Calculus of variations; GELFAND and FOMIN

$\mathcal{F}$: a family of <u>admissible</u> <u>functions</u>

$\mathcal{F} \xrightarrow{\Lambda} \mathbb{R}$: a <u>functional</u> on $\mathcal{F}$

Problem: find $\bar{f} \in \mathcal{F}$ s.t.

$$\Lambda(\bar{f}) = \min\{\Lambda(f) : f \in \mathcal{F}\}$$

- Does it exist?
- How to find it?
- What are its properties?

Related problems:

- Local, rather than global, minima (what does it mean "local" in $\mathcal{F}$?)
- If there is no ~~m~~minimum in $\mathcal{F}$, can we find a minimum "near" f ?

Examples.

(1) The closed curve in the plane $\mathbb{R}^2$ enclosing ~~a~~ fixed area, having minimal length.

(2) The surface in 3-space $\mathbb{R}^3$ having as boundary a fixed curve $\Gamma \subseteq \mathbb{R}^3$ and having minimal surface.

We restrict ourselves (for a while) to the following case.

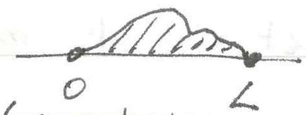
$\mathcal{F} = \{\varphi \in C^1([a,b],\mathbb{R}) : \varphi(a) = \alpha \text{ and } \varphi(b) = \beta\}$

$$\varphi \longmapsto \Lambda(\varphi) = \int_a^b F(x, \varphi(x), \dot\varphi(x))\,dx,$$

where $F \in C^2(\mathbb{R}^3, \mathbb{R})$.

Examples.

(1) Dido's problem.



Among all curves (graphs) joining $(0,0)$ and $(L,0)$ in $\mathbb{R}^2$, having fixed length $\lambda > L$, find the one enclosing the maximal area.
This is not in the above class, but it is interesting nonetheless.

$\mathcal{F} = \{\varphi \in C^1([0,L],\mathbb{R}) : \varphi(0) = \varphi(L) = 0\}$

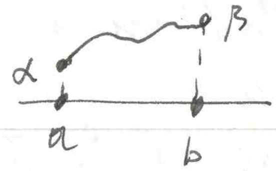
and we want to <u>maximize</u>

$$\Lambda(\varphi) = \int_0^L \varphi(x)\,dx$$

under the costraint

$$G(\varphi) = \int_0^L \sqrt{1+\dot\varphi^2(x)}\,dx = \lambda > L.$$

(2)

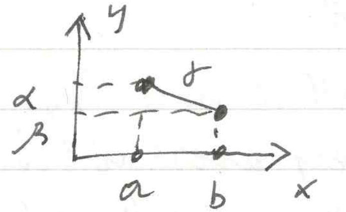


Find the shortest curve (graph) joining (a,α) and (b,β) in $\mathbb{R}^2$.

$$\Lambda(\varphi) = \int_a^b \sqrt{1+\dot\varphi^2(x)}\,dx,$$
$$\varphi(a) = \alpha;\ \varphi(b) = \beta$$

$$F(x, y, \dot y) = \sqrt{1+\dot y^2}$$

(3)



Find the curve γ joining (a,α) and (b,β) in $x > 0$, $y > 0$ s.t. the surface of revolution of γ around the y-axis has minimal ~~surface~~ area.

$$\Lambda(\varphi) = 2\pi \int_a^b x\sqrt{1+\dot\varphi^2(x)}\,dx \quad \text{(Exercise)}.$$

$$\varphi(a) = \alpha;\ \varphi(b) = \beta.$$

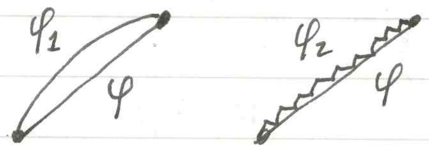
<u>Idea</u> for solving the problem:

$\nabla\Lambda(\varphi_0) = 0$ and $\operatorname{Hess}\Lambda(\varphi_0) = 0$

$\Rightarrow \varphi_0$ is a ~~[illegible]~~ local min.

What does it mean?

- We want to ~~consider~~ define a gradient of A, hence
- We have to consider limits in $\mathcal{F}$, so
- We need a distance (a measure of proximity) in $\mathcal{F}$.



Is φ_2 closer to φ than φ_1, or the other way around?

