Mechanics - Lecture 9, 4/4/2018

Today's focus: the "Lagrangian" and "Hamiltonian" perspectives on classical mechanics (and some discn about the insights they provide)

Lagrangian perspective first.

Big picture: in solid mechanics configs in elastic equilibrium were critical pts of an "elastic energy" functional. In mechanics we are also led to consider var'l principles, but now they involve paths (functions of time) taking values in $\mathbb{R}^{6N}$ (for N interacting particles in $\mathbb{R}^3$) and the crit pts are the paths that solve eqns of classical mechanics.

Some warmup examples:

① arc length, for paths in $\mathbb{R}^n$:

$$\text{If } y(t) \in \mathbb{R}^n, \text{ arc length} = \int_{t_0}^{t_1} |\dot{y}(t)|\, dt$$

Suppose $y(t_0)$ + $y(t_1)$ are fixed. Then (assuming $\dot{y} \neq 0$)

$$\text{crit pt} \iff \frac{d}{dt}\left(\frac{\dot{y}}{|\dot{y}|}\right) = 0$$

$\iff$ path traverses a straight line.

(2) arc length, for paths on sphere $|y(t)| \equiv 1$.

Euler-Lagr eqn is different: it acquires a "Lagr mult" for the constraint:

$$\delta \int_{t_0}^{t_1} |\dot{y}| + \tfrac{1}{2} g(t)(|y|^2 - 1)\, dt = 0$$

$$\Rightarrow \quad \frac{d}{dt}\left(\frac{\dot{y}}{|\dot{y}|}\right) = g(t)\, y(t)$$

<u>Note</u>: rate of change of velocity is <u>normal</u> to sphere at position $y(t)$.

<u>Also note</u>: a path can (wlog) be parametrized by arclength. Then $|\dot{y}| = 1$ and previous relation becomes

$\ddot{y}$ is normal to sphere at $y(t)$.

(3) Does it matter what coord system we use? Of course not! For example, if curve in item 2 is represented in local coords (for ex. spherical coords θ, φ) on $S^2 \subset \mathbb{R}^3$) — let's write $\vec{\xi}(t) = (\xi_1(t), \cdots, \xi_{n-1}(t))$ for a general coord

patch on $S^{n-1} \subset \mathbb{R}^n$ then

$$|\dot{y}| = \left(\sum_{i,j}' g_{ij}(\xi(t))\, \dot{\xi}_i \dot{\xi}_j\right)^{1/2}$$

where g_{ij} is the "metric tensor". Now there's no constraint so EL eqn becomes

$$\frac{\partial F}{\partial \xi_j} - \frac{d}{dt}\frac{\partial F}{\partial \dot{\xi}_j} = 0 \quad , \quad F(\xi, \dot{\xi}) = \left(\sum' g_{ij}(\xi)\dot{\xi}_i \dot{\xi}_j\right)^{1/2}$$

We see here several features:

- var'l viewpt makes inclusion of constraints easy
- var'l viewpt makes it easy to write EL eqn in any coord system (by starting from the var'l pbm — avoiding an orgy of changes of vars)
- critical pts need not be minima (consider the sphere)

OK, now back to mechanics: our "Newtonian" eqns $\quad m_i \ddot{x}_i = -\dfrac{\partial U}{\partial x_i}$ where each $x_i \in \mathbb{R}^3$ and $U = U(x_1, \ldots, x_N)$

are identical to the EL eqn for the <u>action</u>

$$\int_{t_0}^{t_1} L(q, \dot{q})\, dt \quad \text{where } q = (x_1, \ldots, x_N) \in \mathbb{R}^{3N}$$
$$\dot{q} = (\dot{x}_1, \ldots, \dot{x}_N) \in \mathbb{R}^{3N}$$

(so $q(t)$ records all positions + $\dot{q}(t)$ records all velocities).
and

$$L(q, \dot{q}) = \underbrace{\tfrac{1}{2} \sum m_i |\dot{x}_i|^2}_{T = \text{kinetic energy}} - U(x_1, \ldots, x_N)$$

note the minus sign!

Pf is elementary (of course): EL eqn is
$\dfrac{\partial L}{\partial q_i} = \dfrac{d}{dt} \dfrac{\partial L}{\partial \dot{q}_i}$, which reduces to

$$-\frac{\partial U}{\partial x_i} = \frac{d}{dt}(m_i \dot{x}_i) = m_i \ddot{x}_i$$

(I'm assuming m_i is const in time; otherwise correct Newtonian eqn is $-\dfrac{\partial U}{\partial x_i} = \dfrac{d}{dt}(m_i \dot{x}_i)$ not $m_i \ddot{x}_i$.)

