Hamiltonian for Central Potentials

Lecture 4

Physics 411
Classical Mechanics II

September 5th 2007

We turn now to Hamilton's formulation of the equations of motion. This discussion parallels our Lagrange studies, but as in strict classical mechanics, new avenues of discovery are open to us using the Hamiltonian. In particular, there is almost no better place to discuss invariance and conservation, a beautiful correspondence that is incredibly important to GR.

As you probably know already, general relativity can be viewed as a "theory without forces", and here the Hamiltonian plays an even more interesting role, because it is numerically identical to the Lagrangian. So it behooves us to use as much of one or the other approach as we find useful. In order to set the stage, we re-derive and solve the equations of motion for the Keplerian ellipse. I hope the clothing is different enough to hold your interest.

4.1 Legendre Transform

To develop the Hamiltonian form and the generators themselves, we need the notion of a Legendre transformation. In general, this is a special transformation that allows us to replace variables in a function in a consistent manner. We'll start with the definition and some examples in one dimension.

4.1.1 One Dimensional Legendre Transformation

Consider an arbitrary function of x: $f(x)$. We know that locally, the slope of this curve is precisely its derivative w.r.t. x, so the change in the function $f(x)$ at the point x for a small change in the argument, dx is

$$df = \frac{df}{dx}\, dx \equiv p(x)\, dx \qquad p(x) \equiv \frac{df(x)}{dx} \tag{4.1}$$

as usual. Now suppose we want to find a function that reverses the roles of the slope and infinitesimal, i.e. a function $g(p)$ such that $dg = x\,dp$ (where we now view x as a function of p defined by the inverse of $p = f'(x)$). It is easy to see that the function:

$$\boxed{g(p) = p\,x(p) - f(p),} \tag{4.2}$$

called the Legendre Transform, has

$$dg = x\,dp + p\,dx - df = x\,dp, \tag{4.3}$$

as desired. Notice that since $g(p)$ is a function of p only, we must have $x = \frac{dg}{dp}$ just as we had $p = \frac{df}{dx}$ before. So we have a pair of functions $f(x)$ and $g(p)$ related in the following way:

$$\begin{aligned} f(x) &\longrightarrow g(p) = x(p)\,p - f(x(p)) \quad x(p) : \frac{df(x)}{dx} = p \\ g(p) &\longrightarrow f(x) = p(x)\,x - g(p(x)) \quad p(x) : \frac{dg(p)}{dp} = x, \end{aligned} \tag{4.4}$$

where the pair $f(x)$ and $g(p)$ are Legendre transforms *of each other*. There is a nice symmetry here, the same transformation takes us back and forth. Let's look at an example to see how this works out.

Consider the function $f(x) = \alpha x^m$ for integer m. We can define p from the derivative of $f(x)$ as prescribed above:

$$\frac{df}{dx} = \alpha\,m\,x^{m-1} = p \longrightarrow x(p) = \left(\frac{p}{\alpha\,m}\right)^{\frac{1}{m-1}}. \tag{4.5}$$

With this assignment, we can construct $f(x(p)) \equiv f(p)$ – it is

$$f(p) = \alpha\left(\frac{p}{\alpha\,m}\right)^{\frac{m}{m-1}}, \tag{4.6}$$

and the Legendre transform is:

$$\begin{aligned} g(p) = p\,x(p) - f(p) &= \left(\frac{1}{\alpha\,m}\right)^{\frac{1}{m-1}} p^{\frac{m}{m-1}} - \alpha\left(\frac{p}{\alpha\,m}\right)^{\frac{m}{m-1}} \\ &= p^{\frac{m}{m-1}}\left[\left(\frac{1}{\alpha\,m}\right)^{\frac{1}{m-1}} - \alpha\left(\frac{1}{\alpha\,m}\right)^{\frac{m}{m-1}}\right]. \end{aligned} \tag{4.7}$$

Now take the reverse point of view – suppose we were given $g(p)$ as above, if we define $x = \frac{dg}{dp}$, and form the Legendre transform of $g(p)$: $h(x) = p(x)\,x - g(x)$ (with $g(x) \equiv g(p(x))$), then we would find that $h(x) = \alpha\,x^m$, precisely our starting point.

