Physics 5153 Classical Mechanics
The Hamiltonian

1 Introduction

So far in our studies of analytical mechanics have been based on Lagrangian, which is a function of n generalized coordinates and velocities, in addition can also depend on the time. The Lagrangian is used to derive a set of n second order differential equations of motion. In most cases this is more than adequate to solve any problem. In fact, this method extends into fields that Newton, Lagrange, Euler, ..., would never have dreamed of. For instance, the Lagrangian can be used to describe relativistic motion, classical fields, and relativistic quantum fields. But there are cases where a set of $2n$ first order differential equations would be more useful. In addition, there are cases where the momentum is a better description of the motion than the velocity. For these cases we use the Hamiltonian, which is a function of the coordinates and the momentum.

The principle areas where Hamiltonian mechanics comes into play are: Transformation theory, the simple form of the canonical equations of motion, and the added number of independent coordinates increases the types of transformations that are available to simplify a problem. Celestial mechanics, since most problems can not be solved one uses perturbation theory, which is closely related to transformation theory. Statistical mechanics, and quantum mechanics are based on the Hamiltonian for its use of perturbation theory, simple form of the equations of motion, connection of the commutator to the Poisson bracket, ...

1.1 The Hamiltonian

Before we jump into defining the Hamiltonian, let's consider a statement that is made in the introduction, we transform the Lagrangian from position and velocity to position and momentum so that we reduce the equations of motion to first order differential equations. Since for many systems the momentum is proportional to the velocity, why not simply reduce the Lagrange equations of motion to first order differential equations by the substitution $\dot{q}_i \to v_i$. The Lagrange equations of motion are

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = \frac{\partial L}{\partial q_i} \quad \Rightarrow \quad \begin{cases} \displaystyle\sum_{j=1}^{n}\left(\frac{\partial^2 L}{\partial v_i \partial v_j}\frac{dv_j}{dt} + \frac{\partial^2 L}{\partial v_i \partial q_j} v_j\right) + \frac{\partial^2 L}{\partial v_i \partial t} = \frac{\partial L}{\partial q_i} \\ v_i = \dfrac{dq_i}{dt} \end{cases} \tag{1}$$

where we have $2n$ first order differential equations in the n q_i and n v_i. Clearly by making the transformation $\dot{q}_i \to v_i$, we reduce the Lagrange equations of motion to first order differential equations, but the time derivatives of the velocities are mixed up in the equations, and the other terms are messy.

Notice that Lagrange equations of motion can be written as first order differential equations

$$\dot{p}_i = \frac{\partial L}{\partial q_i}$$

But L is not a function of p_i, therefore we need a transformation to take $L(q, \dot{q}, t)$ to a function of the n momenta, the coordinates and possibly the time. Let's start by examining the total derivative of the Lagrangian

$$dL = \sum_{i=1}^{n} \left[\frac{\partial L}{\partial \dot{q}_i} d\dot{q}_i + \frac{\partial L}{\partial q_i} dq_i \right] + \frac{\partial L}{\partial t} dt = \sum_{i=1}^{n} \left[p_i \, d\dot{q}_i + \dot{p}_i \, dq_i \right] + \frac{\partial L}{\partial t} dt \tag{2}$$

The next step is to remove the dependence on $\dot{q}_i$. This can be accomplished by realizing that the term $p_i \, d\dot{q}_i$ can be written as a total derivative of the product and a derivative of p_i

$$p_i \, d\dot{q}_i = d(p_i \dot{q}_i) - \dot{q}_i \, dp_i \tag{3}$$

Hence, the total derivative of the Lagrangian becomes

$$dL = \sum_{i=1}^{n} \left[d(p_i \dot{q}_i) + \dot{p}_i \, dq_i - \dot{q}_i \, dp_i \right] + \frac{\partial L}{\partial t} dt \quad \Rightarrow \quad d\left(\sum_{i=1}^{n} p_i \dot{q}_i - L \right) = \sum_{i=1}^{n} \left[\dot{q}_i \, dp_i - \dot{p}_i \, dq_i \right] - \frac{\partial L}{\partial t} dt \tag{4}$$

with the second equation defines a function of q_i, p_i, and t

$$H(q, p, t) = \sum_{i=1}^{n} p_i \dot{q}_i - L \tag{5}$$

which we call the Hamiltonian.

