New topic: 1D problems

\[ \min \int_a^b F(t, u(t), \dot{u}(t)) \, dt \quad u : [a, b] \to \mathbb{R}^n \]

- geodesics, as a key example
- importance of \( F \) being convex w.r.t \( u \)
- role of 2nd variation; conjugate pts

[Students taking Mechanics will see additional examples there, associated with solving eqns of Hamiltonian mechanics by "action minimization"][3]

Reasonable source for most of this material: Jost+Li-Jost, Sections 1.1-1.3 and 2.1.

Key example: geodesics. By defn: a geodesic is a curve that (locally) minimizes arc length. In local coordinates, if the curve is \( \tilde{x}(t) \),

\[
1 \dot{x}^2 = \left| \dot{x}(t) \right|^2 \, dt = \left( \sum g_{ij}(x(t)) \ddot{x}_i \ddot{x}_j \right)^{1/2}
\]

where \( g_{ij} \) is the Riemannian metric's associated problem

\[ L = \int_a^b \left| \dot{x}(t) \right| \, dt \]
(note: we're interested in critical pts, not just minima).

Two issues:

1. This has the form \( \int F(x(t), \dot{x}(t)) \, dt \), but \( F \) is not smooth in \( \dot{x} \) near \( \dot{x} = 0 \).

2. Arc length is width of parametrization, so would plan chooses a "curve" but not any particular parametrization (thus: a dramatic but geometrically-natural failure of uniqueness).

Both issues can be fixed by considering instead the different functional

\[
E = \frac{1}{2} \int_a^b \dot{x}^2 \, dt
\]

where \( \dot{x}^2 = \sum q_i(x(t)) \dot{x}_i(t) \dot{x}_i(t) \). To see why, observe that for any parametrized curve \( \vec{x}(t) \),

\[
L[\vec{x}] \leq \sqrt{2(b-a)} \sqrt{E[\vec{x}]}
\]

(with strict inequality unless \( \dot{x}^2 \) is constant), as a consequence of

\[
\int_a^b \dot{x} \, dt \leq \left( \int_a^b \dot{x}^2 \, dt \right)^{1/2} \left( \int_a^b 1 \, dt \right)^{1/2}
\]
Thus
\[ \min \text{ value of } L \leq \sqrt{2(b-a)} \cdot \left( \min \text{ value of } E \right)^{1/2}. \]

But opposite \( \neq \) is easy: given any curve with length \( L \), its constant-speed parametrization has
\[ |\dot{x}| = \frac{L}{b-a}, \quad \text{to} \]
\[ \frac{1}{2} \int_a^b |\dot{x}|^2 \, dt = \frac{1}{2} (b-a) \frac{L^2}{(b-a)^2} = \frac{1}{2(6a)} \frac{L^2}{2} \]
Thus
\[ \left( \min \text{ value of } E \right)^{1/2} \leq \frac{1}{2(6a)} \left( \min \text{ value of } L \right) \]

**Conclusion:** minimizers of \( E \) has min length and constant speed.

(Exercise: use the EL eqn for \( E \) to give a different proof that extremals - even critical pts! - of \( E \) have constant speed, by showing that \( \frac{\dot{x}}{L} |\dot{x}|^2 = 0 \) if \( x(t) \) solves the EL eqn.)

**Key properties of geodesics:**

a) They're smooth
b) They're locally paths of shortest length
c) Globally, they may not be paths of shortest length (e.g., on a sphere the geodesics are arcs of great circles)
Rule: Discussed assumed we had a single "coordinate chart" valid along entire curve. Locally true, but not necessarily globally so. In general, must use different local charts on different parts of curve (see §22.1 of Joy/Ho-Po for detail on what this means).

