Mechanics – Lecture 10 – 4/10/2019

Remaining task: explain relation between Lagrangian + Hamiltonian viewpts; also a little about how this is related to optimal control. Specifically:

A) Equivalence of Lagr + Hamilt viewpts can be seen by brute-force calculation (though we don't see very clearly <u>why</u> it's true by this method)

B) Better insight is obtained by identifying a link to Hamilton-Jacobi pde

C) <u>In</u> mechanics, we often get <u>critical pts</u> (not minima) of $\int_{t_1}^{t_2} L(q,\dot{q})\,dt$. The analogue of (B) but insisting on minimization is an example of an optimal control problem. (The Hamilton-Jacobi eqn assoc to this can easily have non-smooth solns; the theory of <u>viscosity</u> solns was created to deal with that.)

Concerning (A): recall our hypoth (from Lecture 9) that $L(q,\dot{q})$ is convex + superlinear in $\dot{q}$. Our claim is that a sol'n of

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = \frac{\partial L}{\partial q_i}$$

can be written in Hamiltonian form

$$\dot{p}_i = -\frac{\partial H}{\partial q_i}, \quad \dot{q}_i = \frac{\partial H}{\partial p_i}$$

by taking $p_i = \frac{\partial L}{\partial \dot{q}_i}$ and

$$(*) \qquad H(q,p) = \max_{\xi} \langle p, \xi \rangle - L(q,\xi)$$

[Note that the optimal ξ for $(*)$ satisfies $p = \frac{\partial L}{\partial \xi}$ so that

$$(**) \qquad H(q,p) = \langle p, \dot{q} \rangle - L(q,\dot{q})$$

when $p = \frac{\partial L}{\partial \dot{q}}$. Also, recall that

$$(q,\dot{q}) \xrightarrow[p = \frac{\partial L}{\partial \dot{q}}]{} (q,p)$$

is a well-defined and invertible change of coordinates on phase space.]

<u>step 1</u>: take differentials –

$$dH = \sum \frac{\partial H}{\partial p_i} dp_i + \frac{\partial H}{\partial q_i} dq_i$$

by chain rule (here $\frac{\partial H}{\partial p_i}$ is calculated with q_i fixed, etc)

Also, from (**),

$$dH = \sum \dot{q}_i dp_i + \cancel{p_i d\dot{q}_i} - \sum \frac{\partial L}{\partial q_i} dq_i + \cancel{\frac{\partial L}{\partial \dot{q}_i} d\dot{q}_i}$$

(here $\frac{\partial L}{\partial q_i}$ is calculated with $\dot{q}_i$ fixed!)

So: $\quad \frac{\partial H}{\partial p_i} = \dot{q}_i \quad , \quad \frac{\partial H}{\partial q_i} = \frac{-\partial L}{\partial q_i}$ as fns on phase space.

<u>Step 2</u>: Now consider how p + q change along a trajectory satisfying $\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = \frac{\partial L}{\partial q_i}$: we find

$$\frac{d}{dt} p_i = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = \frac{\partial L}{\partial q_i} = \frac{-\partial H}{\partial q_i}$$

$$\frac{d}{dt} q_i = \dot{q}_i = \frac{\partial H}{\partial p_i}$$

Now (B). To start, I observe that over a sufficiently <u>short</u> time interval, a crit pT of

$$\int_{t_1}^{t_2} L(q, \dot{q})\, ds$$

is assured to be a minimum. The reason is that changing variables to $s = t/\varepsilon$,

$$\int_0^\varepsilon L(q, \tfrac{dq}{dt})\, dt = \varepsilon \int_0^1 L(q, \tfrac{dq}{ds}\cdot\tfrac{1}{\varepsilon})\, ds$$

As $\varepsilon \to 0$ the convexity (and superquadratic growth) of L wrt $\dot{q}$ dominates, making the RHS a convex function of $q(s)$. [Exercise: fill in the details - perhaps some additional conds on L are needed.]

With this in mind, we can consider the min action as fn of the final time + position (fixing t_1 = initial time, & leaving initial position free):

$$u(t_2, x_2) = \min_{\substack{q(t_2) = x_2 \\ q(t_1) = x_1 \text{ arbitrary}}} \int_{t_1}^{t_2} L(q, \dot{q})\, d\tau$$

The optimizer will (evidently) be a soln of the Lagrangian formulation of mechanics.