Let's see if we can now find a set of differential equations that defines the motion of the system. Given that we have found a function of the coordinates, momenta, and time, we calculate its total derivative

$$dH = \sum_{i=1}^{n} \left[\frac{\partial H}{\partial q_i} dq_i + \frac{\partial H}{\partial p_i} dp_i \right] + \frac{\partial H}{\partial t} dt \tag{6}$$

Comparing this equation to Eq. 4 and noting that q_i, p_i, and t are independent variables, and any variation in these quantities is independent and arbitrary, hence

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \qquad \dot{p}_i = -\frac{\partial H}{\partial q_i} \qquad \frac{\partial L}{\partial t} = -\frac{\partial H}{\partial t} \tag{7}$$

a set of $2n$ first order differential equations that are normally called the canonical equations of motion. The first equation defines the velocity, while the second equation defines the dynamics of the system, being effectively Newton's second law, and the third equation states the time dependence of both the Lagrangian and Hamiltonian are the same. If the system is composed of forces that are not derivable from a potential or nonholonomic constraints, the canonical equations of motion become

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \qquad \dot{p}_i = -\frac{\partial H}{\partial q_i} + Q_i + \sum_{j=1}^{n} \lambda_j a_{ji} \qquad \frac{\partial L}{\partial t} = -\frac{\partial H}{\partial t} \tag{8}$$

1.2 The Legendre Transform

The transformation defined in Eq. 5 is called a Legendre transform. This type of transformation allows some of the variables to participate in the transformation, while some remain the same. Assume that we have a function

$$f(u_i, w_i) \tag{9}$$

where the u_i are transformed and the w_i do not[1]. The complete differential of this function is given by

$$df = \frac{\partial f}{\partial u_i}\,du_i + \frac{\partial f}{\partial w_i}\,dw_i \tag{10}$$

We introduce the new set of variables having the following property

$$v_i = \frac{\partial f}{\partial u_i} \tag{11}$$

Note that this is exactly what we want, since we want to remove $\dot{q}_i$ and replace it with p_i. Next we define a symmetric function in terms of the old and new variables and functions

$$g(v_i) = u_i v_i - f(u_i) \tag{12}$$

where for the time being we ignore the passive variables. The complete differential for this function is

$$dg = v_i\,du_i + u_i\,dv_i - \frac{\partial f}{\partial u_i}\,du_i = \left(v_i - \frac{\partial f}{\partial u_i}\right)du_i + u_i\,dv_i = u_i\,dv_i \tag{13}$$

where Eq. 11 was used to obtain the final equality and we have assumed that Eq 11 can be inverted so that u can be written as a function of v. Note that dg can also be written as

$$dg = \frac{\partial g}{\partial v_i}\,dv_i \quad \Rightarrow \quad u_i = \frac{\partial g}{\partial v_i} \tag{14}$$

Therefore the transformation is entirely symmetric.

The Legendre transformation is give by the following steps

Old system	New system
Variables: u_i	v_i
Function: $f(u_i, w_i)$	$g(v_i, w_i)$
Transformation	
$v_i = \frac{\partial f}{\partial u_i}$	$u_i = \frac{\partial g}{\partial v_i}$
$g = u_i v_i - f$	$f = u_i v_i - g$
$\frac{\partial f}{\partial w_i} = -\frac{\partial g}{\partial w_i}$	

[1]The variables that participate in the transformation are called active variables, while those that do not participate are passive variables.

where the last expression can be derived by including the partial derivatives with respect to the variable w_i in Eqs. 13 and 14.

We will now answer the question, what is the meaning of the Legendre transform. For simplicity consider a function of a single variable $L(v)$. Next calculate the slope tangent to the point v_o

$$\text{slope} = p(v_o) = \left.\frac{\partial L}{\partial v}\right|_{v_o} \tag{15}$$

Given the slope and a specific point, we can calculate the intercept of the tangent line

$$H(v_o, p) = L(v_o) - p(v_o)v_o \quad \Rightarrow \quad H(v, p) = L(v) - p(v)v \tag{16}$$

where the second expression generalizes the first to any point on the curve. Suppose that $p(v)$ is invertible, so that we can obtain $v(p)$. In this case, the function H becomes a function of p only

$$H(p) \equiv H(v(p), p) = L(v(p)) - pv(p) \tag{17}$$

Clearly, we can start with $H(p)$ and determine $L(v)$ by the same procedure as long as $p(v)$ is invertible.