<u>Hamiltonian perspective</u> <u>now</u>: in suitable local coordinates (p_i, q_i), the paths we're interested in satisfy the ODE's

$$\dot{p}_i = -\frac{\partial H}{\partial q_i}$$

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

for some $H = H(q_1, \ldots, q_N, p_1 \ldots p_N)$. For our elementary example

$$m_i \ddot{x}_i = -\frac{\partial U}{\partial x_i}$$

the soln has this form when $p_i = m\dot{x}_i$, $q_i = x_i$, and

$$H = \underbrace{\frac{1}{2} \sum_i \frac{1}{m_i} |p_i|^2}_{\text{kinetic energy (again!)}} + \underbrace{U(q_1, \ldots, q_N)}_{\text{potential energy}}$$

(note the plus sign this time). Pf is elementary as usual:

$$p_i = m_i \dot{x}_i \Rightarrow \dot{p}_i = m_i \ddot{x}_i = -\frac{\partial U}{\partial x_i} = \frac{-\partial H}{\partial q_i}$$

$$\dot{q}_i = \dot{x}_i \text{ while } \frac{\partial H}{\partial p_i} = \frac{1}{m_i} p_i = \dot{x}_i \quad \text{(they're equal).}$$

Our tasks for the immediate future:

a) explore some results that are most easily seen in Lagrangian viewpt (eg link btwn symmetries + cons laws)

b) explore some results that are most easily seen in Hamiltonian viewpt (eg Liouville's Thm)

c) explore how the Lagrangian + Hamiltonian viewpts are connected (This will lead us to links with optimal control and Hamilton-Jacobi eqns)

<u>Disc'n of (a)</u>. We'll do this through <u>examples</u>

<u>1st example</u>. We mentioned that var'l prin is esp. useful for changing coords. So let's revisit that 2D particle in a central force field, ie $x(t) \in \mathbb{R}^2$ and $U = U(r)$. Since $\dot{x} = \dot{r} e_r + \dot{\theta} r e_\theta$,

$$L = \frac{1}{2} m (\dot{r}^2 + r^2 \dot{\theta}^2) - U(r)$$

and the EL eqn for $\int L(r, \theta, \dot{r}, \dot{\theta})\, dt$ becomes

$$\frac{\partial L}{\partial r} = \frac{d}{dt} \frac{\partial L}{\partial \dot{r}}, \quad \underbrace{\frac{\partial L}{\partial \theta}}_{0} = \frac{d}{dt} \frac{\partial L}{\partial \dot{\theta}}$$

$$\Rightarrow \quad -\frac{\partial U}{\partial r} + m r \dot{\theta}^2 = \frac{d}{dt}(m \dot{r})$$

$$0 = \frac{d}{dt}(m r^2 \dot{\theta})$$

leading quickly to the conclusion proved more laboriously before that

$$m \ddot{r} = -U'(r) + \frac{\text{const}}{r^3}$$

Essential mechanism of this calcn: L is indep of $\theta \Rightarrow$ "cons. law" $r^2\dot{\theta} = \text{constant}$

2nd example Whenever $L(q_i, \dot{q}_i)$ has no explicit dependence on time, trajectories "conserve energy" in the sense that

$$E = \left(\sum_i \dot{q}_i \frac{\partial L}{\partial \dot{q}_i}\right) - L$$

is constant in time. (Note that if $L = T - U$ with $T = \frac{1}{2}\sum a_{ij}(q)\dot{q}_i\dot{q}_j$, then $\sum \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} = 2T$, so $E = 2T - (T-U) = T + U$, which is precisely the Hamiltonian.)

Pf: EL eqn for $\int L(q,\dot q)\,dt$ is

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot q_i}\right) - \frac{\partial L}{\partial q_i} = 0$$

so (since L doesn't depend on t)

$$\frac{d}{dt}L = \sum \frac{\partial L}{\partial q_i}\dot q_i + \sum \frac{\partial L}{\partial \dot q_i}\ddot q_i$$
$$= \sum \frac{d}{dt}\left(\frac{\partial L}{\partial \dot q_i}\right)\dot q_i + \sum \frac{\partial L}{\partial \dot q_i}\ddot q_i$$
$$= \frac{d}{dt}\left(\sum \dot q_i \frac{\partial L}{\partial \dot q_i}\right)$$

3rd example Cons of lin momentum comes similarly from translation invariance, i.e. from hypoth that $x_i \to x_i + \vec a$ leaves L invariant. Indeed, if

$$\frac{d}{dt} L(x_1 + t\vec a, \dots, x_N + t\vec a, \dot x_1, \dots, \dot x_N) = 0$$

for any $\vec a$ then

$$\sum_{i=1}^{N} \frac{\partial L}{\partial x_i} = 0 \qquad \text{as vector in } \mathbb{R}^3$$

Princ of least action combines with this to give

$$\sum_{i=1}^{N} \frac{d}{dt}\frac{\partial L}{\partial \dot x_i} = 0$$

so

$$\sum_i' \frac{\partial L}{\partial \dot{x}_i} = \text{constant along trajectories}$$

4th example: Cons of angular momentum follows similarly from isotropy, ie $L(x_1, \dots, x_N, \dot{x}_1, \dots, \dot{x}_N) = L(Rx_1, \dots, R\dot{x}_N)$ for any rotn R. (For a proof see eg §9 of Landau + Lifshitz.)

