1 The Legendre Transformation

1.1 Definition

Consider a smooth real-valued function f on $\mathbb{R}$ that is strictly convex, i.e. $\frac{d^2 f}{dx^2} > 0$ for all x. The Legendre transformation is the operation that assigns to any pair $(x, f(x))$ the new pair $(p, F(p))$ according to

$$F(p) = \max_{x \in \mathbb{R}} [px - f(x)]$$

Since $-\frac{d^2 f}{dx^2} < 0$ for all x, a necessary and sufficient condition for p to exist and be unique is

$$p = f'(x) \tag{1}$$

In other words, the Legendre transform can be written as

$$(x, f(x)) \mapsto \begin{cases} F(p) = px - f(x) \\ \qquad\quad p = f'(x) \end{cases} \tag{2}$$

The top equality on the right-hand side has to be viewed as a function of p only, which implies that the relation $p = f'(x)$ has to be inverted to find x as a function of p. This is what makes the Legendre transformation a nonlinear transformation, and what sometimes makes the transformation difficult to evaluate in practice.

Here is an example for which the Legendre transformation is easily calculated. Consider the function $f(x) = e^x$. The Legendre transformation of f is given by

$$\begin{cases} F(p) = px - e^x \\ p = e^x \end{cases} \tag{3}$$

The second equation is easily inverted for $p > 0$: $x = \ln p$. Plugging this into the first equation, we have

$$F(p) = p(\ln p - 1)$$

Note that F, defined for $p > 0$, is convex. This is not an artifact of the example we picked, but in fact true in general, as we will prove next.

1.2 Derivatives and Convexity

By taking a total differential of the first equation in (2)

$$dF = pdx + xdp - f'(x)dx$$

and plugging the second equality, $p = f'(x)$ in the relation above, we get the important equation

$$\frac{dF}{dp} = F'(p) = x \tag{4}$$

The second derivative of F therefore is

$$\frac{dF'(p)}{dp} = \frac{dx}{dp} = \left(\frac{dp}{dx}\right)^{-1} = \left(\frac{df'(x)}{dx}\right)^{-1} > 0$$

We conclude that the Legendre transformation preserves convexity. We can therefore apply the Legendre transformation to F. Let us see what we obtain.

1.3 Involution

Let us consider the Legendre transform pair $(p, F(p))$ of the function f. We have

$$f(x) = px - F(p) = pF'(p) - F(p)$$

Since F is strictly convex, by the same arguments as we used in Section 1.1, the expression above is equivalent to

$$f(x) = \max_p [xp - F(p)]$$

and $x = F'(p)$ for the expression of p in terms of x where the maximum is reached. In other words, the Legendre transform is its own inverse. We say that it is an *involution*.

1.4 Multivariable case

Consider a function f of two variables whose matrix of second derivatives is positive definite, so that f is globally strictly convex. The Legendre transformation is defined by

$$F(p_1, p_2) = \max_{p_1, p_2}[p_1 x_1 + p_2 x_2 - f(x_1, x_2)] \tag{5}$$

It is quite easy to show that the definition (5) is equivalent to

$$\begin{cases} F(p_1, p_2) = p_1 x_1 + p_2 x_2 - f(x_1, x_2) \\ p_1 = \partial_1 f(x_1, x_2) \\ p_2 = \partial_2 f(x_1, x_2) \end{cases} \tag{6}$$

Furthermore, all the properties proved above naturally extend to the multidimensional case: preservation of convexity, involution, and

$$\partial_1 F(p_1, p_2) = x_1 \qquad \partial_2 F(p_1, p_2) = x_2$$

1.5 Legendre transformation with passive arguments

Consider a function $f(x, y)$ of two variables and its Legendre transform with respect to the first slot:

$$F(p, y) = \max_x [px - f(x, y)]$$

The expression above can be reexpressed as the following coupled equations, which are more easily to work with:

$$\begin{cases} F(p, y) = px - f(x, y) \\ p = \partial_1 f(x, y) \end{cases} \tag{7}$$

We now write two expressions for the total differential of F. First, by definition, we have

$$dF(p, y) = \partial_1 F(p, y)dp + \partial_2 F(p, y)dy$$

On the other hand, from Equation (7), we can also say

$$dF(p, y) = \partial_1 f(x, y)dx + xdp - \partial_1 f(x, y)dx - \partial_2 f(x, y)dy = xdp - \partial_2 f(x, y)dy$$

Comparing the two expressions, we get the following relations. First,

$$x = \partial_1 F(p, y)$$

This can be seen as the natural extension of the relation $x = F'(p)$ we derived in Section 1.2. The other relation is

$$\partial_2 F(p, y) = -\partial_2 f(x, y) \tag{8}$$

Note the - sign in Equation (8)