To insure invertiblity, the function $p(v)$ must be one-to-one. That is for every possible value of v in the allowed range, $p(v)$ must have a unique value. For this to be the case, $p(v)$ can not have a maximum or minimum otherwise multiple values of v will give the same value of p. This requirement can be stated as follows, the curve $L(v)$ can not have a point of inflection, since

$$\frac{\partial p}{\partial v} \neq 0 \quad \Rightarrow \quad \frac{\partial^2 L}{\partial v^2} \neq 0 \tag{18}$$

All of this can be generalized to n-dimensions. Instead of L representing a curve, it represents a hypersurface. The derivatives are now calculated along each axis as is the intercept

$$H(p) = L(v) - \sum_{i=1}^{n} p_i v_i \tag{19}$$

Finally, we require that the Hessian determinant not equal zero

$$\left|\frac{\partial^2 L}{\partial v_i \partial v_j}\right| \neq 0 \tag{20}$$

1.3 Interpretation of the Hamiltonian

In order to determine an explicit form for the Hamiltonian, let's start with the term $p_i\dot{q}_i$. The generalized momentum was previously shown to be

$$p_i = \sum_{j=1}^{n} M_{ij}\dot{q}_j + M_i \tag{21}$$

The product of velocity and momentum is therefore given by

$$\sum_{i=1}^{n} p_i\dot{q}_i = \sum_{i=1}^{n}\left[\sum_{j=1}^{n}(M_{ij}\dot{q}_j\dot{q}_i) + M_i\dot{q}_i\right] = 2T_2 + T_1 \tag{22}$$

Based on this last relation, the Hamiltonian is

$$H = \sum_{i=1}^{n} p_i \dot{q}_i - T + V = 2T_2 + T_1 - (T_2 + T_1 + T_0) + V = T_2 - T_0 + V \tag{23}$$

note, this is not the Jacobi integral, unless the system is conservative. Using matrix notation, we write the term T_2 as

$$T_2 = \frac{1}{2}\dot{\mathbf{q}}^T \mathbf{M} \dot{\mathbf{q}} \tag{24}$$

From Eq. 21, the velocity is

$$\dot{\mathbf{q}} = \mathbf{b}(\mathbf{p} - \mathbf{a}) \tag{25}$$

where $\mathbf{b} = \mathbf{M}^{-1}$, $\mathbf{M}$ is the matrix M_{ij} and $\mathbf{a}$ is the column vector M_i. The term T_2 can therefore be expanded out to be

$$T_2 = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij} p_i p_j - \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij} M_i p_j + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij} M_i M_j \tag{26}$$

The Hamiltonian can therefore be written in the following form

$$H = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij} p_i p_j - \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij} M_i p_j + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij} M_i M_j - T_0 + V \tag{27}$$

or combining like powers of momentum, we get

$$H = H_2 + H_1 + H_0 \tag{28}$$

where

$$H_2 = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij} p_i p_j$$

$$H_1 = -\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij} M_i p_j$$

$$H_0 = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij} M_i M_j - T_0 + V$$

If the system we are considering is scleronomic, then the M_i terms are all zero. In this case $T = T_2$, and the Hamiltonian is

$$H = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij} p_i p_j + V = T + V \tag{29}$$

which is the total mechanical energy.

Next we consider the time dependence of the Hamiltonian. Its total time derivative is given by

$$\dot{H} = \sum_{i=1}^{n}\left(\frac{\partial H}{\partial q_i}\dot{q}_i + \frac{\partial H}{\partial p_i}\dot{p}_i\right) + \frac{\partial H}{\partial t} \tag{30}$$

For a holonomic system, where the canonical equations of motion apply, the time derivative of the Hamiltonian becomes

$$\dot{H} = \sum_{i=1}^{n}\left(\frac{\partial H}{\partial q_i}\frac{\partial H}{\partial \dot{p}_i} - \frac{\partial H}{\partial p_i}\frac{\partial H}{\partial q_i}\right) + \frac{\partial H}{\partial t} = \frac{\partial H}{\partial t} \tag{31}$$

The total time derivative of the Hamiltonian is equal to the partial time derivative. Hence, if the Lagrangian is independent of time, then the Hamiltonian is a constant.

The result can be easily extended to a nonholonomic conservative system. In this case, the canonical equations of motion are

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \qquad\qquad \dot{p}_i = -\frac{\partial H}{\partial q_i} + \sum_{j=1}^{n}\lambda_j a_{ji} \tag{32}$$

where we recall, for a conservative system the constraints are given by

$$\sum_{i=1}^{n} a_{ji}\dot{q}_i = 0 \quad \text{with} \quad (j = 1, 2, 3, \ldots, m) \tag{33}$$

Therefore, the Hamiltonian is again a constant in this case. Therefore we conclude that every conservative system, holonomic or nonholonomic, has a constant Hamiltonian, and from Eq. 23 it is equal to the Jacobi integral. In the special case of a natural system, then $T_0 = 0$, and the Hamiltonian is the total energy.

