Mechanics - Lecture 10, 4/11/2018 (typos fixed 4/25)

These notes:

a) discuss some "applications" of the Hamiltonian perspective (i.e. properties that are clear from this perspective)

b) explanations - more conceptual + appealing than the one in Lecture 9 - about why the Lagrangian + Hamiltonian perspectives are equivalent. (These explanations focus on the "action" integral + its links to Hamilton-Jacobi eqns + optimal control)

Also, at the end of these notes:

c) Fermat's principle of "least travel time" (and a related connection between mechanics + geometry, via geodesics)

d) Some facts about Fenchel transforms that are needed to make the arguments at end of Lecture 9 notes (relating Lagrangian + Hamiltonian approaches) honest.

"Applns" of the Hamiltonian perspective — that in a well-chosen coord system our evolution is

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

where $H = H(\vec{q}, \vec{p})$. (Recall: for particles interacting by a potential U, $p_i = m_i \dot{q}_i$ and $H = \frac{1}{2}\sum^n \frac{1}{m_i}|p_i|^2 + U$ = kinetic + potential energy.)

<u>1st consequence</u>: H is constant along trajectories:

$$\frac{dH}{dt} = \sum \frac{\partial H}{\partial p_i}\dot{p}_i + \frac{\partial H}{\partial q_i}\dot{q}_i = 0$$

using chain rule + Hamilton's eqns. (We are of course assuming that H is a fn of q + p <u>only</u>, indep of t)

[<u>Note</u>: this is the <u>same</u> "cons of energy" law that we obtained in the Lagrangian setting. There we got

$$\left(\sum \dot{q}_i \frac{\partial L}{\partial \dot{q}_i}\right) - L = \text{const along trajectories.}$$

But from pg 9.16 of Lecture 9, together

with the defn $p_i = \frac{\partial L}{\partial \dot{q}_i}$, we have

$$H(p,q) = \sum \frac{\partial L}{\partial \dot{q}_i}\dot{q}_i - L(q,\dot{q})$$

as expected.]

<u>2nd consequence</u>: "dimension reduction": if the Lagrangian is independent of q_1 then so is the Hamiltonian. As a result p_1 is constant & the problem reduces to Hamilton's eqns in $(q_2, \dots, q_n; p_2, \dots, p_n)$.

In fact: $\frac{dp_1}{dt} = \frac{-\partial H}{\partial q_1} = 0$ by hypoth; so $p_1 = \text{const}$,

So: solve $\dot{p}_j = \frac{-\partial H}{\partial q_j}$, $\dot{q}_j = \frac{\partial H}{\partial p_j}$ $(j \geq 2)$

by substituting the const value of p_1 into H. Finally, get $q_1(t)$ at the end by integrating

$$\frac{dq_1}{dt} = \frac{\partial H}{\partial p_1}$$

along the resulting path.

This argument can be repeated, So: if L & H depend on just one spatial variable

Then evolution can be reduced to phase plane analysis.

3rd consequence: Liouville's Thm: In the (q,p) coordinates, the flow assoc to our evolution is volume-preserving

Pf: for any flow we can consider its "infinitesimal generator"

$$\text{image of } \vec{x} \text{ after time } t = \vec{x} + \underbrace{\vec{f}(x)}_{\text{inf'l generator}} t + O(t^2)$$

and the flow is vol-preserving iff $\operatorname{div} \vec{f} = 0$, since

$$\frac{d}{dt}\Big|_{t=0} (\text{vol of image of } D) = \int_D \frac{d}{dt}\Big|_{t=0} \det(I + t\nabla f)$$

$$= \int_D \operatorname{div} \vec{f}$$

Apply this to Hamiltonian flow: $\vec{x} = (\vec{q}, \vec{p})$ and $\vec{f} = \left(\frac{\partial H}{\partial p}, -\frac{\partial H}{\partial q}\right)$ so

$$\operatorname{div} f = \sum \frac{\partial}{\partial q_i}\left(\frac{\partial H}{\partial p_i}\right) - \frac{\partial}{\partial p_i}\left(\frac{\partial H}{\partial q_i}\right) = 0$$

Liouville's thm has some surprising consequences, eg through

<u>Poincaré's recurrence thm</u>: if g is a vol preserving map (for example, the time-1 map of a Hamiltonian flow) + $g(D) = D$ for some set D of finite volume, then it is "recurrent" in sense that:

for any set B of pos measure (eg a tiny ball) $\exists\, x_0 \in B$ st $g^n(x_0)$ is again in B for some $n < \infty$.

(Instructive examples: rotn of S^1 by rational or irrational angle.)