1.6 Local Legendre transformation

Consider a smooth function f and a point x in the domain of f. One can locally define a pair $(p, F(p))$ that can be seen as the local Legendre transformation of $(x, f(x))$ through the local mapping

$$\begin{cases} F(p) = px - f(x) \\ p = f'(x) \end{cases}$$

The local transformation can then be extended in the neighborhood of x by letting $x \to x + dx$ in the expressions above. We find

$$dp = f''(x)dx \qquad dF = F'(p)dp = xf''(x)dx$$

As long as $f''(x) \neq 0$ (it can be positive *or* negative), this procedure can be used to construct the Legendre transform increment by increment. If a function is globally strictly convex, then this construction agrees with the global definition of the Legendre transformation given previously. However, the construction fails as soon as $f''(x)$ goes through 0 since $F(p)$ then becomes multivalued.

2 Hamiltonian Formulation of Mechanics

2.1 Hamilton's equations

Consider a one-dimensional Lagrangian $L(q, \dot{q}, t)$. The Hamiltonian function $H(q, p, t)$ is related to the Lagrangian L through the Legendre transformation with respect to the second slot in L:

$$H(q, p, t) = \max_{\dot{q}} [p\dot{q} - L(q, \dot{q}, t)]$$

which can be rewritten, as we know,

$$\begin{cases} H(q, p, t) = p\dot{q} - L(q, \dot{q}, t) \\ p = \partial_2 L(q, \dot{q}, t) \end{cases}$$

Note that the Hamiltonian H has the same value as the energy of the system, but the Hamiltonian depends on (q, p, t), while the energy depends on $(q, \dot{q}, t)$. The relationships above hold for any path, whether it is physically realizable or not. According to Equation (4), we can then also write

$$\dot{q} = \partial_2 H(q, p, t) \tag{9}$$

Along a physically realizable path, we can derive an additional relation. Indeed, following Section 1.5, we know that

$$\partial_1 H(q, p, t) = -\partial_1 L(q, \dot{q}, t)$$

The E-L equation for q can thus be rewritten as

$$\frac{d}{dt}(\partial_2 L(q, \dot{q}, t)) = \frac{dp}{dt} = \partial_1 L(q, \dot{q}, t) = -\partial_1 H(q, p, t)$$

In summary, in the Hamiltonian formulation, the dynamics of the system is given by the following coupled set of differential equations for $q(t)$ and $p(t)$:

$$\begin{cases} \dfrac{dq}{dt} = \dfrac{\partial H}{\partial p}(q, p, t) \\ \dfrac{dp}{dt} = -\dfrac{\partial H}{\partial q}(q, p, t) \end{cases} \tag{10}$$

These are two first-order ODEs that have to be solved as an initial-value problem by specifying $q(0)$ and $p(0)$. The first equation in the system is a statement of the relationship between the momenta and the velocities in terms of the Hamiltonian, and holds for any path. The second equation in the system only holds for realizable paths and is the one that contains the information on the dynamics of the system.

The system of equations (10) is known as *Hamilton's equations*. The equations are also sometimes referred to as *canonical equations*. In the remainder of this lecture and in the coming lectures, we will see why and in which situations the Hamiltonian formulation of mechanics is particularly convenient.

Our little study of the Legendre transformation also tells us that the following holds:

$$\partial_3 H(q, p, t) = -\partial_3 L(q, \dot{q}, t)$$

This means that if the Lagrangian does not have an explicit time dependence, then the Hamiltonian does not have an explicit time dependence either. In Lecture 3, we saw that the energy of a system whose Lagrangian does not depend explicitly on time is conserved. We conclude that *if the Hamiltonian does not depend explicitly on time, it is a conserved quantity* along the physically realizable path(s): $dH/dt(q, p, t) = 0$.

2.2 Hamiltonian action principle

The system of equations (10) are the E-L equations of an action principle called the *canonical action principle* or *Hamiltonian action principle*. The action principle states that the physically realizable path(s) is such that the following functional

$$\mathcal{J}[q, p] = \int_0^T \left[p\frac{dq}{dt} - H(q, p, t)\right] dt$$

is extremal with respect to independent variations of $q(t)$ and $p(t)$ such that the end points $q(0)$ and $q(T)$ are held fixed. To see this, let us take variations $q \to q + \delta q$, $p \to p + \delta p$ and evaluate $\delta \mathcal{J}$:

$$\begin{aligned}
\delta \mathcal{J} &= \int_0^T \left[\frac{dq}{dt} - \partial_2 H(q,p,t)\right] \delta p + \int_0^T \left[p\frac{d\delta q}{dt} - \partial_1 H(q,p,t)\delta q\right] \\
&= \int_0^T \left[\frac{dq}{dt} - \partial_2 H(q,p,t)\right] \delta p - \int_0^T \left[\frac{dp}{dt} + \partial_1 H(q,p,t)\right] \delta q + [p\delta q]_0^T \\
&= \int_0^T \left[\frac{dq}{dt} - \partial_2 H(q,p,t)\right] \delta p - \int_0^T \left[\frac{dp}{dt} + \partial_1 H(q,p,t)\right] \delta q
\end{aligned}$$

Using the usual arguments, we see that $\delta \mathcal{J} = 0$ if and only if

$$\frac{dq}{dt} = \partial_2 H(q,p,t) \qquad \frac{dp}{dt} = -\partial_1 H(q,p,t)$$

As desired, we recover Hamilton's equations.

