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

A) Equivalence of Lagr. + Hamilt viewpoints can be seen by brute-force calculations (though we don't see very clearly why it's true by this method).

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

C) In mechanics, we often get critical pts (not minima) if \( \int L(q, \dot{q}) dt \). The analogue of (B) but insisting on minimization is an example of an optimal control problem. (The Hamilton-Jacobi equation associated to this can easily have non-smooth solutions; the theory of viscosity solutions was created to deal with that.)

Concerning (A): recall our hypothesis (from Lecture 9) that \( L(q, \dot{q}) \) is convex & super-linear in \( \dot{q} \). Our
claim is that a solv 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

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

Note that the optimal \( \dot{q} \) for (4) satisfies \( p = \frac{\partial L}{\partial \dot{q}} \)

so that

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

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

\[
(q, \dot{q}) \rightarrow (q, p) \quad p = \frac{\partial L}{\partial \dot{q}}
\]
is a well-defined and invertible change of coordinates on phase space.

step 1: take differentials -
\[ dH = \sum \frac{\partial H}{\partial q_i} dq_i + \frac{\partial H}{\partial p_i} dp_i \]

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

Also, from (4.11),

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

(here \( \frac{\partial H}{\partial p_i} \) is calculated with \( q_i \) fixed!)

So:

\[ \frac{\partial H}{\partial q_i} = \dot{q}_i \quad , \quad \frac{\partial H}{\partial p_i} = -\frac{\partial L}{\partial q_i} \quad \text{as functions of phase space.} \]

Step 2: Now consider how \( p + q \) change along a trajectory satisfying \( \frac{d}{dt} (\frac{\partial L}{\partial \dot{q}_i}) = \frac{\partial L}{\partial q_i} \), we find

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

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

\[ \text{Now (B). To start, I observe that over a sufficiently short time interval, a crit pt of} \]

\[ \int_{t_1}^{t_2} L(q, \dot{q}) \, dt \]
is asserted to be a minimum. The reason is that changing variables to $s = t/\varepsilon$,

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

As $\varepsilon \to 0$ the convexity (and super-quadratic growth) of $L$ with $q$ dominates, making the RHS a convex function of $q(t)$. [Exercise: fill in the details — perhaps some additional conditions on $L$ are needed.]

With this in mind, we can consider the min action as a function of the final time and position (fixing $t_i$ = initial time, $x_i$ = leaving initial position free):

$$u(t_f, x_f) = \min_{t_1, \ldots, t_i} \int_{t_1}^{t_f} L(q, \dot{q}) \, dt$$

$q(t_f) = x_f$
$q(0) = x_0$ arbitrary

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

By “principle of dynamic programming”

$$u(t, x) \geq \min_x u(t - \Delta t, x - \Delta x) + L(x, \dot{x}) \Delta t$$

by taking paths whose last little bit has $\dot{q} = x$. Proceeding formally:
$$u(t,x) = \min_{\partial} u(x,t) + \Delta \frac{\xi}{2} - u_{+} - \nabla H u_{\|} + L(x,\xi)$$

Cancel $u(t,x)$ and divide by $\Delta t$ to get

$$u_{+} = \min_{\partial} \{ L(x,\xi) - u_{\|} \}$$

$$= - \max_{\partial} \{ u_{\|} - L(x,\xi) \}$$

$$= - H(x,\xi)$$

Thus $u(t,x)$ solves $u_{+} + H(x,\xi) = 0$ for $t > t_{f}$ and $u = 0$ at $t = t_{f}$.

(Note: $t_{f}$ was fixed throughout the preceding discussion.)

More: along the optimal paths we have

$$\frac{d\xi}{dt} = \nabla H$$

and

$$\frac{dp}{dt} = - \nabla H.$$
(Thus, solving PDE along this well-chosen curve reduces to an ODE.)

\[
\frac{\partial^2 u}{\partial x_i \partial t} + \sum_i \left( \frac{\partial \mathcal{H}}{\partial x_i} \frac{\partial u}{\partial x_i} + \frac{\partial \mathcal{H}}{\partial p_i} \frac{\partial u}{\partial p_i} \right) = 0
\]

Now,

\[
\frac{d}{dt} \nabla_i u(x(t), t) = \frac{\partial^2 u}{\partial x_i \partial t} + \sum_i \frac{\partial^2 u}{\partial x_i \partial p_i} \frac{dx_i}{dt}
\]

along any curve. If we choose \( \frac{dx_i}{dt} = \frac{\partial \mathcal{H}}{\partial p_i} \) then we get

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

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

\[
\frac{d}{dt} u(x(t), t) = \langle \nabla u, \dot{x} \rangle + u_t
\]

\[
= \langle p, \dot{x} \rangle - H
\]

as asserted.

Now (c): in optimal control, the goal is to optimize something similar to our "action", but typically it's the final time \( t_f = T \) that's fixed. A classic example is
10.7
\[ u(x, t) = \min_y \int_0^T \frac{1}{2} |\dot{y}|^2 \, dt + g(y(T)) \]
\[ y(t) = x \]
\[ \text{cost + penalty at final time} \]
\[ \text{travel from initial location} \]
\[ x \text{ to final location } y(T) \]

Optimal paths have constant velocity (by Jensen), so in fact
\[ u(x, t) = \min_{z \in \mathbb{R}^n} \left\{ \frac{1}{2} \frac{|z-x|^2}{(t-T)} + g(z) \right\} \]

But arguing as we did earlier gives a Hamilton-Jacobi eqn for \( u \):
\[ u(x, t) = \min_\Delta u(t + \Delta t, x + \Delta x) + \frac{1}{2} \lambda^2 \Delta t \]
\[ = \min_\Delta u(x, t) + \Delta t \left\{ \frac{1}{2} \lambda^2 + \langle \lambda, 72 \dot{y} \rangle + \frac{1}{2} |x|^2 \right\} \]

So (finally)
\[ u_t + \min_{\Delta} \left\{ \lambda^2 + \frac{1}{2} \lambda^2 \right\} = 0 \]

The optimal \( \lambda \) is \(-7z\), giving the "final-value form"
\[ u_t - \frac{1}{2} |z|^2 = 0 \quad \text{for } t < T \]
\[ u = g \quad \text{at } t = T \]
However, when optimal $z$ for $(\star \star)$ is nonuniquely (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 solutions (first-order PDEs) was designed for precisely such problems. Evans' PDE book (chapters 3-10) is a good place to read about this.