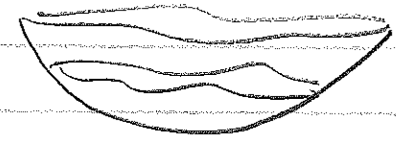
Pf of recurrence: clearly $B, g(B), g^2(B), \ldots$ cannot all be disjt, so $\exists\, x_1 \in g^k(B) \cap g^l(B)$ for some $l < k$. Then $x_0 = g^{-k} x_1$ satisfies $x_0 \in B \cap g^{l-k}(B)$. So

$$x_0 \in B \ \& \ g^{k-l}(x_0) \in B,$$

as asserted.

<u>Typical mechanical consequence</u>:

Consider motion of a ball in an asymmetrical bowl



Region of phase space st $T+U \le$ const is invariant + has finite volume. So ball returns to _almost_ its initial position + velocity.

Turning to (b) = understanding link btwn Hamiltonian + Lagrangean viewpts, key is the "action"

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

Recall that in Lagr mechanics path is a crit pt for this, and (due to strict convexity of L in $\dot q$) this crit pt is a (local) min if $t_2 - t_1$ is small enough. If we fix final time + final position, we can consider

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

and the optimizer will be a soln of Lagrangian 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:

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

$$\Rightarrow \quad u_t = \min_\alpha L(x,\alpha) - \alpha \cdot \nabla u$$

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

$$= -H(x, \nabla u).$$

Thus: $u(t,x)$ evolves (for $t > t_1$ time) by HJ eqn $u_t + H(x, \nabla u) = 0$, with $u = 0$ at $t = t_1$.

More: along the optimal paths we have $du/ds = L(q, \dot{q})$, so we expect a connection to the <u>method of characteristics</u>.

In fact: Hamilton's eqns are the characteristic eqns for $u_t + H(x, \nabla u) = 0$; more specifically, if

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

then along the resulting curve

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

(Thus: solving the pde along this well-chosen curve requires only solving an ODE.)

Explain: if $u_t + H(x, \nabla u) = 0$ then by differentiation

$$\frac{\partial^2 u}{\partial x_i \partial t} + \sum_j \frac{\partial H}{\partial p_j} \frac{\partial^2 u}{\partial x_i \partial x_j} + \frac{\partial H}{\partial x_i} = 0$$

so

$$\frac{d}{dt} \nabla_i u(x(t), t) = \frac{\partial^2 u}{\partial x_i \partial t} + \sum \frac{\partial^2 u}{\partial x_i \partial x_j} \frac{dx_j}{dt}$$

along any curve $x(t)$. If we choose $\frac{dx_j}{dt} = \frac{\partial H}{\partial p_j}$ then we get

$$\frac{d}{dt} \nabla_i u(x(t), t) = \frac{-\partial H}{\partial x_i} \qquad \text{(the eqn for } p_i\text{!)}$$

and

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

as asserted.

Thus:

* Lagrangian viewpt leads naturally to considering action minimization (the var'l prob defining u)

* This leads naturally to the HJ eqn $u_t + H(x, \nabla u) = 0$ where H = Fenchel transform of L

* Hamilton's eqns give paths along which evoln reduces to an ODE,

* But along Lagrangian paths u also reduces to an ode ($du/ds = L(q_0, \dot q_0)$). So of course these paths are the same. (And indeed the ode's are consistent, since along the paths $\langle p, \dot q_0 \rangle - H(p, q_0) = L(q_0, \dot q_0)$.)

(c) You might get the idea that the "action" is useful only as a theoretical tool. Actually it's also useful in more practical ways. Let's discuss the "principle of least travel time" (good source: §1.10–1.12 of Bühler's book).

First example: Geodesics on a hypersurface $S \subset \mathbb{R}^n$.

Main pt: a particle constrained to stay on S (but subject to no other forces) travels along a geodesic, at constant speed. To see this, consider:

var'l prbm 1: let $A = \int_{t_1}^{t_2} \frac{1}{2}|\dot{x}|^2\,dt$ (the action!). Particle has $\delta A = 0$ for perturbations that stay on S. So

$$\delta x \text{ tangent to } S \Rightarrow \int_{t_1}^{t_2} \langle \dot{x}, \delta\dot{x}\rangle\,dt = 0$$

$$\Rightarrow -\int_{t_1}^{t_2} \langle \ddot{x}, \delta x\rangle\,dt = 0$$

provided perturbation vanishes at t_1, t_2.
True for all vars $\Rightarrow \ddot{x} \perp S$.

<u>var'l pbm 2</u>: let $\mathcal{L}$ = arclength = $\int_{t_1}^{t_2} |\dot{x}|\,dt$.