2.3 Examples

2.3.1 The simple pendulum

We saw in Lecture 2 that a Lagrangian for the simple pendulum is

$$L(\theta, \dot{\theta}, t) = \frac{1}{2} l^2 \dot{\theta}^2 + gl\cos\theta$$

The Hamiltonian for this system is given by

$$\begin{aligned}
& H(\theta, p_\theta, t) = p_\theta \dot{\theta} - L(\theta, \dot{\theta}, t) \qquad p = \partial_2 L(\theta, \dot{\theta}, t) \\
\Leftrightarrow\ & H(\theta, p_\theta, t) = p_\theta \dot{\theta} - \frac{1}{2} l^2 \dot{\theta}^2 - gl\cos\theta \qquad \dot{\theta} = \frac{p_\theta}{l^2} \\
\Leftrightarrow\ & H(\theta, p_\theta, t) = \frac{1}{2}\frac{p_\theta^2}{l^2} - gl\cos\theta
\end{aligned}$$

We see that the Hamiltonian function does not depend explicitly on the time variable, so energy is conserved and we can write

$$H(\theta, p_\theta, t) = \frac{1}{2}\frac{p_\theta^2}{l^2} - gl\cos\theta = Cst = E \tag{11}$$

In some sense, this equation is sufficient to understand the important features of the dynamics of the simple pendulum, as we have already seen in Lecture 1, and discuss in more detail in future lectures. Still, if one insists on deriving the differential equation for the time evolution of θ, it is given by the canonical equations

$$\begin{cases} \dfrac{d\theta}{dt} = \dfrac{\partial H}{\partial p_\theta} \\ \dfrac{dp_\theta}{dt} = -\dfrac{\partial H}{\partial \theta} \end{cases} \quad \Leftrightarrow \quad \begin{cases} \dfrac{d\theta}{dt} = \dfrac{p_\theta}{l^2} \\ \dfrac{dp_\theta}{dt} = -gl\sin\theta \end{cases} \Rightarrow \frac{d^2\theta}{dt^2} = -\frac{g}{l}\sin\theta$$

This is the expected equation.

2.3.2 Point mass in a central potential

In Lecture 2, we also studied the case of a point mass m in the gravitational potential $-GmM/r$. A Lagrangian for the system was

$$L(r, \theta, \dot{r}, \dot{\theta}) = \frac{m}{2}\left(\dot{r}^2 + r^2\dot{\theta}^2\right) + \frac{GmM}{r}$$

The corresponding Hamiltonian is

$$
\begin{aligned}
&H(r,\theta,p_r,p_\theta) = p_r\dot r + p_\theta\dot\theta - L(r,\theta,\dot r,\dot\theta) \qquad p_r = \partial_3 L(r,\theta,\dot r,\dot\theta) \qquad p_\theta = \partial_4 L(r,\theta,\dot r,\dot\theta)\\
\Leftrightarrow &H(r,\theta,p_r,p_\theta) = p_r\dot r + p_\theta\dot\theta - L(r,\theta,\dot r,\dot\theta) \qquad \dot r = \frac{p_r}{m} \qquad \dot\theta = \frac{p_\theta}{mr^2}\\
\Leftrightarrow &H(r,\theta,p_r,p_\theta) = \frac{1}{2}\frac{p_r^2}{m} + \frac{1}{2}\frac{p_\theta^2}{mr^2} + \frac{GmM}{r}
\end{aligned}
$$

The time variable does not appear explicitly in H, so energy is conserved:

$$
\frac{1}{2}\frac{p_r^2}{m} + \frac{1}{2}\frac{p_\theta^2}{mr^2} - \frac{GmM}{r} = Cst = E \tag{12}
$$

Let us write the canonical equations for (r,p_r) and (θ,p_θ):

$$
\begin{cases}
\dfrac{dr}{dt} = \dfrac{\partial H}{\partial p_r} = \dfrac{p_r}{m}\\
\dfrac{dp_r}{dt} = -\dfrac{\partial H}{\partial r} = \dfrac{p_\theta^2}{mr^3} - \dfrac{GmM}{r^2}
\end{cases}
\qquad
\begin{cases}
\dfrac{d\theta}{dt} = \dfrac{\partial H}{\partial p_\theta} = \dfrac{p_\theta}{mr^2}\\
\dfrac{dp_\theta}{dt} = -\dfrac{\partial H}{\partial \theta} = 0
\end{cases}
$$

