These notes:

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

b) explanations - were conceptual and appealing than the one in Lecture 9 - about why the Lagrangian and Hamiltonian perspectives are equivalent. (These explanations focus on the "action" integral & its links to Hamilton-Jacobi equations & 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)
"Applying the Hamiltonian perspective — that is, a well-chosen coordinate system, the evolution is

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

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

1st consequence: \( H \) is constant along trajectories:

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

using chain rule + Hamilton's equations. (We are of course assuming that \( H \) is a \( \text{fn} \) of \( q + p \) only, \( \text{with} \) \( \frac{\partial H}{\partial t} \).)

**Note:** this is the same "conservation of energy" law that we obtained in the Lagrangian setting. Therefore, we get

\[ (\sum \frac{\partial H}{\partial q_i} \dot{q}_i) - L = \text{const along trajectories.} \]

Best from pg 9, 16 of Lecture 9, together
with the data \( p_i = \frac{\partial L}{\partial \dot{q}_i} \), we have

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

as expected.

2nd consequence: "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 equations in \( (q_2, \ldots, q_n, p_2, \ldots, p_n) \).

In fact: \( \frac{\partial p_1}{\partial \dot{q}_1} = -\frac{\partial H}{\partial q_1} = 0 \) by hypothesis; so \( p_1 = \text{const.} \)

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

by substituting the constant value of \( p_1 \) into \( H \).

Finally, get \( q_j(t) \) 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 Theorem: In the $(q, p)$ coordinates, the flow of $\mathbf{f}$ is volume-preserving.

For any flow we can consider its "infinitesimal generator"

\[
\text{map of } x \text{ after } t = \overline{x} = \overline{f(x)} + O(t^2)
\]

with

\[
\text{generator}
\]

and the flow is volume-preserving iff $\text{div } \mathbf{f} = 0$.

Since

\[
\frac{D}{Dt} \left( \text{vol of } \omega \text{ at } t = 0 \right) = \frac{D}{Dt} \left( \text{vol of } \omega \text{ at } t = 0 \right)
\]

\[
= \int_D \text{div } \mathbf{f}
\]

Apply this to Hamiltonian flow: $\overline{x} = (\overline{q}, \overline{p})$ and

\[
\mathbf{f} = \left( \frac{\partial H}{\partial p}, -\frac{\partial H}{\partial q} \right)
\]

\[
\text{div } \mathbf{f} = \sum_i \frac{\partial}{\partial q_i} \left( \frac{\partial H}{\partial p_i} \right) - \frac{\partial}{\partial p_i} \left( \frac{\partial H}{\partial q_i} \right) = 0
\]
Liouville's theorem has some surprising consequences, e.g. through

Poincaré's recurrence theorem: if \( g \) is a vol preserving map (for example, the time-1 map of a Hamiltonian flow) and \( g(D) = D \) for some set \( D \) of finite volume, then it is "recurrent" in a sense that:

for any set \( B \) of pos measure (e.g. a tiny ball) \( \exists x_0 \in B \) s.t. \( g^n(x_0) \) is again in \( B \) for some \( n < \infty \).

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

Proof of recurrence: clear, \( B, g(B), g^2(B), \ldots \) cannot all be disjoint, so \( \exists x_j \in g^j(B) \) for some \( j \leq k \), then \( x_0 = g^{-k}x_j \) satisfies \( x_0 \in B \cap g^{-k}(B) \). So

\[ x_0 \in B \cap g^{-k}(x_0) \subset B, \]

as asserted.

Typical mechanical consequence:
Consider motion of a ball on an asymmetrical bowl.

Region of phase space at $T+U \leq \text{const}$ is invariant & has finite volume. So ball returns to almost its initial position & velocity.

Turning to (b) = understanding link between Hamiltonian & Lagrangian viewpoints, key is the "action"

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

Recall that in Lagrangen 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

$$U(t_2, x_2) = \min \int_{t_1}^{t_2} L(q, \dot{q}) \, dt$$

$\dot{q}(t_2) = x_2$ & arbitrary
and the optimizer will be as solved
Lagrangean mechanics.