1.4 Application of the Hamiltonian

Recall that earlier we had derived an expression similar to Eq. 5 under the assumption that the Lagrangian is independent of time and the forces derivable from a potential. The equation expressed the conservation of energy ($h = T + V$), but unlike the Hamiltonian it was a function of q_i and $\dot{q}_i$ not q_i and p_i. In the case of the Hamiltonian, we what to allow for the general case. Therefore, we start by writing the Lagrangian then applying Eq. 5. The formal procedure for arriving at the Hamiltonian is:

1. Construct the Lagrangian using an appropriate set of generalized coordinates;

2. Calculate the conjugate momenta;

3. Calculate the Hamiltonian from Eq. 5, note at this point H has both $\dot{q}_i$ and p_i;

4. Invert the generalized momentum, $p_i = \frac{\partial L}{\partial \dot{q}_i}$, to obtain $\dot{q}_i(p_i, q_i, t)$;

5. Finally, replace all $\dot{q}_i$ in H with $\dot{q}_i(p_i, q_i, t)$, so that H is a function of p_i, q_i, and t only.

Let's derive the Hamitonian for a Lagrangian given in the following general form

$$L = L_0(q_i, t) + L_1(q_i, t)\dot{q}_i + L_2(q_i, t)\dot{q}_i\dot{q}_j \tag{34}$$

where $L(q_i, t)$ are the coefficients of the various velocity terms; recall that the kinetic energy has the following general form $T = M_0 + M_i\dot{q}_i + M_{ij}\dot{q}_i\dot{q}_j$. If the generalized coordinates do not explicitly depend on time then $T = L_2\dot{q}_i\dot{q}_j$. If the forces are given by potentials then $L_0 = -V$. These conditions lead to

$$p_i = 2L_2\dot{q}_j \tag{35}$$

So finally, the Hamiltonian is

$$H = 2L_2\dot{q}_j\dot{q}_i - (-V + L_2\dot{q}_i\dot{q}_j) = T + V \tag{36}$$

so as expected, the Hamiltonian is the total energy, but only in this case. In general the procedure given above must be followed.

We can generalized this a bit further. The Lagrangian, as stated above, in many cases can be written as

$$L(\dot{q}_i, q_i, t) = L_0(q,t) + \dot{q}_i a_i(q,t) + \dot{q}_i\dot{q}_j T_{ij}(q,t) \tag{37}$$

where we allow velocity dependent potentials and explicit time dependence in the coordinates. This equation can be written in matrix notation as follows

$$L(\dot{q}_i, q_i, t) = L_0(q,t) + \dot{\mathbf{q}}^T\mathbf{a} + \frac{1}{2}\dot{\mathbf{q}}^T\mathbf{T}\dot{\mathbf{q}} \quad \Rightarrow \quad H = \dot{\mathbf{q}}^T(\mathbf{p}-\mathbf{a}) - \frac{1}{2}\dot{\mathbf{q}}^T\mathbf{T}\dot{\mathbf{q}} - L_0 \tag{38}$$

where $\mathbf{T}$ has been previously defined, and the canonical momentum is

$$\mathbf{p} = \mathbf{T}\dot{\mathbf{q}} + \mathbf{a} \tag{39}$$

By inverting Eq. 39, we can write the Hamiltonian in a more compact form. The generalized velocity is

$$\dot{\mathbf{q}} = \mathbf{T}^{-1}(\mathbf{p}-\mathbf{a}) \quad \Rightarrow \quad \dot{\mathbf{q}}^T = (\mathbf{p}^T - \mathbf{a}^T)\mathbf{T}^{-1} \tag{40}$$

where we use the fact that $\mathbf{T}$ is symmetric. Substituting these expressions into the Hamiltonian, we arrive at

$$\begin{aligned} H &= (\mathbf{p}^T - \mathbf{a}^T)\mathbf{T}^{-1}(\mathbf{p}-\mathbf{a}) - \frac{1}{2}(\mathbf{p}^T - \mathbf{a}^T)\mathbf{T}^{-1}(\mathbf{p}-\mathbf{a}) - L_0 \\ &= \frac{1}{2}(\mathbf{p}^T - \mathbf{a}^T)\mathbf{T}^{-1}(\mathbf{p}-\mathbf{a}) - L_0 \end{aligned} \tag{41}$$