- $\text{Length}(\varphi) < \text{Length}(\varphi_1) < \text{Length}(\varphi_2)$
  $= \sqrt{2} \cdot \text{Length}(\varphi)$, but
- $\max_x |\varphi_2(x) - \varphi(x)| < \max_x |\varphi_1(x) - \varphi(x)|$

Evidence is controversial: the answer depends upon the notion of closeness we need in our problem.

We choose the nicest possible stage, that of normed linear spaces.

Linear space (over $\mathbb{R}$).

$(V, +, \cdot)$ $\quad \forall u, v \in V: u + v \in V$

$\forall u \in V \ \forall \alpha \in \mathbb{R}: \alpha u \in V$

(1) $\forall u, v, w \in V: (u+v)+w = u+(v+w)$

(2) $\forall u, v \in V: u+v = v+u$

(3) $\exists 0 \in V \ \forall u \in V: 0 + u = u$

(4) $\forall v \in V \ \exists v' := -v \in V: v + v' = 0$

(5) $\forall \alpha \in \mathbb{R} \ \forall u, v \in V: \alpha(u+v) = \alpha u + \alpha v$

(6) $\forall \alpha, \beta \in \mathbb{R} \ \forall u \in V: (\alpha+\beta) u = \alpha u + \beta u$

(7) $\forall \alpha, \beta \in \mathbb{R} \ \forall u \in V: (\alpha\beta) u = \alpha(\beta u)$

(8) $\forall u \in V: 1 \cdot u = u$.

Examples.

- $\mathbb{R}^n$
- $P[x]$: polynomials in the x-variable
- $C([a,b]) = \{\varphi : [a,b] \to \mathbb{R}, \text{ continuous}\}$
- More generally,
  $\mathbb{R}^{[a,b]} = \{\varphi : [a,b] \to \mathbb{R}\}$.

If V is a linear space and $U \subseteq V$ is a linear space, U is a (linear) subspace of V.

- $C^1([a,b])$ is a subspace of $C([a,b])$, which is a subspace of $\mathbb{R}^{[a,b]}$.

Messages

(i) Most examples are ∞-dimensional

(ii) All examples are (or can be thought of) spaces of functions

$\mathbb{R}^n \equiv \{ \varphi : \{1, 2, \dots, n\} \to \mathbb{R} \}$,

the vector $(x_1, x_2, \dots, x_n)$ can be identified with the function

$$1 \longmapsto x_1$$
$$2 \longmapsto x_2$$
$$\vdots$$
$$n \longmapsto x_n$$

(iii) Hence, functions are better thought of as points of a large space.

Exercise. Verify that $P[x]$ is not finite dimensional.

Hint. Show that the polynomials $1, x, x^2, \dots, x^n, \dots$ are linearly independent, i.e. that if $n \in \mathbb{N}$ and $\alpha_0, \dots, \alpha_n \in \mathbb{R}$,

$$\alpha_0 + \alpha_1 x + \dots + \alpha_n x^n = 0 \text{ in } P[x] \Rightarrow \alpha_0 = \dots = \alpha_n = 0.$$

Norm. Let $V = (V, +, \cdot)$ be a linear space over $\mathbb{R}$. A norm on V is a function

$$x \longmapsto \|x\|$$
$$V \longrightarrow [0, +\infty)$$

s.t. (1) $\|x\| = 0 \Leftrightarrow x = 0 \quad \forall x \in V$

(2) $\forall \alpha \in \mathbb{R} \ \forall x \in V: \|\alpha x\| = |\alpha| \cdot \|x\|$

(3) $\forall x, y \in V \Rightarrow \|x + y\| \leq \|x\| + \|y\|$

Examples.

(1) In $C([a,b])$,

$$\|\varphi\|_0 := \max_{x \in [a,b]} |\varphi(x)|$$

is a norm. (The existence of MAX is Weierstrass' Theorem).

(2) In $C^1([a,b])$, $\|\cdot\|_0$ is a norm. A better one is:

$$\|\varphi\|_1 := \max_{x \in [a,b]} |\varphi(x)| + \max_{x \in [a,b]} |\dot{\varphi}(x)|$$

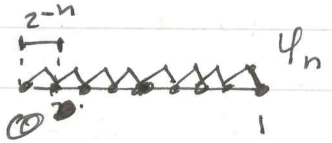
(3) Another norm on $C([a,b])$ is:

$$\|\varphi\|_{L^2} := \left(\int_a^b |\varphi(x)|^2 dx \right)^{1/2}.$$

Norms measure distances:

$\forall u, v \in V$: $\|u - v\|$ is the distance between u and v (for the norm $\|\cdot\|$).

