These notes:

a) more careful treatment of the reln between $H + L$ (more generally: the reln between a convex $f$ and its Fenchel transform), augmenting what was at end of Lect 9 notes.

b) discussion of the "action" link to Hamilton-Jacobi eqns, and just a bit about optimal control.

c) Fermat's principle of "least travel time" (also the connection between mechanics + geometry, via geodesics).

Topic (a) recall that

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

= Fenchel transform of $L$ wrt $\dot{q}$

(holding $q$ fixed)

and is showing the efficacy of Lagr + Hamilton viewpoints we needed that
\[ p_i = \frac{\partial L}{\partial \dot{q}_i} \]

determined a well-defined change of

cords \((q, \dot{q}) \rightarrow (\tilde{q}, \tilde{p})\) also, we can

recover the Lagrangian from the Hamiltonian

by

\[ L(q, \dot{q}) = \text{max} \ <\tilde{q}, \tilde{p}> - H(q, \tilde{p}) \]

and we can get \( \dot{q} \) as a fn of \( q, \tilde{p} \) by

\[ \dot{q}_i = \frac{\partial H}{\partial \tilde{p}_i} \]

What we're using here is a little more than

I explained in the Lecture 9 notes.

Following Evans' treatment, let \( \varphi(\tilde{q}) \) be convex

+ assume

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

so that

\[ \varphi^*(\tilde{q}) = \text{max} \ <q, \tilde{q}> - \varphi(q) \]

is finite for all \( q \). Then the assertions

above about \( H + L \), etc. follows from the facts

That

1) \( \varphi^* \) is convex, \( \frac{\varphi^*(q)}{\|q\|} \rightarrow 0 \) as \( q \rightarrow 0 \).
2) \( \Phi^* \Phi = \Phi \)

**Pf of (1):** \( \Phi^* \) = max of \( \Phi \) on \( \bar{B} \), so certainly convex.

Take \( \varepsilon = \frac{2}{19} \) as test choice \( \Rightarrow \)

\[
\Phi^*(y) \geq \frac{1}{19} - \frac{\Phi(x)}{19}
\]

\[
\Rightarrow \frac{\Phi^*(y)}{19} \geq x - \frac{\max \Phi \text{ on } \bar{B}}{19}
\]

\[
\Rightarrow \text{lim inf}_{19 \to \infty} \frac{\Phi^*(y)}{19} \geq x \text{ for any } x.
\]

**Pf of (2):** Clearly \( \Phi^*(y) + \Phi(x) \geq <\varepsilon, y> \) for all \( \varepsilon, y \).

\[
\Phi(x) \geq <\varepsilon, y> - \Phi^*(y)
\]

where \( \Phi \geq \Phi^* \).

For the reverse, observe that

\[
\Phi^*(y) = <\varepsilon, y> - \Phi(x) \text{ when } \frac{\varepsilon}{\Phi} = y.
\]

So
\[ \varphi(x) = \mathbf{e}^x \cdot 2 + \int_0^x \varphi''(s) \, ds \quad \text{when} \quad \varphi'(0) = g. \]

Thus

\[ \varphi''(x) = \max \left( \varphi(0), \varphi(1) \right) \]

\[ = \varphi(x) \]

Topic (2): The link to HT is given. This can be seen in two different (but related!) ways:

1) as an example of the method of characteristics (e.g., Evans' PDE book, ch. 3)

2) by viewing action min as an optimal control problem (e.g., Bücker's book, ch. 11.10)

Let's discuss both.

For (1): consider the HT eqn. \( \frac{\partial u}{\partial t} + H(\nabla u, x) = 0. \) Method of characteristics solves it (along certain curves in "phase space"), by solving an appropriate ODE, \( x(t) + p(t) = \varphi'' u(x(t)). \)
Claim: The characteristics in this case are

\[ \frac{\partial x}{\partial t} = \nabla p \cdot H, \quad \frac{\partial p}{\partial t} = -\nabla x \cdot H \]

and along the curve that results,

\[ \frac{\partial}{\partial t} u(x(t), t) = \langle p, \dot{x} \rangle - H(p, x). \]

Indeed, suppose \( u_t + H(u, x) = 0 \). Then

\[ \frac{\partial}{\partial t} \sum \frac{\partial u}{\partial x_i} + \sum \frac{\partial H}{\partial p_j} \frac{\partial u}{\partial x_i} + \frac{\partial H}{\partial x_i} = 0. \]