The polynomial example is a good place to see the geometrical implications of the Legendre transformation. Remember the goal, unmotivated at this point: we took a function $f(x)$ with infinitesimal $df = p\,dx$ (for $p \equiv \frac{df}{dx}$), and found a function $g(p)$ with infinitesimal $dg = x\,dp$, reversing the roles of the derivative and argument of the original $f(x)$. The procedure is shown graphically in Figure 4.1.

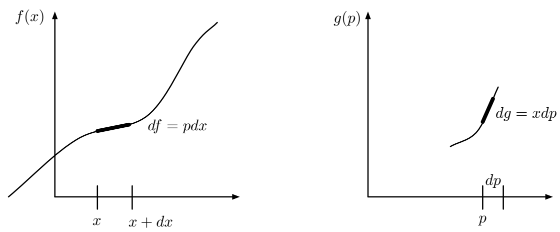

Figure 4.1: The Legendre transform constructs a function $g(p)$ from $f(x)$. We are swapping the role of the local slope of the curve for its argument.

For concreteness, take $m = 2$ and $\alpha = 1$, so that $f(x) = x^2$ – then we find from (4.7) that $g(p) = \frac{p^2}{4}$. But let's try to construct the transform $g(p)$ graphically, using only the local slope and value of $f(x)$. Start by constructing the line tangent to the curve $f(x)$ at a point x_0, say. We know that in the vicinity of x_0, the slope of this line must be $\frac{df}{dx}|_{x=x_0} \equiv p_0$, and then the equation for the line is:

$$\bar{f}(x, x_0) = p_0\,(x - x_0) + f(x_0), \tag{4.8}$$

i.e. it has slope equal to the tangent to the curve at x_0, and takes the value $f(x_0)$ at x_0. From this, we know that at the point p_0, the function $g(p)$ has slope x_0, reversing the roles of x_0 and p_0, but what is the value of the function $g(p_0)$? We don't know from the differential itself – but if we go back to the Legendre transform, we have:

$$g(p) = x(p)\,p - f(x(p)) \longrightarrow g(p_0) = x(p_0)\,p_0 - f(x(p_0)) \longrightarrow g(p_0) = x_0\,p_0 - f(x_0), \tag{4.9}$$

and we can use this value to construct the line going through the point p_0 having slope x_0:

$$\bar{g}(p, p_0) = x_0 (p - p_0) + g(p_0) = x_0 p - f(x_0) \tag{4.10}$$

which means that the p-intercept of this line has value $-f(x_0)$. So we have a graphical prescription: For every point x_0, write down the slope tangent to $f(x_0)$ (called p_0) – on the transform side, draw a line having p-intercept $-f(x_0)$ and slope given by x_0 - mark its value at p_0, and you will have the graph of $g(p)$.

In Figure 4.2, we see the graph of $f(x) = x^2$ with four points picked out, the lines tangent to those points are defined by (4.8), with $x_0 = 0, 1, 2, 3$. For example, for $x_0 = 3$, we have $p_0 = 6$, and $f(x_0) = 9$. Referring to Figure 4.3, the point at $p_0 = 6$ can be found by drawing a line with p-intercept $-f(x_0) = -9$, and slope $x_0 = 3$ – we mark where this line crosses $p_0 = 6$, as shown. In this manner, we can convert a graph of $f(x)$ into a graph of $g(p)$, and, of course, vice versa.

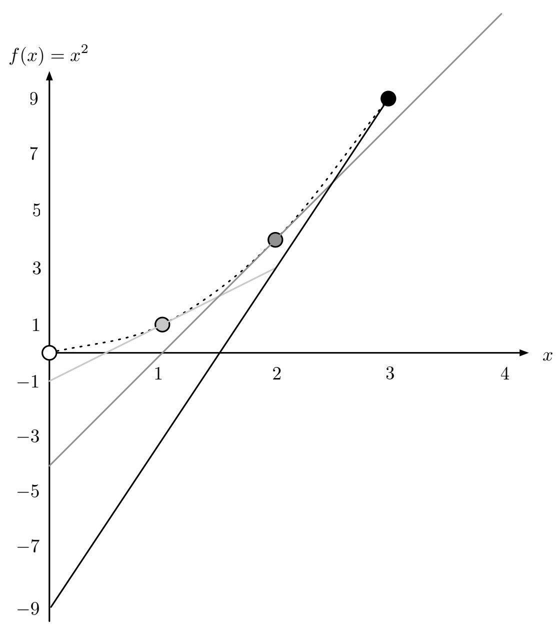

Figure 4.2: A graph of $f(x) = x^2$ with some representative points and tangent lines shown.

