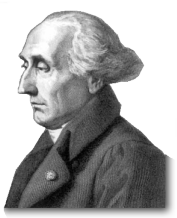
Hamiltonian Mechanics

Wednesday, 7 November 2011



Lagrange developed an alternative formulation of Newtonian mechanics, and Hamilton developed yet another. Of the three methods, Hamilton's proved the most readily extensible to the fields of statistical mechanics and quantum mechanics.

We have already been introduced to the **Hamiltonian**,

$$H = \sum_{j=1}^{N} \left(\dot{q}_j \frac{\partial L}{\partial \dot{q}_j} \right) - L = \sum_j (\dot{q}_j p_j) - L \tag{1}$$

where the generalized momenta are defined by

$$p_j \equiv \frac{\partial L}{\partial \dot{q}_j} \tag{2}$$

and we have shown that

$$\frac{dH}{dt} = -\frac{\partial L}{\partial t}$$

1. Legendre Transformation

We will now show that the Hamiltonian is a so-called "natural function" of the generalized coordinates and generalized momenta, using a trick that is just about as sophisticated as the product rule of differentiation. You may have seen **Legendre transformations** in thermodynamics, where they are used to switch independent variables from the ones you have to the ones you want. For example, the internal energy U satisfies

$$dU = T\,dS - p\,dV = \left(\frac{\partial U}{\partial S} \right)_V dS + \left(\frac{\partial U}{\partial V} \right)_S dV \tag{3}$$

The expression between the equal signs has the physics. It says that you can change the internal energy by changing the entropy S (scaled by the temperature T) or by changing the volume V (scaled by the negative of the pressure p). The final expression is simply a mathematical statement that the function $U(S, V)$ has a total differential. By lining up the various quantities, we deduce that $T = \left(\frac{\partial U}{\partial S} \right)_V$ and $p = -\left(\frac{\partial U}{\partial V} \right)_S$.

In performing an experiment at constant temperature, however, the $T\,dS$ term is awkward. It would be far nicer if the differential applied to the temperature, because then it would vanish in a constant-temperature process. A Legendre transformation trades S for T as the independent variable. Since, by the product rule, $d(TS) = T\,dS + S\,dT$, $T\,dS = d(TS) - S\,dT$. Substituting this into Eq. (3) gives

$$dU = d(TS) - S\,dT - p\,dV \qquad \Longrightarrow \qquad d(U - TS) = dF = -S\,dT - p\,dV \tag{4}$$

In other words, the Helmholtz potential, $F = U - TS$, is a (natural) function of the temperature and the volume. Furthermore, we may directly identify

$$\left(\frac{\partial F}{\partial T}\right)_V = -S \qquad \text{and} \qquad \left(\frac{\partial F}{\partial V}\right)_T = -p$$

Thus, we have found a new potential whose independent variables are T and V. It may be a more convenient quantity to calculate in many circumstances, because temperature is a handier variable than entropy.

Returning to mechanics, first calculate the total differential of the Lagrangian, which we assume not to have explicit time dependence,

$$dL = \frac{\partial L}{\partial q_j}\, dq_j + \frac{\partial L}{\partial \dot{q}_j}\, d\dot{q}_j \tag{5}$$

where we use the summation convention. Using Lagrange's equation,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} = 0 \qquad \Longrightarrow \qquad \dot{p}_j = \frac{\partial L}{\partial q_j}$$

the total differential may be rewritten

$$dL = \dot{p}_j\, dq_j + p_j\, d\dot{q}_j$$

We would now like to transfer the differential in the second term on the right-hand side from the velocity to the momentum, just as we transferred the differential from the entropy to the temperature in the example.

$$d\left(p_j \dot{q}_j\right) = p_j\, d\dot{q}_j + \dot{q}_j\, dp_j \qquad \Longrightarrow \qquad p_j\, d\dot{q}_j = d\left(p_j \dot{q}_j\right) - \dot{q}_j\, dp_j$$

so that

$$dL = \dot{p}_j\, dq_j + d\left(p_j \dot{q}_j\right) - \dot{q}_j\, dp_j$$

Taking the total differential to the left-hand side and multiplying through by -1, we have

$$d\left(\dot{q}_j p_j - L\right) = dH = \dot{q}_j\, dp_j - \dot{p}_j\, dq_j \tag{6}$$

That is, the total differential of the Hamiltonian, $H = p_j \dot{q}_j - L$, is expressed in terms of the variables p_j and q_j. Since

$$dH(q_j, p_j) = \sum_j \left(\frac{\partial H}{\partial p_j}\, dp_j + \frac{\partial H}{\partial q_j}\, dq_j\right)$$

we may readily identify

$$\boxed{\dot{q}_j = \frac{\partial H}{\partial p_j} \qquad \text{and} \qquad \dot{p}_j = -\frac{\partial H}{\partial q_j}} \tag{7}$$

These are **Hamilton's equations**. They are a set of $2N$ coupled first-order differential equations of motion, instead of the set of N coupled second-order differential equations obtained in the Lagrangian formulation.