So,

\[ \frac{\partial}{\partial t} \left( \sum \frac{\partial u}{\partial x_i} \right) + \sum \frac{\partial^2 u}{\partial x_i \partial x_j} \frac{\partial x_j}{\partial t} \]

along any curve. If we choose \( \frac{\partial x_i}{\partial t} = \frac{\partial H}{\partial p_j} \), then we get

\[ \frac{\partial}{\partial t} \sum \frac{\partial u}{\partial x_i}(x(t), t) = -\frac{\partial H}{\partial x_i} \] (This is the eqn for \( p_i \) !)

and we have

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

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

as asserted.
This calculation works, but it seems mysterious. The other viewpoint explains why it works.

**Explain (2):** Recall that over short times, the Lagrangian integral isn't just stationary, it is minimized. Let's consider dependence of the integral ("action") on final time and position:

Let $u(t, x_2) = \min \int_{t_1}^{t_2} L(g, \dot{g}) \, dt$

$g(t_2) = x_2$

$g(t_1) = x_1$ arbitrary

By the "principle of dynamic programming"

$$u(x, t) = \min_x \left\{ u(t + \Delta t, x - x \Delta t) + L(x, x) \Delta t \right\}$$

By taking paths where last little bit has $\Delta t = \Delta$, proceeding formally:

$$u(x, t + \Delta t) = \min_x \left\{ u(x, t) + \Delta \left\{ -2u_x - \alpha \cdot \nabla u + L \right\} \right\}$$

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

$$= -\max_x \left( \alpha \cdot \nabla u - L(x, x) \right)$$
Moreover: optimal paths for this case of various plans must be characteristics of the HJ
equation along these paths.

\[ \frac{\partial}{\partial t} u = L \]

Our previous calculus gave \( \frac{\partial}{\partial t} u = \langle p, x \rangle - H \),
but that's consistent since

\[ \langle p, x \rangle - H = \langle p, \nabla_p H \rangle - H \]

\[ = \langle p, \dot{q} \rangle - H = L. \]

**Discussion:** The link to optimal control starts here. A typical optimal control problem is

\[ \min_{\bar{u}} \int_{t_0}^{T} h(x(t), z(t)) \, dt + F(x(T)) \]

- \( \dot{x} = f(x, z) \)
- \( x(t_0) = x_0 \) \( \uparrow \) \( \text{this is} \)
- \( z(t) \)
- \( \text{the "control"} \)
- \( \dot{x} = f(x, z) \)
- \( \text{the "state eqn"} \)
and the "value for" \( u(x_0, t_0) \) is the min value.

Our mechanics problem has zero or little this term, with \( F = 0 \) then \( z = \text{velocity} \), \( \mathcal{L} = \text{Lagrangian} \). Our choice to fix the final-the location instead of initial-the location was to get \( u_t + H = 0 \) instead of \( u_t - H = 0 \) is essentially cosmetic.

Over large times, method of characteristics isn't very reliable (\( u(x, t) \) is not always smooth) and observation of \( H \) error must also be re-examined. This leads to notion of viscosity solution. (See Evans for a good introduction.)

What is the \( H \) error assoc to problem (1)?

By "principle of duality",

\[
\begin{align*}
    u(x_0, t_0) &\approx \min_x \left\{ \mathcal{L}(x_0, x) \Delta t + u(x_0 + \Delta t f(x_0, x), x_0, t_0 + \Delta t) \right\} \\
    &\quad + u_t(x_0, t_0) \left( \Delta t \right)
\end{align*}
\]

\[
\Rightarrow u_t + H(x, \nabla u, t) = 0
\]

(finally)
where \( H(x, p, t) = \min_x P(x, x) + p \cdot f(x, x) \).  

Here the formula for \( H \) isn't exactly a Fenchel transform (it is concave in \( p \)) but the calculus is clearly very similar to what we did earlier.

**Topic:** The principle of "least travel time."  
(Good source for this: 3.1.10-1.12 of Eichler's book.)  

**First example:** Geodesics on a hypersurface \( S \subset \mathbb{R}^n \)

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

- **Goal 1:** let \( A = \frac{1}{2} \int_1^{t_2} \| \dot{x} \|^2 \, dt \) ("action").

The particle has \( \delta A = 0 \) for all perturbations that stay on \( S \). So

\[
\delta x \text{ tangent to } S \implies \int_1^{t_2} \langle \dot{x}, \delta x \rangle \, dt = 0
\]

\[
\implies -\int_1^{t_2} \langle \ddot{x}, \delta x \rangle \, dt = 0
\]

provided perturbation vanishes at \( t_1 \) and \( t_2 \).
True for all variations ⇒ \( \dot{x} + S \).