1.5 Cyclic Coordinates and Momentum Conservation

If the Lagrangian does not depend on a one of the generalized coordinates, then there is a conserved momentum

$$p_q = \frac{\partial L}{\partial \dot{q}} \quad \Rightarrow \quad \dot{p}_q = \frac{\partial L}{\partial q} = 0 \quad \text{if } L \text{ is independent of } q \tag{42}$$

One would expect the same to hold for the Hamiltonian. To see this, let's write the canonical equation of motion

$$\dot{p}_q = -\frac{\partial H}{\partial q} = \frac{\partial L}{\partial q} \tag{43}$$

therefore if the Lagrangian is independent of q so is the Hamiltonian. One can also see this by examining the transformation of the Lagrangian to Hamiltonian

$$H = \dot{q}_i p_i - L(\dot{q}_i, q_i, t) \tag{44}$$

the coordinates are only contained in the Lagrangian.

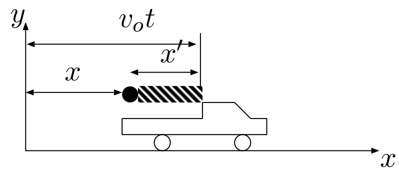

Figure 1: Cart traveling to the right with velocity v_o. The unstretched position of the spring corresponds to $x' = 0$.

1.6 Conservation of Energy and the Hamiltonian

As has already been pointed out, if the Hamiltonian is independent of time and the forces are derivable from a potential, the Hamiltonian is the total energy $(T + V)$ and it is conserved. In addition, if the Hamiltonian is independent of time it is also a constant of the motion

$$\frac{dH}{dt} = \frac{\partial H}{\partial q_i}\dot{q}_i + \frac{\partial H}{\partial p_i}\dot{p}_i + \frac{\partial H}{\partial t} = \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t} \tag{45}$$

where the second equality comes from

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} \quad \text{and} \quad \dot{q}_i = \frac{\partial H}{\partial p_i} \tag{46}$$

Therefore, if the Lagrangian is independent of time so is the Hamiltonian and the Hamiltonian becomes a constant of the motion. Note this does not imply energy conservation, since the Hamiltonian is only the total energy when it does not explicitly depend on time and forces are derivable from a potential. It will turn out that the form of the Hamiltonian will depend on the set of coordinates used.

Let's consider the following example: A cart travels at a constant velocity v_o. On the cart is a mass attached to a spring that is constrained to move in the horizontal direction only. Derive the Hamiltonian. The Lagrangian can be written as follows (see Fig. 1)

$$L = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}k(x - v_o t)^2 \tag{47}$$

where $v_o t$ corresponds to the position of the unstretch spring and x is the position of the mass. To calculate the Hamiltonian for this system, we start by calculating the canonical momentum

$$p_x = \frac{\partial L}{\partial \dot{x}} = m\dot{x} \tag{48}$$

The Hamiltonian is

$$H = p_x\dot{x} - L = \frac{p_x^2}{m} - \frac{p_x^2}{2m} + \frac{1}{2}k(x - v_o t)^2 = \frac{p_x^2}{2m} + \frac{1}{2}k(x - v_o t)^2 \tag{49}$$

Since this has the form of $H = T + V$, the Hamiltonian represents the total energy, but the Hamiltonian is not a constant of the motion.

An alternate approach to this problem is to calculate the coordinates relative to the position of the unstretched spring (moving reference frame $x' = x - v_o t$). The Lagrangian using these coordinates is

$$L = \frac{1}{2}m(\dot{x}' + v_o)^2 - \frac{1}{2}kx'^2 = \frac{1}{2}m\dot{x}'^2 + m\dot{x}v_o + \frac{1}{2}mv_o^2 - \frac{1}{2}kx'^2 \tag{50}$$

We go through the same procedure as above to calculate the Hamiltonian. First determine the canonical momentum

$$p' = \frac{\partial L}{\partial \dot{x}'} = m\dot{x}' + mv_o \quad \Rightarrow \quad \dot{x}' = \frac{p' - mv_o}{m} \tag{51}$$

Next transform the Lagrangian into the Hamiltonian

$$H = p'\dot{x}' - L = \frac{p'^2}{2m} + \frac{1}{2}kx'^2 - p'v_o = \frac{(p' - mv_o)^2}{2m} + \frac{1}{2}kx'^2 - \frac{1}{2}mv_o^2 \tag{52}$$

notice that the Hamiltonian is a constant of the motion, but it is not the total energy, it is the total energy relative to the cart excluding the last term which is a constant.