Examples (1)



$$\|\varphi_n - 0\|_0 = \|\varphi_n\|_0 = \frac{2^{-n}}{\sqrt{2}} \xrightarrow[n\to\infty]{} 0 :$$

φ_n is close to 0 for n large in $\|\cdot\|_0$;

but $\|\varphi_n\|_1 = \frac{2^{-n}}{\sqrt{2}} + 1 \xrightarrow[n\to\infty]{} 1$:

φ_n is not close to 0, even if n is large, for the norm $\|\cdot\|_1$.

Is φ_n close to 0 in $\|\cdot\|_{L^2}$ for large n?

Estimate: $\|\varphi_n\|_{L^2} \approx 2^{-n}$,

where $A_n \approx B_n$ means that there are $0 < C < D < \infty$ s.t.

$$C \leq \frac{A_n}{B_n} \leq D$$

independently of n.

(2) $\varphi_n(x) = \sin(nx)$, $\varphi_n : [0,\pi] \to \mathbb{R}$ ($n \in \mathbb{N}$, $n \geq 1$)

$$\|\varphi_n\|_{L^2([0,\pi])} = \sqrt{\frac{\pi}{2}}$$

$$\|\varphi_n\|_1 = 1 + n \xrightarrow[n\to\infty]{} \infty$$

The nicest norms arise from inner products.

$(V, +, \cdot)$ a linear space.

$$(x, y) \longmapsto \langle x, y \rangle \quad \text{(inner product of } x \text{ and } y\text{)}$$
$$V \times V \longrightarrow \mathbb{R}$$

with the properties

(1) $\langle x, x \rangle \geq 0$ and $\langle x, x \rangle = 0 \iff x = 0$ $\forall x \in V$

(2) $\forall \alpha \in \mathbb{R}$ $\forall x, y \in V$: $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$

(3) $\forall x, y, z \in V$: $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$

(4) $\forall x, y \in V$: $\langle x, y \rangle = \langle y, x \rangle$.

Under such assumptions we have the Cauchy-Schwarz inequality:

$$\forall x, y \in V: \quad |\langle x, y \rangle| \leq \langle x, x \rangle^{1/2} \cdot \langle y, y \rangle^{1/2}.$$

Theorem. If $\langle \cdot, \cdot \rangle$ is a inner product on V, then $\|x\| := \langle x, x \rangle^{1/2}$ defines a norm on V.

Examples.

(1) On $\mathbb{R}^n$, $\langle x, y \rangle := x \cdot y$
$$= x_1 y_1 + \dots + x_n y_n$$
is a inner product, giving
$$\|x\| = \left(\sum_{j=1}^{n} x_j^2 \right)^{1/2}$$

(2) On $C([a,b])$,
$$\langle \varphi, \psi \rangle_{L^2} := \int_a^b \varphi(x)\, \psi(x)\, dx$$
defines a inner product.

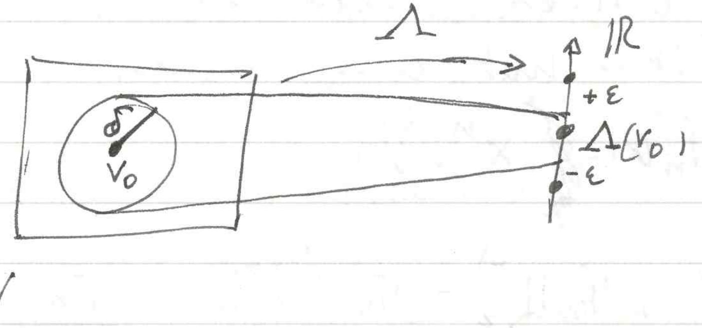
(3) On $C^1([a,b])$,
$$\langle \varphi, \psi \rangle_{W^1} := \int_a^b \varphi(x)\psi(x)\,dx + \int_a^b \dot\varphi(x)\dot\psi(x)\,dx$$
defines a inner product

(2') Let $w \in C([a,b], \mathbb{R})$, $w(x) \geq 0$ $\forall x \in [a,b]$.
$$\langle \varphi, \psi \rangle_{L^2(w)} := \int_a^b \varphi(x)\, \psi(x)\, w(x)\, dx$$

(3') Let $p, q \in C([a,b], \mathbb{R})$; $p(x) \geq 0$, $q(x) > 0$ $\forall x \in [a,b]$
$$\langle \varphi, \psi \rangle_{W^1(p,q)} := \int_a^b \varphi(x)\psi(x)\,p(x)\,dx + \int_a^b \dot\varphi(x)\,\dot\psi(x)\,q(x)\,dx.$$

Continuous functionals.