5th example (really a special case of the 4th): for particles in $\mathbb{R}^3$ with forces $f_i = -\frac{\partial U}{\partial x_i}$, if U is invariant under rotn about $\vec{e}$ axis, then $\vec{e}$ component of ang momentum is conserved

Pf: $\frac{d}{dt} \sum' (x_i \times m_i \dot{x}_i) \cdot \vec{e}$

$$= \frac{d}{dt} \sum' \left(x_i \times \frac{\partial L}{\partial \dot{x}_i}\right) \cdot \vec{e}$$

$$= \underbrace{\sum' \left(\dot{x}_i \times \cancel{\frac{\partial L}{\partial \dot{x}_i}}\right) \cdot \vec{e}}_{0} + \underbrace{\sum \left(x_i \times \frac{d}{dt} \frac{\partial L}{\partial \dot{x}_i}\right) \cdot \vec{e}}_{\sum' \left(x_i \times \frac{\partial L}{\partial x_i}\right) \cdot e}$$

since $\frac{\partial L}{\partial \dot{x}_i} = m_i \dot{x}_i$

Now use fact: if $\varphi_t(x) =$ rotn of x abt $\vec{e}$ axis, then $\frac{d}{dt} \varphi_t(x) = \vec{e} \times \vec{x}$. So hypothesis on L

says

$$\sum_i \frac{\partial L}{\partial x_i} \cdot (e \times x_i) = 0$$

Now finally use vector identity $a \cdot (b \times c) = -b \cdot (a \times c)$ (they're both det (a,b,c) up to sign) to see that

$$\sum_i \left(\frac{\partial L}{\partial x_i} \times x_i\right) \cdot e = 0$$

as asserted.

(These examples are unified by a Thm of E. Noether, asserting that any cont's family of invariances of L implies a cons law for the assoc trajectories.)

What choices of $L(q, \dot{q})$ should we permit?

- Our simplest examples had $L = T - U$

$$T = \tfrac{1}{2} \sum m_i |\dot{q}_i|^2, \quad U = U(q_1, \dots, q_N)$$

- If we change vars (or introduce local coords) we naturally change this to $L = T - U$ where

$$T = \sum a_{ij}(q)\, \dot{q}_i \dot{q}_j$$

is a quadratic form (pos definite!) in $\dot{q}$ with coeffts that depend on position.

- We can of course consider the EL eqns for $\int L(q,\dot{q})\,dt$ for more general choices of L. The key structural hypothesis is that

$$\dot{q} \to L(q,\dot{q}) \quad \text{should be strictly convex}$$

(or something similar). One reason this is useful: it assures that for <u>short</u> time intervals our var'l pbm

$$\int_{t_0}^{t_1} L(q,\dot{q})\,dt \qquad \text{with } q(t_0),\, q(t_1) \text{ fixed}$$

has positive 2nd variation at any crit pt. (so trajectories are at least local minima).

<u>Explain</u>: Given a crit pt, consider 2nd var'n assoc perturbation $\eta(t)$ (nb $\eta(t_0)=0$, $\eta(t_1)=0$)

$$Q[\eta] = \int_{t_0}^{t_1} L_{qq}\,\eta\otimes\eta + 2L_{q\dot{q}}\,\eta\otimes\dot{\eta} + L_{\dot{q}\dot{q}}\,\dot{\eta}\otimes\dot{\eta}$$

where for example

$$L_{qq}\,\eta\otimes\eta = \sum \frac{\partial^2 L}{\partial q_i \partial q_j}(q,\dot{q})\,\eta_i\eta_j$$

Suppose $L_{\dot{q}\dot{q}} \geq c_0 I$ with $c_0 > 0$. Then

$$Q[\eta] \geq c_0 \int_{t_0}^{t_1} |\dot{\eta}|^2 - C_1 \int_{t_0}^{t_1} |\eta|^2 + |\eta||\dot{\eta}| .$$

from which we easily get (using $|\eta||\dot{\eta}| \leq \varepsilon |\dot{\eta}|^2 + C_\varepsilon |\eta|^2$)

$$Q[\eta] \geq \frac{c_0}{2} \int_{t_0}^{t_1} |\dot{\eta}|^2 - C_2 \int_{t_0}^{t_1} |\eta|^2$$