By "principle of dynamic programming"

\[ u(t, x) = \min_{\alpha} \left\{ u(t - \Delta t, x - 2\Delta x) + L(x, x) \Delta t \right\} \]

by taking paths ever-once last little bit
has \( \alpha = 1 \). Proceeding formally,

\[ u(t, x) = \min_{\alpha} [u(x + \delta x, t + \Delta t) - 2\Delta x - 2 \cdot 5u + L(x, x) \Delta t] \]

\[ = \min_{\alpha} \left\{ \delta x \cdot 7u - L(x, x) \right\} \]

\[ = -H(x, 7u) \]

Thus: \( u(t, x) \) evolves (for \( t > t_1 \))
turn) by \( H \) eqn \( u + H(x, 7u) = 0 \),
with \( u = 0 \) at \( t = t_1 \).

Hence: Along the optimal paths we have
\( d/ds = L(q, \dot{q}) \), so we expect a connection
to the method of characteristics.
In fact: Hamilton's equations are the characteristic equations for \( u_x + H(x, u) = 0 \).

More specifically, if
\[
\frac{dx}{dt} = \nabla H \quad \text{and} \quad \frac{dp}{dt} = -\nabla_x H
\]

Then along the resulting curve
\[
\frac{du}{dt} (\mathbf{x}(t), t) = \langle p, \dot{x} \rangle - H(p, x(t))
\]

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

Explain: if \( u_x + H(x, u) = 0 \) then by definition
\[
\frac{\partial^2 u}{\partial x^2} + \sum_{i=1}^n \frac{\partial}{\partial x_i} \left( \frac{\partial}{\partial x_i} \frac{\partial u}{\partial x} + \frac{\partial H}{\partial x_i} \right) = 0
\]

so
\[
\frac{du}{dt} (\nabla u (\mathbf{x}(t), t)) = \frac{\partial^2 u}{\partial x^2} + \sum_{i=1}^n \frac{\partial}{\partial x_i} \frac{\partial u}{\partial x} \frac{dx_i}{dt}
\]
along any curve \( \mathbf{x}(t) \). If we choose \( \frac{dx_i}{dt} = \frac{\partial H}{\partial p_i} \), then we get
\[
\frac{du}{dt} (\mathbf{x}(t), t) = -\frac{\partial H}{\partial x_i}
\]
(Thus the eqn for \( p_i \)).

and
\[
\frac{du}{dt} (\mathbf{x}(t), t) = \langle \nabla u, \dot{x} \rangle + u_t
\]
\[
= \langle p, \dot{x} \rangle - H
\]
as asserted.
Thus:

* Lagrangian recept leads naturally to considering action minimization (the world path defining $\mathbf{u}$)

* This leads naturally to the HT $(\mathbf{u})$

$$\frac{d}{dt} H(x, \dot{x}, \mathbf{u}) = 0$$

where $H$ = Feynman transform of $L$

* Hamilton's equations give paths along which evolves reduces to an ODE,

* But along Lagrangian paths it also reduces to an ODE ($\frac{d\mathbf{u}}{ds} = L(\mathbf{u}, \dot{\mathbf{u}})$). So of course these paths are the same. (And indeed the ODE's are consistent, since along these paths $\langle p, \dot{q} \rangle = H(p, q) = L(q, \dot{q})$.)

(c) You might get the idea that the "action" is useless 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: § 11.10 - 11.12 of Riehl's
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'd plan 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

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

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

provided pertubation vanishes at $t_1$ and $t_2$, i.e. for all vanes $\Rightarrow \dot{x} \perp S$. 
var'd pbw 2: let $L$ = arclength = $\int_1^{t_2} |\dot{x}| \, dt$.

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

$$\delta x \text{ tangent to } S \implies \int_1^{t_2} \left< \frac{\dot{x}}{|\dot{x}|}, \delta x \right> \, dt = 0$$

(varishing at end pt)

ie

$$\frac{\delta}{\delta t} \left( \frac{\dot{x}}{|\dot{x}|} \right) \text{ is normal to } S,$$

connection: solves to "var'd pbw 1" have constant speed + traverse paths assoc "var pbw 2."