Properties (a) - (c) are not special to geodesics; so it's natural to discuss them more generally, in terms of form

\[ \int_a^b F(t, u(t), \dot{u}(t)) \, dt \]

where \( u: [a, b] \to \mathbb{R}^n \). Note that EL eqn in this setting is

\[ \frac{\partial F}{\partial u_j} - \frac{d}{dt} \frac{\partial F}{\partial \dot{u}_j} = 0 \quad 1 \leq j \leq n \]

Discussion of (a) = smoothness of solutions; we clearly need some condition on \( F \), since for \( u: [7, 1] \to \mathbb{R} \)

\[ \min_{u(-1) = 0, \ u(1) = 0} \int_{-1}^{+1} (u^2 - 1)^2 \, dt \]

is solved by

\[ \therefore \]

\[ \min_{u(-1) = 0, \ u(1) = 1} \int_{-1}^{+1} (u - 1)^2 u^2 \, dt \]

is solved by
Convenient hypothesis is that $F(t,u,p)$ is smooth enough (I won't try to define minimal conditions - see Fost+Li; Fost+Li in such things) and strictly convex in $p$. The point: the EL eqn can be written as

$$\frac{\partial F}{\partial u_i} - \frac{\partial^2 F}{\partial u_i \partial t} - \sum_{k} \frac{\partial^2 F}{\partial u_i \partial u_k} u_k = \sum_{k} \frac{\partial^2 F}{\partial u_i \partial u_k} u_k,$$

which we can solve (inverting the strictly pos. def. matrix $\frac{\partial^2 F}{\partial u_i \partial u_k}$) to see that $\ddot{u}$ is bounded if $\dddot{u}$ is bounded. Higher claims can be handled similarly (differentiate eqn in $t$).

Proceeding arg is a bit sloppy, since it assumes $\dddot{u}$ exists. Let's explain why strict convexity $\Rightarrow$ it must exist. Consider

$$\phi_i(t,u,p,q) = \frac{\partial F}{\partial p_i} - q_i$$

and observe that $\ddot{p}$ solves $\phi(t,u,p,q) = 0$ iff it achieves

$$\max_{\dot{p}} \langle q, \dot{p} \rangle - F(t,u,p),$$

(Here $t, u, q$ enter only as parameters, maximizing $\dot{p}$ is unique if $F$ is strictly convex; arg has an implicit hypothesis that $F(t, u, p)$ grows faster than
linearly as $|p| \to \infty$, so optimal $p$ in preceding formula exists [$p = \infty$ is not optimal].

Implicit in these hypotheses that $\frac{\partial F}{\partial p}$ has full rank, so we can (locally) solve $p^* \Phi = 0$ for $p$ as for other vars, say

$$\frac{\partial F}{\partial p_j} = \delta_j \quad \forall j \iff p_j = \psi_j(t, u, \Phi)$$

Now, we know $\Phi = 0$ when $\Phi = \frac{\partial F}{\partial p}$. Evaluating this at $p = \bar{u}$ gives

$$\bar{u}(t) = \psi_j(t, u(t), \frac{\partial F}{\partial u}(t, u, \bar{u}))$$

RHS is differentiable (using EL eqn to know differentiability of $\frac{\partial F}{\partial u}$) so LHS is differentiable (in $t$).

Rest of these notes discuss $p$th (b) & (c) (local minimality, conjugate pts, etc).

Brief summary:

1) 2nd variation provides a convenient necessary condition for minimality.

2) Importance of convexity is visible here too: if $F$ is not convex in $p$, then 2nd var. test is sure to fail.
3) as we work in longer time intervals (e.g. $[a,b]$), with a fixed $+ b^+$ failure of local optimality can be detected by 2nd order test

(Suflit errors in optimality are also interesting of course, but they would lead us too far astray.)

**Defn of 2nd order:** given a solution $\gamma(t)$ of the EL eqn, it is natural to consider

$$\frac{d^2}{ds^2} \int_a^b F(t, u+\delta u, \dot{u}+\delta \dot{u}) \, dt$$

$$s=0$$

where $\delta \gamma(t)$ is arbitrary (except perhaps for retractor case to balance controls). This reduces to

$$Q[\gamma] = \int_a^b F_{uu} \gamma \delta \gamma + 2 F_{u\gamma} \gamma \delta \gamma + F_{\gamma \gamma} \delta \gamma \delta \gamma \, dt$$

where, for example,

$$F_{uu} \gamma \delta \gamma = \sum \frac{d^2 E}{du_i du_j} (t, u, \dot{u}) \gamma_i(t) \gamma_j(t).$$

Focusing on case when $u(a)$, $u(b)$ are fixed (so $\gamma(a) = \gamma(b) = 0$) we see that

$$u \in loc \min \Rightarrow Q[\gamma] \geq 0 \text{ for all } \gamma \text{ s.t. } \gamma(a) = \gamma(b) = 0.$$
Importance of convexity is 2-fold.