Example 1

A simple harmonic oscillator has mass m and spring constant k. Its Lagrangian is therefore $L = \frac{1}{2}kx^2 - \frac{1}{2}m\dot{x}^2$ and its energy is $E = \frac{1}{2}kx^2 + \frac{1}{2}m\dot{x}^2$. To express the Hamiltonian, we need to compute the generalized momentum and re-express the energy in terms of the generalized coordinate and momentum. This is not too tough:

$$p = \frac{\partial L}{\partial \dot{x}} = m\dot{x}$$
$$H(x,p) = \frac{kx^2}{2} + \frac{p^2}{2m}$$

We can now confirm that Hamilton's equations give the proper equations of motion:

$$\dot{x} = \frac{\partial H}{\partial p} = \frac{p}{m} = v$$
$$\dot{p} = -\frac{\partial H}{\partial x} = -kx$$

Since $F = \dot{p}$, these are indeed the appropriate equations.

Exercise 1 Compute the Hamiltonian for a small particle moving frictionlessly inside a spherical bowl of radius R. Your answer should be expressed in terms of θ, ϕ, p_θ, and p_ϕ.

2. Phase Space

The time evolution of an isolated mechanical system is governed by Hamilton's equations, which are first order in time. Each possible configuration of the system is described by a point in a $2N$-dimensional space, called **phase space**, with coordinates $\{q_j\}$ and $\{p_j\}$. At each point in phase space, the rate of change of all the coordinates and momenta is uniquely

specified by Hamilton's equations. The point representing the system moves along a trajectory in phase space. The trajectory for an isolated, conservative system in phase space can never cross itself. Why? Because Hamilton's equations have the form $\dot{q}_j = f(q_j, p_j)$ and $\dot{p}_j = g(q_j, p_j)$. That is, the direction of motion of the point and the rate of change with time are completely determined by the position of the point in phase space. Therefore, if the trajectory returns to the same point, it must move in exactly the same way again. Thus, the system wonders in phase space along a surface of constant energy.

The **ergodic hypothesis** of statistical mechanics assumes that as the (isolated) system evolves in time, it passes arbitrarily close to every allowed point in phase space (that is, consistent with the fixed energy and any other constraints). In other words, all points on the constant energy surface are visited with equal probability. This assumption allows us to replace the time average we would make in performing a measurement with an ensemble average of systems distributed uniformly in the allowed region of phase space. Confidence in this hypothesis is bolstered by Liouville's theorem, to which we now turn.

3. Liouville's Theorem

Liouville's theorem is not about a single mechanical system but a whole collection of them. All the systems are identical—they obey the same equations of motion—but they have different initial conditions. Such collections are called **ensembles** in statistical mechanics. Suppose we have a certain phase-space density of systems in a compact region near a particular point in phase space. That is, we have a number $N = \rho(q_j, p_j)\, dV$ of systems in this region, where the volume in phase space is $dV = dq_1\, dp_1\, dq_2\, dp_2 \cdots$. Liouville's theorem then states that the density of points remains constant as the systems evolve in time (and move through phase space). That is, they flow like an incompressible fluid.

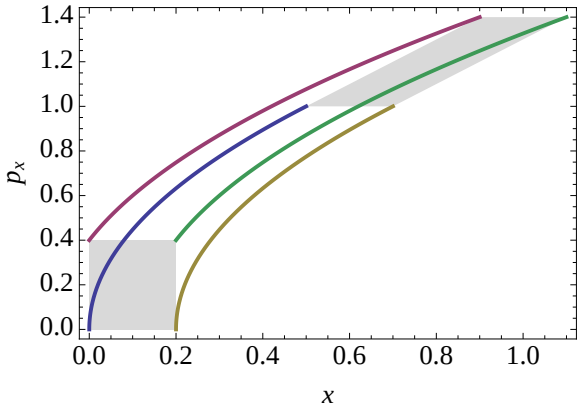


Figure 1: *A mass m in a uniform gravitational field follows a parabolic trajectory in the $x\,p_x$ plane. Systems that begin initially in the gray rectangle at the lower left will be found at a later time in the gray parallelogram at the top right, which has the same area. This illustrates Liouville's theorem.*