If $x(t)$ solves pbw 1, then

$$\frac{d}{dt} \left( \frac{\dot{x}}{|\dot{x}|} \right)^2 = 2 \left< \ddot{x}, \dddot{x} \right> = 0$$

so speed is constant. Evidently

$$\frac{\delta}{\delta t} \left( \frac{\dot{x}}{|\dot{x}|} \right) + \lambda$$

so it solves pbw 2.
Conversely, if path solves pbm 2 then a constant-speed path clearly has 0 LS so is extremal for pbm 1.

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

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

with $f > 0$.

var pbm 1: particles trajectories are extremal for the action

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

var pbm 2: consider paths of "least travel time" whose speed $= \sqrt{f}$, they're extremal for

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

(note: t is just a parameter here, not time.)
Rule: in geometrical optics, "wave-front" is set at fixed travel-time from a given pt.

Claim: correspondence between the two plans is exactly as in Geodesics; along solutions of
plm 1, $f'x'^2 = \text{const}$, + path is a soln of
plm 2.

**Pf:** Since $T = \frac{1}{2} f'^2 x'^2$, $H = \text{Legendre transform}$

$$= \frac{1}{2} f'^2 x'^2 = p^2.$$

From Hamilton's eqns

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

From 1st eqn and constancy of $H$,

$$H = \frac{1}{2} f'^2 x'^2 \dot{p}^2 = \frac{1}{2} f'^2 x'^2 (dx'^2) = \text{const} \ w, t$$

Now $x(t)$ extremal in plm 1 $\Rightarrow \frac{d}{dt} (f'^2 x') + f f'^2 x'^2 = 0$

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

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

$$\Rightarrow x(t) \text{ is extremal in plm 2.}$$
Conversely, if \( x(t) \) is extremal for \( 2 \) then

\[
\frac{\partial}{\partial t} \left( f \frac{x}{|x|} \right) = \nabla f \cdot \mathbf{x}
\]

so a path at \( f |x| = \text{constant} \) will have

\[
\frac{\partial}{\partial t} \left( f^2 \mathbf{x} \right) = \nabla f \cdot f \mathbf{x} |x|^2
\]

has being extremal for \( p \in \text{in} \).

[Preceding calc extends with no essential change to \( T = \int \frac{1}{2} \sum \alpha_i(x) x_i \mathbf{x}_i \cdot \mathbf{x}_i \, dt \).]

\[
L = \int \left( \sum \alpha_i(x) x_i \mathbf{x}_i \cdot \mathbf{x}_i \right)^{\frac{1}{2}} \, dt
\]

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

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

for which Hamiltonian is \( H = \frac{1}{2} |x|^2 + V(x) \).

(Recall that \( H = \text{const along solutions} \).)

Claim: a path \( x(t) \) with energy \( H = E \) is
extremal for
\[ L = \int_{t_1}^{t_2} \sqrt{2(E-V(x))} \, |\dot{x}(t)| \, dt \]

Proof proceeds as usual: if \( x \) solves mechanical eqn \( \ddot{x} = -\nabla V \), then since \( H = E \) along the path
\[ \frac{1}{2} |x|^2 + V(x) = E \]

\[ \Rightarrow \quad |x| = \sqrt{2(E-V)} \]

Now, cond of being extremal for \( L \) is

\[ -\int_{t_1}^{t_2} [2(E-V)]^{-1/2} \langle \nabla V, \delta x \rangle |x| + 12(E-V) \frac{\dot{x}}{|x|} \delta x = 0 \]

ie

\[ - [2(E-V)]^{-1/2} \nabla V |x| - \left( \frac{12(E-V) \dot{x}}{|x|} \right) = 0 \]