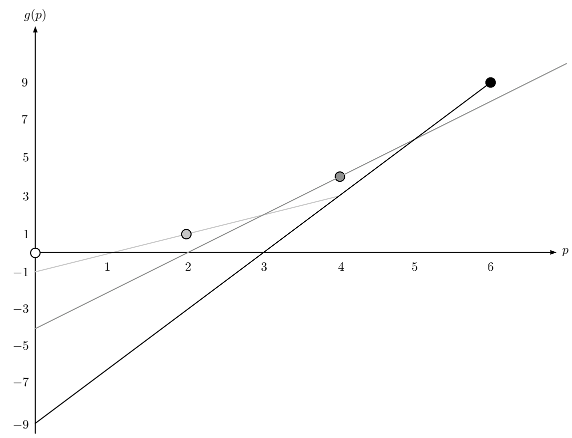

Figure 4.3: The three relevant lines coming from the data in Figure 4.2, used to define the curve $g(p)$.

Notice, finally, that the curve defined by the three points in Figure 4.3 is precisely described by $g(p) = \frac{p^2}{4}$.

Hamiltonian in Classical Mechanics

We can easily extend the above one-dimensional discussion to higher dimensions. In classical mechanics, we often start with a Lagrangian, defined as a function of $x(t)$ and $\dot{x}(t)$, say. Then we have in a sense, two variables in $L(x, \dot{x})$, and we can promote $\dot{x}$ to a full independent variable by setting $p = \frac{\partial L}{\partial \dot{x}}$, and performing a Legendre transform to eliminate $\dot{x}$ in favor of p. Define H via

$$\boxed{H(x,p) = p\,\dot{x}(p) - L(x,p) \qquad p = \frac{\partial L}{\partial \dot{x}} \qquad L(x,p) \equiv L(x, \dot{x}(p)).} \tag{4.11}$$

Notice that we have performed the transformation on only one of the two variables in the Hamiltonian. The prescription is: Use the definition of p to find $\dot{x}(p)$ and then write $H(x,p)$ entirely in terms of p.

For example, if we have a simple harmonic oscillator potential, then

$$L(x, \dot{x}) = \frac{1}{2}\,m\,\dot{x}^2 - \frac{1}{2}\,k\,x^2 \tag{4.12}$$

and $p \equiv \frac{\partial L}{\partial \dot{x}} = m\,\dot{x}$, so that $\dot{x} = \frac{p}{m}$, $L(x,p) = \frac{1}{2}\frac{p^2}{m} - \frac{1}{2}\,k\,x^2$ and we can form $H(x,p)$ as above

$$H(x,p) = p\,\frac{p}{m} - \left(\frac{1}{2}\frac{p^2}{m} - \frac{1}{2}\,k\,x^2\right) = \frac{1}{2}\frac{p^2}{m} + \frac{1}{2}\,k\,x^2. \tag{4.13}$$

We recognize this as the total energy of the system (numerically).

For the usual Lagrangian in three dimensions: $L = \frac{1}{2}\,m\,v^2 - U(\mathbf{x})$, with $v^2 = \mathbf{x}\cdot\mathbf{x}$, we can define the *canonical momentum vector* via $p_x = \frac{\partial L}{\partial \dot{x}}$, $p_y = \frac{\partial L}{\partial \dot{y}}$ and $p_z = \frac{\partial L}{\partial \dot{z}}$ and proceed to the Hamiltonian once again:

$$H(\mathbf{x},\mathbf{p}) = \frac{p^2}{2m} + U(\mathbf{x}), \tag{4.14}$$

at least, in Cartesian coordinates. This is the starting point for Hamiltonian considerations in classical mechanics, and we will begin by looking at some changes that must occur to bring this natural form into usable, relativistic notation.