Let's consider a simple example: a mass m in a uniform gravitational field. Let the gravitational acceleration be g in the positive x direction and consider an ensemble of systems with initial conditions described by the

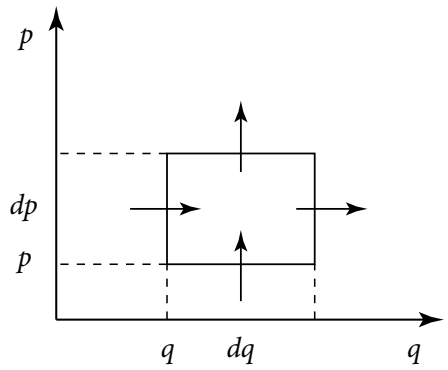


Figure 2: *A small region of phase space. The time rate of change of the number of systems within the region is given by the flux of systems into the region on the left and bottom minus the flux out of the region on the right and top.*

rectangle in the lower left of Fig. 1. The equations of motion of the system(s) are

$$p_x = p_0 + mgt$$
$$x = x_0 + \frac{p_0}{m}t + \frac{g}{2}t^2$$

meaning that the momentum increases linearly with time but the position increases quadratically. As shown in the figure, the trajectories in phase space are all parabolas. As time evolves, the points in the rectangle move along their respective trajectories and after a unit time interval are found in the gray parallelogram at the top right of the figure, which has the same area as the initial rectangular, since it has the same base and the same height.

To analyze the change in phase space density, consider the small region of phase space illustrated in Fig. 2. For simplicity we consider a one-dimensional system, which we will shortly generalize to N dimensions. Let the density of systems inside this region be $\rho(q, p)$. Then the number of systems within the region is

$$n = \rho(q, p)\, dq\, dp \tag{8}$$

In a small interval of time dt the number of systems entering the region on the left is $\rho\dot{q}\, dp\, dt$ and the number entering from the bottom is $\rho\dot{p}\, dq\, dt$. Therefore, the total number entering is

$$dn_{\text{in}} = (\rho\dot{q}\, dp + \rho\dot{p}\, dq)\, dt \tag{9}$$

Using a first-order Taylor series expansion, the number leaving the region through the right and top is

$$dn_{\text{out}} = \left(\rho\dot{q} + \frac{\partial(\rho\dot{q})}{\partial q}\, dq\right) dp\, dt + \left(\rho\dot{p} + \frac{\partial(\rho\dot{p})}{\partial p}\, dp\right) dq\, dt \tag{10}$$

Subtracting the outflow from the inflow gives the net change in the number of systems in the volume in the time interval dt:

$$\frac{dn}{dt} = \frac{dn_{\text{in}}}{dt} - \frac{dn_{\text{out}}}{dt} = \frac{\partial \rho}{\partial t}\, dq\, dp = -\left[\frac{\partial(\rho \dot{q})}{\partial q} + \frac{\partial(\rho \dot{p})}{\partial p}\right] dq\, dp$$

Therefore,

$$\frac{\partial \rho}{\partial t} = -\left[\rho \frac{\partial \dot{q}}{\partial q} + \dot{q}\frac{\partial \rho}{\partial q} + \rho \frac{\partial \dot{p}}{\partial p} + \dot{p}\frac{\partial \rho}{\partial p}\right] \tag{11}$$

From Hamilton's equations,

$$\frac{\partial \dot{q}}{\partial q} + \frac{\partial \dot{p}}{\partial p} = \frac{\partial}{\partial q}\left(\frac{\partial H}{\partial p}\right) + \frac{\partial}{\partial p}\left(-\frac{\partial H}{\partial q}\right) = 0$$

by the equality of mixed partials. We may thus simplify Eq. (11) to get

$$\frac{\partial \rho}{\partial t} = -\left(\dot{q}\frac{\partial \rho}{\partial q} + \dot{p}\frac{\partial \rho}{\partial p}\right)$$

Generalizing now to a set of N generalized coordinates and momenta, we have

$$\frac{\partial \rho}{\partial t} + \sum_{j=1}^{N}\left(\frac{\partial \rho}{\partial q_j}\dot{q}_j + \frac{\partial \rho}{\partial p_j}\dot{p}_j\right) = \frac{d\rho}{dt} = 0$$

That is, ρ remains constant as the systems evolve in time. This completes the proof of Liouville's theorem. If at some time the systems are uniformly distributed over the allowed energy surface in phase space, then they will remain uniformly distributed. Note that this result required the (first-order) Hamiltonian formulation of dynamics in phase space. No analogous result exists in the Lagrangian formulation.