We see that p_θ is conserved, as a direct consequence of the absence of the variable θ in the Hamiltonian. p_θ is the magnitude Γ of the angular momentum introduced in Lecture 2 (to within the mass factor m). In that lecture, we had already proven that Γ was conserved, but we see that in the Hamiltonian formulation, the conservation laws tend to be explicit, through the formulation of the canonical equations themselves.

It is interesting to observe that since p_θ is conserved, the system on the left, for (r,p_r) decouples entirely from the system on the right. The conservation of p_θ allowed us to reduce the analysis to that of a one-dimensional problem. We will go back to this in future lectures.

3 Poisson bracket

3.1 Definition

Consider two functions $f(q,p,t)$ and $g(q,p,t)$. The Poisson bracket of f and g, written $\{f,g\}$ is defined by

$$
\{f,g\} = \partial_1 f\partial_2 g - \partial_2 f\partial_1 g \tag{13}
$$

The definition can be extended to functions $f(q_1,\ldots,q_N,p_1,\ldots,p_N,t)$ and $g(q_1,\ldots,q_N,p_1,\ldots,p_N,t)$ of $2N+1$ variables as follows

$$
\{f,g\} = \sum_{i=1}^{N}\left(\partial_i f\partial_{N+i} g - \partial_{N+i} f\partial_i g\right)
$$

In the following, we will always consider the case where f and g only depend on three variables, $f(q,p,t)$ and $g(q,p,t)$, for the simplicity of the notation. However, all the properties we will present hold for the higher dimensional case as well.

3.2 Properties of the Poisson bracket

A first trivial yet important property of the Poisson bracket is that for any function $f(q,p,t)$, we have

$$
\{f,f\} = 0
$$

If $g(q,p,t)$ is another function, then it is also easy to see that the Poisson bracket is *antisymmetric*:

$$
\{f,g\} = -\{g,f\}
$$

It is clear as well that the Poisson bracket is *bilinear*. If $h(q,p,t)$ is a third function, and c a real constant, we have

$$
\{f,g+h\} = \{f,g\} + \{f,h\} \qquad \{f+g,h\} = \{f,h\} + \{g,h\} \qquad \{f,cg\} = c\{f,g\} = \{cf,g\}
$$

We will need yet another property of the Poisson bracket, known as *Jacobi's identity*, which is more tedious to prove. Jacobi's identity says that

$$
\{f,\{g,h\}\} + \{h,\{f,g\}\} + \{g,\{h,f\}\} = 0 \tag{14}
$$

3.3 Hamiltonian mechanics written with Poisson brackets

Time evolution

The time evolution of any function $f(q(t), p(t), t)$ along a physically realizable path takes a simple form when written with a Poisson bracket:

$$\begin{aligned}\frac{df}{dt} &= \partial_1 f \frac{dq}{dt} + \partial_2 f \frac{dq}{dt} + \partial_3 f \\ &= \partial_1 f \partial_2 H - \partial_2 f \partial_1 H + \partial_3 f \\ &= \{f, H\} + \partial_3 f \end{aligned} \tag{15}$$

Hamilton's equations

Hamilton's equations take an elegant, uniform form when expressed in terms of Poisson brackets. Taking alternatively $f(q, p, t) = q$ and $f(q, p, t)$, the canonical equations can be written as

$$\begin{cases} \dfrac{dq}{dt} = \{q, H\} \\ \dfrac{dp}{dt} = \{p, H\} \end{cases}$$

Energy conservation

Equation (15) can also be used to provide a simple proof that energy is conserved when the Hamiltonian does not depend explicitly on time. Indeed, along any physically realizable path, we have $E = H$ (in value), so

$$\frac{dE}{dt} = \frac{dH}{dt} = \{H, H\} + \partial_3 H = \partial_3 H$$

since $\{f, f\} = 0$ for any f. We see that $dE/dt = 0$ if $\partial_3 H = 0$.

Conserved quantities

Let f and g be two conserved quantities that do not depend on time explicitly. According to Equation (15), we have

$$\{f, H\} = 0 \qquad \{g, H\} = 0$$

Jacobi's identity then implies that

$$\{\{f, g\}, H\} = 0$$

and since $\{f, g\}$ does not depend explicitly depend on time,

$$\frac{d}{dt}\left[\{f, g\}\right] = 0$$

In other words, the Poisson bracket of two conserved quantities is also a conserved quantity.