4.2 Hamiltonian equations of motion

The first thing we have to deal with is Legendre transformations in our generic space (or, later, space-time). We are used to writing the Hamiltonian as (briefly employing "generalized coordinates" q)

$$\boxed{\begin{aligned} H &= \sum_i p_i\,\dot{q}_i - L \\ p_i &\equiv \frac{\partial L}{\partial \dot{q}_i}. \end{aligned}} \tag{4.15}$$

In our new notation, we see there is already an issue here. Cartesian coordinates form a contravariant tensor x^α, but then the canonical momenta are given by

$$p_\alpha = \frac{\partial L}{\partial \dot{x}^\alpha} \tag{4.16}$$

i.e. canonical momenta, unlike *physical* momenta, are covariant. That suggests that we write our Legendre transform as

$$H = p_\alpha\,\dot{x}^\alpha - L. \tag{4.17}$$

Remember the point of the Hamiltonian approach – we want to treat p_α and x^α as independent entities. The variational principle gives us, directly:

$$\begin{aligned} S &= \int dt L = \int dt \, (p_\alpha \, \dot{x}^\alpha - H) \\ \frac{\delta S}{\delta p_\alpha} &= \dot{x}^\alpha - \frac{\partial H}{\partial p_\alpha} = 0 \\ \frac{\delta S}{\delta x^\alpha} &= -\dot{p}_\alpha - \frac{\partial H}{\partial x^\alpha} = 0. \end{aligned} \tag{4.18}$$

Using our tensor notation, $L = \frac{1}{2} \, \dot{x}^\mu g_{\mu\nu} \dot{x}^\nu - U(r)$ for a spherically symmetric potential, we get, as the canonical momenta:

$$\frac{\partial L}{\partial \dot{x}^\alpha} = g_{\mu\alpha} \, \dot{x}^\mu = p_\alpha \tag{4.19}$$

so that

$$\begin{aligned} H = p_\alpha \, \dot{x}^\alpha - L &= p_\alpha \, p^\alpha - \left(\frac{1}{2} \, \dot{x}^\mu \, g_{\mu\nu} \, \dot{x}^\nu - U(r) \right) \\ &= p_\alpha \, g^{\alpha\beta} \, p_\beta - \frac{1}{2} \, p_\alpha \, g^{\alpha\beta} \, p_\beta + U(r) \\ &= \frac{1}{2} \, p_\alpha \, g^{\alpha\beta} \, p_\beta + U(r). \end{aligned} \tag{4.20}$$

From here, we can write down the equations of motion directly:

$$\begin{aligned} \dot{x}^\alpha &= \frac{\partial H}{\partial p_\alpha} = g^{\alpha\beta} \, p_\beta \\ \dot{p}_\alpha &= -\frac{\partial H}{\partial x^\alpha} = -\left(\frac{1}{2} \, p_\mu \, g^{\mu\nu}{}_{,\alpha} \, p_\nu + U_{,\alpha} \right). \end{aligned} \tag{4.21}$$

As we shall soon see, the triply indexed object $g_{\mu\nu,\alpha}$ is *not* a tensor – this will be one of the highlights of next week. The point is, we cannot raise and lower the (μ, ν) indices as we would like – instead I mention the identity

$$\boxed{g^{\mu\nu}{}_{,\alpha} = -g^{\mu\gamma} \, g^{\nu\delta} \, g_{\gamma\delta,\alpha},} \tag{4.22}$$

then the equation for $\dot{p}_\alpha$ from above is

$$\begin{aligned} \dot{p}_\alpha &= \frac{1}{2} \, p_\mu \left(g^{\mu\gamma} \, g^{\nu\delta} \, g_{\gamma\delta,\alpha} \right) \, p_\nu - U_{,\alpha} \\ &= \frac{1}{2} \, p^\gamma \, g_{\gamma\delta,\alpha} \, p^\delta - U_{,\alpha}. \end{aligned} \tag{4.23}$$

Let's introduce coordinates – suppose we are already in the reduced two-dimensional space, (ρ, ϕ) with metric (inspired by our Lagrangian studies)

$$g_{\mu\nu} \doteq \begin{pmatrix} \frac{1}{\rho^4} & 0 \\ 0 & \frac{1}{\rho^2} \end{pmatrix}. \tag{4.24}$$

We have:

$$H = \frac{1}{2} \left(\rho^4 \, p_\rho^2 + \rho^2 \, p_\phi^2 \right) + U(r) \tag{4.25}$$

then the equations of motion are given by

$$\begin{aligned} \dot{\rho} &= \frac{\partial H}{\partial p_\rho} = \rho^4 \, p_\rho \\ \dot{\phi} &= \frac{\partial H}{\partial p_\phi} = \rho^2 \, p_\phi \end{aligned} \tag{4.26}$$

and we recognize the second equation from the definition of J_z. For the rest:

$$\begin{aligned} \dot{p}_\rho &= -\frac{\partial H}{\partial \rho} = -2 \, \rho^3 \, p_\rho^2 - \rho \, p_\phi^2 - \frac{\partial U}{\partial \rho} \\ \dot{p}_\phi &= -\frac{\partial H}{\partial \phi} = 0. \end{aligned} \tag{4.27}$$

Here again, the second equation is familiar – "ignorable coordinates have conserved momenta".