4. The Virial Theorem

There is another general statistical result that is important particularly in astrophysical systems. Consider a system of particles with positions $\mathbf{r}_\alpha$ and momenta $\mathbf{p}_\alpha$, all of which remain bounded as the system evolves in time. Now define the quantity

$$\Lambda \equiv \sum_\alpha \mathbf{p}_\alpha \cdot \mathbf{r}_\alpha$$

and compute its time derivative:

$$\frac{d\Lambda}{dt} = \sum_\alpha \left(\dot{\mathbf{p}}_\alpha \cdot \mathbf{r}_\alpha + \mathbf{p}_\alpha \cdot \dot{\mathbf{r}}_\alpha\right) = \sum_\alpha \left(\mathbf{F}_\alpha \cdot \mathbf{r}_\alpha\right) + 2T$$

Now compute a time average of this expression over a time interval long compared to the time it takes a particle to move across its range of travel,

$$\left\langle \frac{d\Lambda}{dt} \right\rangle = \frac{1}{\tau}\int_0^\tau \frac{d\Lambda}{dt}\, dt = \frac{\Lambda(\tau) - \Lambda(0)}{\tau} = 2\langle T\rangle + \langle \sum_\alpha \mathbf{F}_\alpha \cdot \mathbf{r}_\alpha \rangle$$

Since the motion is assumed to remain bounded, the value of $\Lambda(\tau) - \Lambda(0)$ remains finite (if the motion is periodic, waiting one period causes this difference to vanish) and so the average can be made as small as we like by waiting long enough. Therefore,

$$\boxed{\langle T \rangle = -\frac{1}{2}\Big\langle \sum_\alpha \mathbf{F}_\alpha \cdot \mathbf{r}_\alpha \Big\rangle} \tag{12}$$

Rudolf Clausius called the term on the right the **virial**,[1] so this result is called the **virial theorem**. For particles interacting with a central force of the form $F_\alpha \propto r^n$, so $V_\alpha = kr^{n+1}$, we have the special case that

$$\langle T \rangle = -\frac{1}{2}\Big\langle \sum_\alpha k(n+1) r^n \times r \Big\rangle = \frac{n+1}{2}\langle V \rangle \tag{13}$$

For spring forces, this means that $\langle T \rangle = \langle V \rangle$, while for gravitation ($n = -2$),

$$\langle T \rangle = -\frac{1}{2}\langle V \rangle \qquad \text{(gravity)}$$

5. Summary

Early on we saw that the Hamiltonian,

$$H \equiv p_j \dot{q}_j - L$$

is conserved when the Lagrangian does not depend explicitly on time, and that it is equal to the total energy when the equations of transformation do not depend explicitly on time. Now we appreciate that the Hamiltonian is a function of the generalized coordinates q_j and the generalized momenta p_j, and that the equations of motion are

$$\dot{q}_j = \frac{\partial H}{\partial p_j} \qquad \text{and} \qquad \dot{p}_j = -\frac{\partial H}{\partial q_j}$$

How can you remember which of these two nearly symmetric equations gets the minus sign? Well, the time derivative of a coordinate is a velocity, which should come from the kinetic energy. That's never negative. On the other hand, the time derivative of a momentum is a force, and that should be the negative derivative of the potential.

Hamilton's equations reduce motion to curves in phase space. Because they are first-order equations, trajectories cannot intersect. Isolated systems move around on a constant-energy surface in $2N$-dimensional phase space, flowing like an incompressible fluid. In some simple systems, the flow is placid, orderly, and laminar. In nonlinear systems, it is often akin to taffy in a taffy machine, which stretches, folds, and remixes the fluid over and over until bits that started out very close to one another can be on "opposite sides" of phase space. In that case, the system exhibits **Hamiltonian chaos**.

[1] From the Latin *vires*, plural of *vis*, meaning strength, force, vigor.