Let $(V, \|\cdot\|)$ be a normed linear space.
A functional $V \xrightarrow{\Lambda} \mathbb{R}$ is continuous at $v_0 \in V$ if
$$\forall \varepsilon > 0\ \exists \delta > 0 : \forall h \in V : \|h\| \leq \delta \Rightarrow |\Lambda(v_0 + h) - \Lambda(v_0)| \leq \varepsilon$$



Obs. In $\mathbb{R}^n$ all linear functionals
$$\mathbb{R}^n \xrightarrow{L} \mathbb{R}$$
have the form $L(x) = x \cdot a$ $(\exists a \in \mathbb{R}^n)$, hence, they are continuous.

Def. Let V be a linear space. A functional $V \xrightarrow{L} \mathbb{R}$ is linear if $\forall \alpha, \beta \in \mathbb{R}$ $\forall x, y \in V \Rightarrow L(\alpha x + \beta y) = \alpha L(x) + \beta L(y)$.

Obs. Not all linear functional are continuous in (∞-dimensional) normed vector spaces.

Example. Let $V = C([0,1])$,

$\|\cdot\| = \|\cdot\|_{L^2}$ and let

$L : C([0,1]) \to \mathbb{R}$,

$L\varphi = \varphi(1)$.

L is linear (exercise), but it is not continuous.

Let $\varphi_n(x) = x^n \cdot \sqrt[4]{n}$.

Then, $\|\varphi_n\|^2_{L^2} = \sqrt{n} \cdot \int_0^1 x^{2n}\,dx = \dfrac{\sqrt{n}}{2n+1} \xrightarrow[n\to\infty]{} 0$,

but $\varphi_n(1) = \sqrt[4]{n} \xrightarrow[n\to\infty]{} \infty$

Nomina nuda tenemus

- The study of linear ~~vector~~ spaces with norm is called (linear) functional analysis
- The study of nonnecessarily linear functionals on normed linear space is called nonlinear functional analysis.

Theorem. Let $(V, \|\cdot\|_V)$ be a normed linear space and let $V \xrightarrow{L} \mathbb{R}$ be a linear functional on V. Then, L is continuous iff L is bounded, i.e. if

$$\|L\|_{V^*} := \sup_{\substack{x\in V\\ x\neq 0}} \frac{|Lx|}{\|x\|_V} = \sup_{\substack{x\in V\\ \|x\|_V=1}} |Lx| < \infty$$

Obs. The second equality follows from L's homogeneity.

- Let $V^* = \{L : L \text{ is a continuous linear functional on } V\}$.

Then (i) V^* is a linear space

(ii) $\|\cdot\|_{V^*}$ is a norm on V.

$(V^*, \|\cdot\|_{V^*})$ is the linear space dual to V (or simply "the dual" of V).

Examples. (1) Let $V = C([a,b])$ and $\|\cdot\|_V = \|\cdot\|_0$. Let $\alpha \in C([a,b])$.

Then,

$L\varphi := \int_a^b \varphi(x)\,\alpha(x)\,dx$

defines an element $L \in V^*$.

check: $|L\varphi| = \left|\int_a^b \varphi(x)\,\alpha(x)\,dx\right|$

$\leq \int_a^b |\varphi(x)|\cdot|\alpha(x)|\,dx$

$\leq \int_a^b \|\varphi\|_0 \cdot |\alpha(x)|\,dx = \int_a^b |\alpha(x)|\,dx \cdot \|\varphi\|_0.$

i.e. $\frac{|L\varphi|}{\|\varphi\|_0} \leq \int_a^b |\alpha(x)|\,dx$, hence

$\|L\|_{V^*} \leq \int_a^b |\alpha(x)|\,dx.$

Obs. It can be proved that

$\|L\|_{V^*} = \int_a^b |\alpha(x)|\,dx$

(Can you prove it? Not a trivial exercise, not too hard either).

(2) Same $(V, \|\cdot\|) = (C([a,b]), \|\cdot\|_0)$.

$L\varphi := \varphi(a)$ defines an element of V^*:

$|L\varphi| = |\varphi(a)| \leq \max_{x\in[a,b]} |\varphi(x)| = \|\varphi\|_0$

Obs. that the same holds for

$L_{x_0}\varphi = \varphi(x_0)$

no matter what $x_0 \in [a,b]$ is.