Since \( |x| = \sqrt{2(E-V)} \) This says

\[ -\nabla V - \ddot{x} = 0 \]

which is true! Next, in opposite direction is easy as usual (extremal for \( L \) \Rightarrow with given \( \dot{x} \) at \( |x| = \sqrt{2(E-V)} \) we get a solution of the mechanical eqns)
Finally, as promised, let me fill in a detail that was asserted without proof in Lecture 9, when we first discussed the Euler-Lagrange-Hamilton approach. Recall that

\[ H(q, p) = \max \langle \dot{q}, p \rangle - L(q, \dot{q}) \]

\[ = \text{Fenchel transform of } L \text{ wrt } \dot{q} \]

(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 variables

\[ (q, \dot{q}) \rightarrow (q, p) \]

2. we can recover the Lagrangian from the Hamiltonian by

\[ L(q, \dot{q}) = \max \langle \dot{q}, p \rangle - H(q, p) \]

3. we can get \( \dot{q} \) as a function of \( p \) by

\[ \dot{q} = \frac{\partial H}{\partial p} \]
The following explanation follows Craig Evans' book. We need to assume (not stated clearly in Lecture 9) that \( L(\epsilon, \tilde{\epsilon}) \) is not only convex but also superlinear in \( \tilde{\epsilon} \), i.e.

\[
\frac{L(\epsilon, \tilde{\epsilon})}{|\tilde{\epsilon}|} \rightarrow \infty \quad \text{as} \quad |\tilde{\epsilon}| \rightarrow \infty.
\]

Claim: 0-5 all follow from the following assertions about convex fun \( \varphi(\tilde{\epsilon}) \):

\[
\frac{\varphi(\tilde{\epsilon})}{|\tilde{\epsilon}|} \rightarrow \infty \quad \text{as} \quad |\tilde{\epsilon}| \rightarrow \infty
\]

(Note that this implies)

\[
\varphi^*(y) = \max_{\tilde{\epsilon}} \left< y, \tilde{\epsilon} \right> - \varphi(\tilde{\epsilon})
\]

is finite for all \( y \).)

Claim: A) \( \varphi^* \) is convex, and \( \varphi^*(y) \rightarrow \infty \) as \( |y| \rightarrow \infty \).

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

Proof (A): \( \varphi^* \) is weak* continuous, so it's certainly convex.
Take $\delta = \frac{\gamma}{171}$ as $\delta > \frac{1}{\delta}$.

\[ \delta^*(\gamma) \geq \frac{\gamma}{171} - \delta \left( \frac{\gamma}{171} \right) \]

\[ \Rightarrow \frac{\delta^*(\gamma)}{171} \geq \frac{1}{\delta} - \max \text{ of } \delta \text{ on } B \]

\[ \Rightarrow \frac{\delta^*(\gamma)}{171} \rightarrow 0 \text{ as } 171 \rightarrow \infty \]

So, limit $\frac{\delta^*(\gamma)}{171} \rightarrow \infty$ for any $\gamma$.

**Proof (R)**: Clearly $\delta^*(\gamma) \cdot \delta(\delta) \leq \sqrt{\varepsilon, \gamma} \Rightarrow \frac{\delta^*(\gamma)}{171} \rightarrow 0$ for all $\varepsilon, \gamma, \delta$.

\[ \delta(\delta) \leq \sqrt{\varepsilon, \gamma} - \delta^*(\gamma) \]

\[ \therefore \delta \geq \delta^* \]

For the reverse, observe that

\[ \delta^*(\gamma) = \sqrt{\varepsilon, \gamma} - \delta(\delta) \text{ when } \frac{\gamma}{\delta} = \gamma \]

So

\[ \delta(\delta) = \sqrt{\varepsilon, \gamma} - \delta^*(\gamma) \text{ when } \frac{\gamma}{\delta} = \gamma \]

Thus

\[ \delta^*(\delta) = \max \leq \sqrt{\varepsilon, \gamma} - \delta^*(\gamma) \]

\[ \geq \sqrt{\varepsilon, \delta^*(\delta)} - \delta^*(\sqrt{\varepsilon, \delta^*(\delta)}) = \varphi(\delta) \]