Putting the derivatives together, we can write the equation of motion in terms of $\ddot{\rho}$:

$$\dot{p}_\rho = \frac{d}{dt} \left(\frac{\dot{\rho}}{\rho^4} \right) = \frac{\ddot{\rho}}{\rho^4} - 4 \, \frac{\dot{\rho}^2}{\rho^5} = -2 \, \rho^3 \left(\frac{\dot{\rho}^2}{\rho^8} \right) - \rho \, p_\phi^2 - \frac{\partial U}{\partial \rho} \tag{4.28}$$

or

$$\frac{\ddot{\rho}}{\rho^4} - 2 \, \frac{\dot{\rho}^2}{\rho^5} + \rho \, p_\phi^2 = -\frac{\partial U}{\partial \rho} \tag{4.29}$$

which is the same as our Lagrange equation of motion (with $p_\phi = J_z$).

We haven't yet gotten to the point of using H rather than L – it's coming, and to set the stage we must discuss transformations.

4.3 Canonical Transformations

A transformation takes a set of coordinates and momenta (x^α, p_α) and changes each (in theory independently) into a new set (X^α, P_α). The transformation is called "canonical" if the new system is a Hamiltonian system with Hamiltonian $H'(X, P) = H(x(X, P), p(X, P))$ – that is to say, if the new system has the Hamiltonian "equations of motion":

$$\begin{aligned} \frac{\partial H'}{\partial P_\alpha} &= \dot{X}^\alpha \\ \frac{\partial H'}{\partial X^\alpha} &= -\dot{P}_\alpha. \end{aligned} \tag{4.30}$$

Basically what we are doing is allowing any coordinate transformation of X and P. Of course, these are linked through the canonical momenta to $\dot{X}$, and we want to make sure that that linkage respects the Hamiltonian-ness.

Take a simple harmonic oscillator, we have $H = \frac{1}{2}\,p^2 + \frac{1}{2}\,x^2$. Now if we "transform" using $y = \frac{1}{2}\,x$ then H' "=" $\frac{1}{2}\,p^2 + 2\,y^2$ and the equations of motion, if we just viewed this as a Hamiltonian, give the wrong frequency of oscillation.

If instead we went back to the Lagrangian $L = \frac{1}{2}\,\dot{x}^2 - \frac{1}{2}\,x^2$, then we carry the transformation through the $\dot{x}$ term, and $L = 2\,\dot{y}^2 - 2\,y^2$ and we find out we should have had $\frac{\partial L}{\partial \dot{y}} = 4\,\dot{y} = p$ as the canonical momentum so that

$$H = (4\,\dot{y})\,\dot{y} - L = 4\,\dot{y}^2 - (2\,\dot{y}^2 - 2\,y^2) = 2\,\dot{y}^2 + 2\,y^2 = \frac{p^2}{8} + 2\,y^2 \tag{4.31}$$

which again gives the correct equation of motion.

While we are free to make any transformation we like, somehow the X and P variables are coupled if we want to make a Hamiltonian out of the resulting system. What constraints can we place on (X, P) to ensure this happens?

4.3.1 Generating Functions

Let's go back – we had a variational principle that generated the equations of motion, so equivalent to asking that the *form* of the equations of motion be retained is the requirement that the variations be identical:

$$\delta\left(\int dt\,(p_\alpha\,\dot{x}^\alpha - H(x,p))\right) = \delta\left(\int dt\,\left(P_\alpha\,\dot{X}^\alpha - H'(X,P)\right)\right). \tag{4.32}$$

This is one of those cases where it is not enough to just set the integrands equal (although that would certainly be *a* solution). What if we had a total time derivative on the right? That would look something like:

$$\int_{t_0}^{t_f} dt\,\dot{K} = K(t_f) - K(t_0) \tag{4.33}$$

i.e. it would contribute a constant – the variation of a constant is zero, so evidently, we can add any total time derivative to the integrand on the right-hand side without changing the equations of motion.