You did prove 2: let \( L = \text{arc length} = \int_{t_1}^{t_2} \sqrt{1 \dot{x}^2} \ dt \).

A geodesic has \( \delta L = 0 \) for all perturbations that stay on \( S \), arguing as above,

\[
\delta x \text{ tangent to } S \implies \frac{d}{dt} \langle \dot{x}, \delta \dot{x} \rangle dt = 0 \\
(\text{vanishing at endpoints}) = 0.
\]

i.e.

\[
\frac{d}{dt} \left( \frac{\dot{x}}{\sqrt{\dot{x}^2}} \right) \text{ is normal to } S.
\]

Connection: solns to "world pln 1" have constant speed + traverse paths assoc "world pln 2".

Prove: If \( x(t) \) solves pln 1, then

\[
\frac{\dot{\dot{x}}}{\sqrt{\dot{x}^2}} = 2 \langle \dot{x}, \ddot{x} \rangle = 0
\]

so speed is constant. Evidently

\[
\frac{d}{dt} \left( \frac{\dot{x}}{\sqrt{\dot{x}^2}} \right) + S
\]

so it solves pln 2.
Conversely, if path solves plan 2 then a constant-speed path clearly has \( \dot{x} + \ddot{s} \) so is extremal for plan 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 \( f > 0 \).

For plan 1: particles trajectories are extremal for the action

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

For plan 2: consider paths of "least travel time" where speed = \( \sqrt{f} \). They're extremal for

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

(note: it 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 planes is exactly as for Geodesics i.e. along solns of
plane 1, $f^2 x^1 = \text{const}$, + path is a soln of
plane 2.

Pf: Since $T = \frac{1}{2} f^2 (x^1 x^1)$, $H = \text{Legendre transform}$

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

From Hamilton's eqns

$x = -\frac{\partial H}{\partial p} = \frac{1}{f^2} p$, \quad $p = \frac{\partial H}{\partial x}$

From $1^{st}$ eqn and constancy of $H$,

$H = \frac{1}{2} f^2 (x^1 p^2) = \frac{1}{2} f^2 (x^1 x^1) = \text{const \ w/}$

Now $x(t)$ extremal for plane 1 $\Rightarrow \frac{\partial}{\partial t} (f^2 x^1) + \frac{\partial}{\partial x^1} f^2 x^1 = 0$

$\Rightarrow \frac{\partial}{\partial t} \left( \frac{f^2 x^1}{f^2 x^1} \right) = \nabla f \cdot x^1$ (since denominator

is constant)

$\Rightarrow \frac{\partial}{\partial t} \left( \frac{f x^1}{x^1} \right) = \nabla f \cdot x^1$

$\Rightarrow x(t)$ is extremal for plane 2.
Conversely, if $x(t)$ is extremal for $\frac{d}{dt}$, then

$$\frac{d}{dt} \left( \frac{x}{|x|} \right) = \frac{1}{|x|} \frac{d}{dt} |x|$$

so a path at $\frac{x}{|x|} = \text{constant}$ will have

$$\frac{d}{dt} \left( \frac{x^2}{|x|} \right) = \frac{1}{|x|} \frac{d}{dt} |x|^2$$

This being extremal for $\frac{d}{dt}$.

[previous calculus extends with no essential change to]

$$T = \int \frac{1}{2} \sum a_{ij}(x) \dot{x}_i \dot{x}_j \, dt$$

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

**Example 3**: what about mechanical systems with a potential? And: 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 solns.)

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 \( H = E \)

during the path

\[ \frac{1}{2} (\dot{x}^2 + V(x)) = E \]

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

Now, consider being extremal for \( L \) is

\[ -\int_{t_1}^{t_2} \left[ 2(E-V) \right]^{-1/2} \langle V, \delta x \rangle \dot{x} \, dt \]

\[ + \frac{12(E-V)}{2(E-V)} \frac{\ddot{x}}{\dot{x}} \delta \dot{x} = 0 \]

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

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

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

which is true! But in opposite direction.

is easy as usual (extremal for \( L \) \( \Rightarrow \)

with \( \text{ergo} \) at \( |\dot{x}| = \sqrt{2(E-V)} \) we get a soln of

the mechanical eqn.)