**First:** If \( F_p \geq c_0 I \) with \( c_0 > 0 \) (this is a little stronger than strict convexity) then

\[
F_p \dot{\gamma} \circ \dot{\gamma} \geq c_0 \| \dot{\gamma} \|^2
\]

and it's easy to see that \( Q \) is strictly positive if \( \sqrt{c} - a \) is small enough. (Hint: \( \| \dot{\gamma} \|^2 \leq C(\sqrt{c} - a)^2 \| \dot{\gamma} \|^2 \) if \( \gamma(a) = \gamma(b) = 0 \), with \( C \) depending on \( b - a \).

**Second:** If, for some \( t_0 \), the matrix

\[
\frac{\partial^2 F}{\partial p_i \partial p_j}(t_0, u(t_0), \dot{u}(t_0))
\]

is not \( \geq 0 \) (roughly: \( F \) is not convex in \( p \) at some point along the curve) then \( \exists \gamma \) s.t. \( Q[\gamma] < 0 \). In fact, it suffices to choose \( \xi \in \mathbb{R}^n \) s.t.

\[
\sum \frac{\partial^2 F}{\partial p_i \partial p_j}(t_0, u(t_0), \dot{u}(t_0)) \xi_i \xi_j < 0
\]

and then take \( \gamma(t) \) supported in a (small) neighborhood of \( t_0 \) at

\[
\dot{\gamma}(t) \in \{ 0, \pm \xi \}
\]

Pictorial:

\[
\begin{align*}
&x \quad \downarrow \quad t_0 \quad \downarrow b \\
&\dot{\gamma} = \begin{cases} 
\xi, & t_0 - 3 < t < t_0 \\
-\xi, & t_0 < t < t_0 + 3
\end{cases}
\end{align*}
\]
If $\epsilon$ is small enough then $Q[E]$ is strictly negative (since $\int_{\gamma} \tilde{p}$ scales like $\epsilon$ and is negative, while the other terms in $Q$ scale like $\epsilon^2$).

On long "time intervals" local optimality can be lost.

Recall example of geodesics on a sphere.

We can detect loss of optimality using the 2nd order quadratic form.

Observe that it makes sense to ask: does $E_L$ for

$$
\min_{\gamma} \int_{\gamma} \left( \dddot{x} + \ddot{x} \right) dt \\
\text{subject to:} \\
\gamma'(a) = 0 \\
\gamma'(b) = 0 \\
\gamma(t) \in \text{a non-zero sub-$g(t)$?} \quad \text{(Such $g(t)$ solves the "homogeneous" 2nd order ODE)}
$$

$$
\dddot{x} + \ddot{x} - \frac{\partial}{\partial t} \left( \dddot{x} \right) - \frac{\partial}{\partial t} \left( \ddot{x} \right) = 0
$$

and is called a "Jacobi field".) If such $g(t)$ exists, we say $b$ is conjugate to $a$.

Thus: After 1st conjugate pt, an extremal (ie a rel of $E_L$ eq) ceases to be a minimizer.
**Pf:** Let $b_0$ be conjugate to $a$, and $b > b_0$.

Try

\[ \gamma = \begin{cases} \text{nonzero Jacobian held on } (a, b_0) \\ 0 \text{ on } (b_0, b) \end{cases} \]

I claim that the 2nd variational evaluated at this $\gamma$ vanishes.

Accepting this for a moment, the rest is easy: the $\gamma$ just defined is not $C^2$, so it cannot minimize the 2nd variational (note: we have assumed that $F_{pp} > 0$ so the term $F_{pp} \gamma \delta \gamma$ in the 2nd variational makes the form is strictly convex). Thus I come other $\gamma(t)$ for which 2nd variational is negative.

To see why 2nd variational at (*) vanishes, consider

\[ \phi(t, \gamma, \delta \gamma) = F_{uu} \gamma \delta \gamma + 2 F_{up} \gamma \delta \gamma + F_{pp} \gamma \delta \gamma \]

and observe that

\[ \phi(t, \alpha \gamma, \alpha \delta \gamma) = \alpha^2 \phi(t, \gamma, \delta \gamma) \]