A geodesic has $\delta\mathcal{L}=0$ for all perturbations that stay on S. Arguing as above,

$$\begin{array}{l}\delta x \text{ tangent to } S \\ \text{(vanishing at endpts)}\end{array} \implies \int_{t_1}^{t_2} \left\langle \frac{\dot{x}}{|\dot{x}|}, \delta\dot{x} \right\rangle dt = 0$$

ie
$$\frac{d}{dt}\left(\frac{\dot{x}}{|\dot{x}|}\right) \text{ is normal to } S.$$

<u>Connection</u>: solns to "var'l pbm 1" have constant speed + traverse paths assoc "var pbm 2".

<u>Pf</u>: If $x(t)$ solves pbm 1, then

$$\frac{d}{dt}|\dot{x}|^2 = 2\langle \dot{x}, \ddot{x} \rangle = 0$$

so speed is constant. Evidently

$$\frac{d}{dt}\left(\frac{\dot{x}}{|\dot{x}|}\right) \perp S$$

so it solves pbm 2.

Conversely, if path solves pbm 2 then a constant-speed param clearly has $\ddot{x} \perp S$ so is extremal for pbm 1.

2nd Example: Mechanical system in $\mathbb{R}^n$ with no potential, and kinetic energy

$$T = \frac{1}{2} f^2(x) |\dot{x}|^2$$

with $f > 0$.

var'l pbm 1: particles trajectories are extremal for the action

$$A = \int_{t_1}^{t_2} \frac{1}{2} f^2(x(t)) |\dot{x}(t)|^2 \, dt$$

var pbm 2: consider paths of "least travel time" where speed $= 1/f$. They're extremal for

$$\mathcal{L} = \int_{t_1}^{t_2} f(x(t)) \, |\dot{x}(t)| \, dt$$

(note: t is just a parameter here, not time.)

Rmk: in geometrical optics, "wave-front" is set at fixed travel-time from a given pt.

<u>Claim</u>: correspondence btwn the two pbms is exactly as for Geodesics: along solns of pbm 1, $f^2|\dot{x}|^2 = $ const, + path is a soln of pbm 2

<u>Pf</u>: Since $T = \frac{1}{2} f^2(x)|\dot{x}|^2$, $H = $ Legendre transf $= \frac{1}{2} f^{-2}(x)|p|^2$.

From Hamilton's eqns

$$\dot{x} = -\frac{\partial H}{\partial p} = f^{-2} p, \qquad \dot{p} = \frac{-\partial H}{\partial x}$$

From 1st eqn and constancy of H,

$$H = \frac{1}{2} f^{-2}(x)|p|^2 = \frac{1}{2} f^2(x)|\dot{x}|^2 \text{ is const in } t$$

Now $x(t)$ extremal for pbm 1 $\Rightarrow \frac{-d}{dt}(f^2\dot{x}) + f\nabla f|\dot{x}|^2 = 0$

$$\Rightarrow \frac{d}{dt}\left(\frac{f^2\dot{x}}{f|\dot{x}|}\right) = \nabla f \cdot |\dot{x}| \text{ (since denominator is constant)}$$

$$\Rightarrow \frac{d}{dt}\left(f\frac{\dot{x}}{|\dot{x}|}\right) = \nabla f|\dot{x}|$$

$\Rightarrow x(t)$ is extremal for pbm 2.

Conversely, if $x(t)$ is extremal for 2 then

$$\frac{d}{dt}\left(f\frac{\dot{x}}{|\dot{x}|}\right) = \nabla f\,|\dot{x}|$$

so a param st $f|\dot{x}| = \text{constant}$ will have

$$\frac{d}{dt}\left(f^2\dot{x}\right) = f\nabla f\,|\dot{x}|^2$$

thus being extremal for pbm 1.

[Preceding calc extends with no essential change to $T = \int \frac{1}{2}\sum a_{ij}(x)\,\dot{x}_i\,\dot{x}_j\,dt$

$$\mathcal{L} = \int\left(\sum a_{ij}(x)\,\dot{x}_i\,\dot{x}_j\right)^{1/2} dt \qquad]$$

Example 3: what about mechanical system with a potential? Ans: we can still do something very similar! Consider Lagrangian

$$L = T - V = \frac{1}{2}|\dot{x}|^2 - V(x)$$

for which Hamiltonian is $H = \frac{1}{2}|\dot{x}|^2 + V(x)$.
(Recall that $H = \text{const}$ along solns.)