Now use $\int_a^b |\eta|^2 \leq \lambda (b-a)^2 \int_a^b |\dot{\eta}|^2$ to see that if $t_1 - t_0$ is small then the 2nd term is dominated by the first. Conclusion: in this setting the "extremal" $q_*(t)$ has

$$\frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \int_{t_0}^{t_1} L(q_* + \varepsilon\eta, \dot{q}_* + \varepsilon\dot{\eta})\, dt = 0$$

$$\frac{d^2}{d\varepsilon^2}\Big|_{\varepsilon=0} \int_{t_0}^{t_1} L(q_* + \varepsilon\eta, \dot{q}_* + \varepsilon\dot{\eta})\, dt \geq \text{const} \cdot \int_{t_0}^{t_1} |\dot{\eta}|^2$$

Looking ahead: I told you that "Lagrangian" + "Hamiltonian" viewpts are equivalent. So every L (convex in $\dot{q}$) should have an assoc

H. The case when

$$(*) \quad L = \frac{1}{2} \sum a_{ij}(q)\, \dot{q}_i \dot{q}_j - U(q)$$

(which arises easily from change of vars) is especially simple: Then

$$(**) \quad H = \frac{1}{2} \sum [a^{-1}(q)]_{ij}\, p_i p_j + U(q)$$

where $a^{-1}(q)$ is the matrix inverse to $a(q)$.
[Exercise: show, by direct calculation, that when L has the form $(*)$ and $q(t)$ solves its EL eqns, then $q(t)$ and $p_i(t) = \sum a_{ij}(q)\,\dot{q}_j(t)$ solve the Hamiltonian eqns

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

More generally: H is the "Legendre transform" of L

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

$= \langle p, \dot{q} \rangle - L(q, \dot{q})$ where $\dot{q}$ is the unique sol'n of

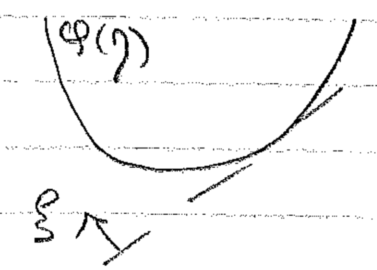
$$\frac{\partial L}{\partial \dot{q}}(q, \dot{q}) = p.$$

But when I tell you this (even in the quadratic case) it seems like magic. We'll discuss the logic behind this (namely: action minimization) soon.

Meanwhile some intuition about Legendre transforms:

- a convex fn is a sup of linear fns; thus: for each slope we can adjust the intercept so assoc plane just touches the graph. Envelope of these planes must be the given convex fn

- in eqns: $\varphi(\eta)$ convex $\Rightarrow \varphi(\eta) = \max_{\xi} \langle \eta, \xi \rangle - \varphi^*(\xi)$



$\leftarrow$ dotted line is graph of $\langle \eta, \xi \rangle - \varphi^*(\xi)$

Formula for φ^*? Well, by convexity
$\varphi(\eta) \geq \langle \eta, \xi \rangle - \varphi^*(\xi)$ with equality for each ξ at some η

so

$\varphi^*(\xi) \geq \langle \eta, \xi \rangle - \varphi(\eta)$ with equality for each ξ at some η

so

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

Let's justify the claim made a little earlier: for any Lagrangian $L(q, \dot{q})$ (convex + superlinear at ∞ in $\dot{q}$) the assoc evolution

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0$$

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

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

Some observations first:

1) $p_i = \frac{\partial L}{\partial \dot{q}_i}$ is a well-defined fn of q and $\dot{q}$

so the choice to work with (q,p)

instead of $(q, \dot{q})$ is just a convenient change of vars. (The convexity of L wrt to $\dot{q}$ $\Rightarrow$ this change of vars is invertible.)

2) $H(p,q) = \langle p, \dot{q} \rangle - L(q, \dot{q})$ when $p = \frac{\partial L}{\partial \dot{q}}(q, \dot{q})$ by opt'y cond in defn of H (we noted this already on pg 9.13)

Now take differentials:

$$dH = \sum \frac{\partial H}{\partial p_i} dp_i + \frac{\partial H}{\partial q_i} dq_i \quad \text{by chain rule}$$

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

by note (2)

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

(Warning: in calculating $\frac{\partial H}{\partial q_i}$, p_i is held fixed; in calculating $\partial L / \partial q_i$, $\dot{q}_i$ is held fixed!)

Now finally: The Lagr trajectory $\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = \frac{\partial L}{\partial q_i}$ has property that along this curve

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

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

as asserted.

We haven't yet explained

- why Hamiltonian perspective is useful (a key feature is Liouville's thm: in the (q,p) coords the flow is volume-preserving)

- how link between Lagrangian + Hamiltonian perspectives can be made less mysterious (explanation is by way of optimal control + Hamilton-Jacobi pde)

We'll see that in the next lecture.