So (by differentiating at $\alpha = 1$)

\[ \gamma \cdot \phi_{\gamma} + \gamma \cdot \phi_{\delta \gamma} = 2 \phi \]
Integrating:

\[ \int_a^b \Phi(t, \eta, \dot{\eta}) \, dt = \int_a^b \Phi(t, \eta, \dot{\eta}) \, dt \]

\[ = \frac{1}{2} \int_a^b \eta \dot{\eta} + \dot{\eta} \dot{\eta} \, dt \]

\[ = \frac{1}{2} \int_a^b \eta \left( \dot{\eta} - \frac{\partial}{\partial t} \Phi \right) \, dt \]

using that \( \eta(a) = \eta(b) = 0 \). But because \( \eta \) is a restriction of a\( 2^n \) variational quadratic form on \([a, b] \) (giving min value 0), so it solves the ELE eqn:

\[ \eta \dot{\eta} - \frac{\partial}{\partial t} \Phi = 0 \quad \text{on } [a, b] \]

Therefore

\[ \int_a^b \Phi(t, \eta, \dot{\eta}) \, dt = 0 \]

as asserted.

---

I only proved in these notes that \( 2^n \) var is \( > 0 \) on short intervals. But it is also true that crit ps are minimizers on short intervals if \( \Phi \) is ps def.

For special case of geodesics on \( S^2 \subset \mathbb{R}^3 \),
The antipodal pt is the conjugate pt. To see why, observe that

a) for antipodal pts, there is a 1-pert family of shortest paths (great semicircles)

b) if there's a 1-pert family of minimizers \( u^\theta(t) \), then \( \gamma = \frac{d}{d\theta} u^\theta \) is a Jacobi field.

Point (a) is obvious. For (b): each \( u^\theta \) solves EL

\[
F_u(t, u^\theta, \dot{u}^\theta) = \frac{d}{d\theta} F_p(t, u^\theta, \dot{u}^\theta)
\]

differentiate \( \gamma \) w.r.t. \( \theta \) to get

\[
F_{u\gamma} \gamma + F_{u\dot{u}} \dot{\gamma} = \frac{d}{d\theta} \left( F_{u\gamma} \gamma + F_{u\dot{u}} \dot{\gamma} \right)
\]

with \( \gamma = \frac{d}{d\theta} u^\theta \). This is precisely the equation characterizing a Jacobi field (note that \( \gamma = 0 \) at end points, i.e. poles).

[More conceptual pt: value of \( \int_0^v F(t, u^\theta, \dot{u}^\theta) \) wrt initial \( \theta = 0 \), both 1st and 2nd derivatives vanish wrt \( \theta \). So (assuming \( u^\theta \) is extremal) \( \gamma = \frac{d}{d\theta} u^\theta \) achieves value 0 at the 2nd variate field. Since the min value is zero, then \( \gamma \) must be a Jacobi field.]
Suggested Exercises:

1. Show directly (using the EL eqns) that any extremal for \( \int_a^b 1 \times 1^2 \, dt \) has constant speed (see pp. 2-3).

2. Show that if \( b \) is conjugate to \( a \) then

\[
\min\left\{ \int_a^b F \, \gamma + 2 \int_a^b \dot{\gamma} \cdot \vec{F} + \int_a^b F \, \delta \dot{\gamma} \, \delta \dot{\gamma} = 0 \right. \\
\left. \gamma(a) = 0 \right. \\
\left. \gamma(b) = 0 \right. \\
\right\}
\]

3. When studying waves it is useful to consider paths that minimize "travel time," where the wave speed \( v(x) \) is a known function of location \( x \). Show that this amounts to considering geodesics in the metric

\[
g_{ij} = \frac{1}{v^2(x)} \delta_{ij}.
\]

4. In these notes we focused on Dir 6c, i.e., \( \min \int_a^b F(t, u, \dot{u}) \, dt \) subject to \( u(a) = \alpha \) and \( u(b) = \beta \) being given. Suppose instead we impose \( u(a) = \alpha \) and \( u(b) \in M \) where \( M \) is a submanifold. What end condition does the EL get at \( t = b \)? What is the proper notion of Jacobi field in this case?
(5) Show that the only critical pts of
\[ \int_a^b u_x^2 + (u - 1)^2 \]
(no boundary condition!) with non-negative 2nd variation are the "trivial ones" namely, \( u = -1 \) and \( u = +1 \). (Hint: let \( u \) be a critical point. Show that \( y = u_x \) achieves value 0 in the 2nd variation quadratic form. Then argue that this \( y \) can't be a minimizer of that quadratic form.)