Messages: (i) linearity depends on V only, but boundedness varies with $\|\cdot\|$. (However, completeness of $(V, \|\cdot\|)$ changes the state of affairs: see p. ....)

(ii) $\|\cdot\|_{L^2}$ does not "see" points in $[a,b]$: $\varphi \longmapsto \varphi(x_0)$ is not bounded from $C([a,b])$ $\|\cdot\|_{L^2}$ (w.r.t. $\|\cdot\|_{L^2}$)

On the contrary, $\|\cdot\|_0$ can see points, for the same/opposite reason.

(iii) An approximating algorithm which converges in the $\|\cdot\|_{L^2}$ norm will not give the correct value of functions at single points, but

(iv) $\varphi \longmapsto \int_a^b \varphi(x)\,dx$ is bounded on $C([a,b])$ w.r.t. $\|\cdot\|_{L^2}$:

the same algorithm will give the correct mean.

check it: $\left|\int_a^b \varphi(x)\,dx\right| \leq$

$\leq \left(\int_a^b dx\right)^{1/2} \cdot \left(\int_a^b |\varphi|^2 dx\right)^{1/2} = \sqrt{b-a}\cdot\|\varphi\|_{L^2}$

1^{st} variation of a functional
(Differential in ∞ dimensions).

Let $(V, \|\cdot\|)$ be a normed linear space and let

$\Lambda : V \to \mathbb{R}$ be a functional, let $\varphi_0 \in V$. Λ is differentiable at φ_0 if

$$-\Lambda(\varphi_0) + \Lambda(\varphi_0 + h) = Lh + \varepsilon \|h\|,$$

when $\varepsilon(h) \to 0$ as $\|h\| \to 0$, and L is a bdd linear functional, $L \in V^*$.

$\left.\begin{array}{l} h \mapsto Lh \\ V \to \mathbb{R} \end{array}\right\}$ is the 1^{st} variation of Λ at φ_0.

Def. Let $(V, \|\cdot\|)$ be a normed linear space and $V \xrightarrow{\Lambda} \mathbb{R}$ be a functional on V, $\varphi_0 \in V$. φ_0 is a local maximum for Λ if $\exists \varepsilon > 0 : \forall h \in V \ \|h\| \le \varepsilon \Rightarrow \Lambda(\varphi_0 + h) \le \Lambda(\varphi_0)$.

[Similar def. for local min.]

Theorem. Let $(V, \|\cdot\|)$ be normed; $\varphi_0 \in V$; $V \xrightarrow{\Lambda} \mathbb{R}$ a functional which is differentiable at φ_0, with 1^{st} variation L.

If φ_0 is a local minimum for Λ, then

$$Lh = \delta\Lambda(h) = 0 \ \forall h \in V$$

(This is Fermat Thm. in ∞ dim.).

We apply the above right away to the simplest variational problem.

$F \in C^2(\mathbb{R}^3, \mathbb{R})$

$A = \{ y = y(x) \in C^1([a,b], \mathbb{R}) : y(a) = \alpha, y(b) = \beta \}$.

$\Lambda(y) = \int_a^b F(x, y, \dot{y})\, dx$.

Obs. that $A \subseteq C^1([a,b], \mathbb{R})$ is not a linear space: $y_1, y_2 \in A \not\Rightarrow y_1 + y_2 \in A$.

This is not a big problem: A is an affine subspace of $C^1([a,b])$.

Let $V = \{ h \in C^1([a,b]) : h(a) = h(b) = 0 \}$.

If $y_0 \in A$ and $y_1 \in C^1([a,b])$, $y_1 \in A \Leftrightarrow y_1 - y_0 \in V$, which is linear

We can work on A "as if" it were a linear space: we only have to restrict the ~~variations~~ "increment" h to the linear V.

Obs./Def. More seriously, we have to choose a norm on $C^1([a,b])$. The "natural" norm is

$$\|y\|_1 = \|y\|_{W^1} = \|y\|_0 + \|\dot y\|_0$$

introduced before

(because V is complete w.r.t. $\|\cdot\|_1$).

In this case we say that a local extremum y_0 is a weak extremum:

$\exists \varepsilon > 0$ s.t. $\forall h \in V: \|h\|_1 < \varepsilon \Rightarrow \Lambda(y_0+h) \leq \Lambda(y_0)$ (for a max).