By "prin of dynamic programming"

$$u(t, x) \simeq \min_{\alpha} \; u(t - \Delta t, x - \alpha \Delta x) + L(x, \alpha)\Delta t$$

by taking paths whose last little bit has $\dot{q} = \alpha$. Proceeding formally:

$$u(t,x) \simeq \min_\alpha \; u(X,t) + \Delta t \left\{ -u_t - \alpha \cdot \nabla u + L(x,\alpha) \right\}$$

Cancel $u(t;x)$ and divide by Δt to get

$$u_t = \min_\alpha \left\{ L(x,\alpha) - \alpha \cdot \nabla u \right\}.$$

$$= -\max_\alpha \left\{ \alpha \cdot \nabla u - L(x,\alpha) \right\}$$

$$= -H(X, \nabla u)$$

Thus: $u(t,x)$ solves $\quad u_t + H(x,\nabla u) = 0$ for $t > t_1$
$u = 0$ at $t = t_1$.

(Note: t_1 was fixed throughout the preceding discussion.)

More: along the optimal paths we have $du/d\tau = L(q,\dot{q})$, so we expect a connection to the method of characteristics. In fact, Hamilton's eqns are the characteristic eqns for $u_t + H(x,\nabla u) = 0$; more specifically, if

$$\frac{dx}{dt} = \nabla_p H \quad \text{and} \quad \frac{dp}{dt} = -\nabla_x H$$

Then along the resulting curve

$$\frac{d}{dt} u(x(t),t) = \langle p, \dot{x} \rangle - H(p, x(t)).$$

(Thus: solving pde along this well-chosen curve reduces to an ODE.)

<u>Explain</u>: if $u_t + H(x, \nabla u) = 0$ then by defn

$$\frac{\partial^2 u}{\partial x_i \partial t} + \sum_i{}' \left(\frac{\partial H}{\partial x_i} + \frac{\partial H}{\partial p_j} \frac{\partial^2 u}{\partial x_i \partial x_j} \right) = 0$$

Now,

$$\frac{d}{dt} \nabla_i u(x(t), t) = \frac{\partial^2 u}{\partial x_i \partial t} + \sum{}' \frac{\partial^2 u}{\partial x_i \partial x_j} \frac{dx_j}{dt}$$

along <u>any</u> curve. If we choose $\frac{dx_j}{dt} = \frac{\partial H}{\partial p_j}$ then we get

$$\frac{d}{dt} \nabla_i u(x(t), t) = \frac{-\partial H}{\partial x_i}$$

(so we have $\nabla_i u(x(t), t) = p_i(t)$) and

$$\begin{aligned} \frac{d}{dt} u(x(t), t) &= \langle \nabla u, \dot{x} \rangle + u_t \\ &= \langle p, \dot{x} \rangle - H \end{aligned}$$

as asserted.

Now (c): in optimal control, the goal is to optimize something similar to our "action", but typically it's the final time $t_2 = T$ that's fixed. A classic example is

$$u(x,t) = \min_{y(t)=x} \int_t^T \tfrac{1}{2}|\dot{y}|^2\,dt + g(y(T)).$$

cost of travel from initial location x to final location $y(T)$

penalty at final time

Optimal paths have constant velocity (by Jensen), so in fact

$$(***) \qquad u(x,t) = \min_{z\in\mathbb{R}^N} \left\{ \frac{1}{2}\frac{|z-x|^2}{(T-t)} + g(z) \right\}.$$

But arguing as we did earlier gives a Hamilton-Jacobi eqn for u:

$$u(x,t) \simeq \min_{\alpha} u(t+\Delta t, x+\alpha\Delta t) + \tfrac{1}{2}|\alpha|^2\Delta t$$

$$\simeq \min_{\alpha} u(x,t) + \Delta t\left\{ u_t + \langle \alpha, \nabla u\rangle + \tfrac{1}{2}|\alpha|^2 \right\}.$$

So (formally)

$$u_t + \min_{\alpha} \left\{ \langle \alpha, \nabla u \rangle + \tfrac{1}{2}|\alpha|^2 \right\} = 0$$

The optimal α is $-\nabla u$, giving the "final-value pblm"

$$u_t - \tfrac{1}{2}|\nabla u|^2 = 0 \quad \text{for } t<T$$

$$u = g \quad \text{at } t=T.$$

However: when optimal z for (***) is nonunique (which can easily happen!) $u(x,t)$ is not smooth (nor even differentiable!), so it's not clear how to justify our formal calculation.

The theory of viscosity solns (of 1st order eqns) was designed for precisely such problems. Evans' pde book (chapters 3+10) is a good place to read about this.