent to the other Hamiltonian equation (3.1). □