A stronger result involves the norm $\|\cdot\|_0$, since $\|h\|_0 \leq \|h\|_1$. In this case, we talk of a strong extremum:

$\exists \varepsilon > 0$ s.t. $\forall h \in V: \|h\|_0 < \varepsilon \Rightarrow \Lambda(y_0+h) \leq \Lambda(y_0)$

y_0 is a strong ext. $\Rightarrow$ y_0 is a weak ext., because $\|\cdot\|_0 \leq \|\cdot\|_1$ (Exercise).

Theorem. If y is a weak extremum, then y satisfies Euler-Lagrange eq.:

$$\text{(EL)} \quad \partial_y F(x, y(x), \dot y(x)) - \frac{d}{dx} \partial_{\dot y} F(x, y(x), \dot y(x)) = 0.$$

Obs. (EL) is (formally!) a 2nd order differential equation, whose solutions are called extremals for the variational problem (they correspond to critical points in finite dimensions).

Proof. Assume $h \in C^1([a,b])$, $h(a) = h(b) = 0$.

$$\Lambda(y+h) - \Lambda(y) = \int_a^b [F(x, y+h, \dot y + \dot h) - F(x, y, \dot y)]\, dt$$

$$= \int_a^b \left[F_y(x,y,\dot y) h + F_{\dot y}(x,y,\dot y)\dot h + \varepsilon(h,\dot h)(h^2+\dot h^2)^{1/2}\right] dt$$

where $\varepsilon(p,q) \xrightarrow[(p,q)\to 0]{} 0$ [which means: $\forall \varepsilon > 0\ \exists \delta > 0: (p^2+q^2)^{1/2} < \delta \Rightarrow \varepsilon(p,q) < \varepsilon$]

$$= \int_a^b [F_y(x,y,\dot y) h + F_{\dot y}(x,y,\dot y)\dot h]\, dt + \text{Error}$$

About $\text{Error} = \int_a^b (h^2 + \dot h^2)^{1/2} \cdot \varepsilon(h, \dot h)\, dt$, the integrand is continuous and $\|h\|_1^{-1} \cdot \text{Error} \xrightarrow[\|h\|_1 \to 0]{} 0$: $\text{Error} = o(\|h\|_1)$

Since $Lh = \int_a^b (F_y \cdot h + F_{\dot y} \dot h)\,dt$ is linear, $Lh = \delta\Lambda(h)$. Integrating by parts: (*)

$$Lh = \int_a^b F_y h\,dt + [F_{\dot y} h]_a^b - \int_a^b \frac{d}{dt}(F_{\dot y}) \cdot h\,dt$$

$$= \int_a^b \left(F_y - \frac{d}{dt} F_{\dot y}\right) h\,dt, \text{ since } h(a) = h(b) = 0.$$

We have found the 1st variation.

Since y is a (weak) extremum, $\forall h$ (admissible) we have $0 = \int_a^b h g\,dt$,

with $g = F_y - \frac{d}{dt} F_{\dot y}$.

We are left with the task of proving:

$$0 = \int_a^b h(t) g(t)\,dt \quad \forall h \in C^1([a,b]) : h(a) = h(b) = 0$$

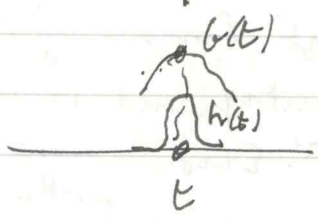
and knowing that $g \in C([a,b])$,

we must have $g \equiv 0$ in $[a,b]$.

In fact, $0 = \int_a^b h(t) g(t)\,dt = -\int_a^b \dot h(t) G(t)\,dt$,

with $G(t) = \int_a^t g(s)\,ds$, hence $G \equiv$ const

hence $g \equiv 0$



(*) We can integrate by parts if $F_{\dot y} \in C^1$, but we only know that $F_{\dot y} \in C$, a priori. But if $p, q \in C([a,b])$ and $\forall h \in C^1([a,b])$ s.t. $h(a) = h(b) = 0$ we have $0 = \int_a^b (ph + q\dot h)\,dt =$ (if $\dot P = p$)

$$= (Ph)_a^b + \int_a^b (-P + q)\dot h\,dt =$$

$$= \int_a^b (-P + q)\dot h\,dt \Rightarrow q = P + c \in C^1; \text{ as before}$$

i.e. $F_{\dot y} \in C^1$ as wished ☒