Our expanded notion is that

$$p_\alpha\,\dot{x}^\alpha - H(x,p) = P_\alpha\,\dot{X}^\alpha - H'(X,P) + \dot{K}. \tag{4.34}$$

$\dot{K}$ is fun to write down – but what is it? Hamiltonians are generally functions of position and momentum, so K must be some function of these, possibly with explicit time dependence thrown in there as well.

Remember the goal – we want to connect two sets of data: (x,p) and (X,P). If we are to have a prayer of using K efficiently, we must make it a function of at least one variable from the original (x,p) set, one variable from the (X,P) set. There are four ways to do this, and it doesn't much matter which one we pick. For now, let $K = K(x,X,t)$, then

$$\frac{d}{dt}K = \frac{\partial K}{\partial x}\,\dot{x} + \frac{\partial K}{\partial X}\,\dot{X} + \frac{\partial K}{\partial t}. \tag{4.35}$$

Inputting this into (4.34) gives us

$$\left(p_\alpha - \frac{\partial K}{\partial x^\alpha}\right)\dot{x}^\alpha - \left(P_\alpha + \frac{\partial K}{\partial X^\alpha}\right)\dot{X}^\alpha = H - H' + \frac{\partial K}{\partial t}. \tag{4.36}$$

We can make this true by setting:

$$\boxed{p_\alpha = \frac{\partial K}{\partial x^\alpha} \quad P_\alpha = -\frac{\partial K}{\partial X^\alpha} \quad \frac{\partial K}{\partial t} = 0 \quad H = H'} \tag{4.37}$$

While this does have the correct counting, $K(x,X)$ itself is not the most useful generating function. For reasons that will become clear next time, we would prefer to use a generator $\bar{K}(x, P)$. How can we find such a function? Well, from our discussion of Legendre transformations, it is apparent that we could replace X with P via a Legendre transformation taking $K(x, X) \longrightarrow \bar{K}(x, P)$. But in order for these to be Legendre duals, we must have $\frac{\partial K}{\partial X} = P$ – that is precisely (modulo a minus sign) the requirement in (4.37). Suppose, then, that we construct $\bar{K}(x, P)$ as the Legendre transform of $K(x, X)$:

$$\bar{K}(x, P) = P_\alpha \, X^\alpha + K(x, X) \qquad \frac{\partial K}{\partial X^\alpha} = -P_\alpha, \tag{4.38}$$

as an explicit check, we can show that our $\bar{K}(x, P)$ is independent of X:

$$\frac{\partial \bar{K}}{\partial X^\alpha} = P_\alpha + \frac{\partial K}{\partial X^\alpha} = 0. \tag{4.39}$$

Now (4.34) reads:

$$\begin{aligned} p_\alpha \, \dot{x}^\alpha - H &= P_\alpha \, \dot{X}^\alpha - H' + \frac{d}{dt} \left(\bar{K} - P_\alpha \, X^\alpha \right) \\ \frac{d}{dt} \, \bar{K} &= \frac{\partial \bar{K}}{\partial x^\alpha} \, \dot{x}^\alpha + \frac{\partial \bar{K}}{\partial P_\alpha} \, \dot{P}_\alpha, \end{aligned} \tag{4.40}$$

so

$$\left(p_\alpha - \frac{\partial \bar{K}}{\partial x^\alpha} \right) \dot{x}^\alpha + \left(X^\alpha - \frac{\partial \bar{K}}{\partial P_\alpha} \right) \dot{P}_\alpha = H - H' \tag{4.41}$$

and the relevant transformation connection is

$$\boxed{p_\alpha = \frac{\partial \bar{K}}{\partial x^\alpha} \qquad X^\alpha = \frac{\partial \bar{K}}{\partial P_\alpha} \qquad H = H'.} \tag{4.42}$$

For our harmonic oscillator example, we want $x = 2\,X$ as the transformation, so from the second of the above we have $\bar{K} = \frac{1}{2}\, x\, P$, then the first tells us $p = \frac{1}{2}\, P$, and we can transform the Hamiltonian:

$$H = \frac{1}{2}\, p^2 + \frac{1}{2}\, x^2 \rightarrow H' = \frac{1}{8}\, P^2 + 2\, X^2, \tag{4.43}$$

which is what we got from Lagrangian considerations.