Claim: a path $x(t)$ with energy $H = E$ is extremal for

$$\mathcal{L} = \int_{t_1}^{t_2} \sqrt{2(E-V(x))}\ |\dot{x}(t)|\, dt$$

Proof proceeds as usual: if x solves mechanical eqn $\ddot{x} = -\nabla V$, then since $H = E$ along the path

$$\tfrac{1}{2}|\dot{x}|^2 + V(x) = E$$

$$\Rightarrow \quad |\dot{x}| = \sqrt{2(E-V)}$$

Now, cond of being extremal for $\mathcal{L}$ is

$$-\int_{t_1}^{t_2} [2(E-V)]^{-1/2} \langle V, \delta x\rangle |\dot{x}| \;+\; \sqrt{2(E-V)} \left\langle \frac{\dot{x}}{|\dot{x}|}, \delta\dot{x} \right\rangle = 0$$

ie

$$-[2(E-V)]^{-1/2}\, \nabla V\, |\dot{x}| - \left(\sqrt{2(E-V)}\ \frac{\dot{x}}{|\dot{x}|}\right)^{\cdot} = 0$$

Since $|\dot{x}| = \sqrt{2(E-V)}$ this says

$$-\nabla V - \ddot{x} = 0$$

which is true! Argt in opposite direction is easy as usual (extremal for $\mathcal{L}$ $\Rightarrow$ with param st $|\dot{x}| = \sqrt{2(E-V)}$ we get a soln of the mechanical eqns)

Finally, as promised, let me fill in a detail that was asserted w/o proof in Lecture 9, when we 1st discussed both the Lagr. + Hamilt. viewpts. Recall that

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

$$= \text{Fenchel transform of } L \text{ wrt } \dot q$$
(holding q fixed).

Discussion at end of Lecture 9 used that

① $p_i = \frac{\partial L}{\partial \dot q_i}$ determines a well-defined change of coords $(q,\dot q) \to (q,p)$

② we can recover the Lagrangian from the Hamiltonian by

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

③ we can get $\dot q$ as fn of p by

$$\dot q = \frac{\partial H}{\partial p_i}$$

The following explanation follows Craig Evans' pde book. We need to assume (not stated clearly in Lecture 9) that $L(q,\dot{q})$ is not only convex but also superlinear in $\dot{q}$, i.e.

$$\frac{L(q,\dot{q})}{|\dot{q}|} \to \infty \quad \text{as } |\dot{q}| \to \infty.$$

Claims ①–⑤ all follow from the following assertions about convex fns $\varphi(\xi)$ s.t.

$$\frac{\varphi(\xi)}{|\xi|} \to \infty \text{ as } |\xi| \to \infty$$

(Note that this implies

$$\varphi^*(\eta) = \max_{\xi} \langle \eta, \xi \rangle - \varphi(\xi)$$

is finite for all η.)

<u>Claim</u>: A) φ^* is convex, and $\frac{\varphi^*(\eta)}{|\eta|} \to \infty$ as $|\eta| \to \infty$.

B) $\varphi^{**} = \varphi$

<u>Pf of (A)</u>: φ^* = max of lin fns, so it's certainly convex.

Take $\xi = \lambda\, \eta/|\eta|$ as test choice $\Rightarrow$

$$\varphi_*(\eta) \geq \lambda|\eta| - \varphi(\lambda\, \eta/|\eta|)$$

$$\Rightarrow \frac{\varphi_*(\eta)}{|\eta|} \geq \lambda - \underbrace{\frac{\max \text{ of } \varphi \text{ on } B_\lambda}{|\eta|}}_{\to 0 \text{ as } |\eta| \to \infty}$$

So $\liminf\limits_{|\eta|\to\infty} \dfrac{\varphi^*(\eta)}{|\eta|} \geq \lambda$ for any λ.

Pf of B) Clearly $\varphi^*(\eta) + \varphi(\xi) \geq \langle \xi, \eta \rangle$ for all ξ, η. So

$$\varphi(\xi) \geq \langle \xi, \eta\rangle - \varphi^*(\eta)$$

whence

$$\varphi \geq \varphi^{**}$$

For the reverse, observe that

$$\varphi^*(\eta) = \langle \xi, \eta\rangle - \varphi(\xi) \quad \text{when } \nabla_\xi \varphi = \eta.$$

So $\quad \varphi(\xi) = \langle \xi, \eta\rangle - \varphi^*(\eta) \quad$ when $\nabla_\xi \varphi = \eta$.

Thus $\quad \varphi^{**}(\xi) = \max\limits_{\eta} \langle \xi, \eta \rangle - \varphi^*(\eta)$

$$\geq \langle \xi, \nabla\varphi(\xi)\rangle - \varphi^*(\nabla\varphi(\xi)) = \varphi(\xi).$$