Variational Calculus - Week 1 - The Arclength Functional and Geodescis

Antonio León Villares

September 2022

Contents

Based on the notes by Jelle Hartong, Chapter 2

1 Geodesics and the Method of Finite Differences

1.1 The Geodesic Problem and Functionals

- **What is a geodesic?**
  - consider 2 points $\underline{P}, \underline{Q} \in \mathbb{R}^2$
  - the **geodesic** is the **shortest path** between $\underline{P}$ and $\underline{Q}$
- **What is a functional?**
  - a mapping between a **function space** and the **real numbers**
- **What is the arclength functional?**
  - a **functional**, which given a **path** $x(t)$, returns the arclength of the path:

$$S[x] = \int_0^1 \|\dot{x}(t)\| dt$$

$$x : [0, 1] \to \mathbb{R}^2$$

  - if $\mathcal{C}_{\underline{P},\underline{Q}}$ is the space of all paths between $\underline{P}$ and $\underline{Q}$, then:

$$S : \mathcal{C}_{\underline{P},\underline{Q}} \to \mathbb{R}$$

- **How can we find the geodesic using the arclength functional?**
  - finding a **geodesic** between $\underline{P}, \underline{Q}$ is equivalent to finding the **minimum** of S on $\mathcal{C}_{\underline{P},\underline{Q}}$
  - however, this isn't simple: functionals can be thought of as functions of infinitely many variables (after all their domain is an infinite dimensional space), so our normal calculus won't work

1.2 Euler's Method of Finite Differences

*Euler tried solving this problem by using **finite differences**:*

1. *Partition* $[0,1]$ *using* N *endpoints:*

$$t_n = \frac{n}{N}$$

2. *Approximate* $x : [0,1] \to \mathbb{R}$ *by using line segments from* $x_{n-1} = x(t_{n-1})$ *to* x_n

3. *The arclength of this piecewise curve is:*

$$S_N(x_1, \dots, x_{N-1} = \sum_{n=1}^{N} \|x_n - x_{n-1}\|$$

4. S_N *is a function of several variables, so can be optimised with calculus*

However, we would require $N \to \infty$ *(so optimising a functional involves optimising a function of infinitely many variables). Instead, the **method of variations** is preferred.*

2 Extrema of Functions of Several Variables

- **What is a directional derivative?**
  -
  - consider a function:

$$f : U \to \mathbb{R}, \qquad U \subset \mathbb{R}^n$$

  - the **directional derivative** of f at $\underline{a}$ in the direction of $\underline{v}$ is given by:

$$f'(\underline{a};\underline{v}) = \lim_{t \to 0} \frac{f(\underline{a} + t\underline{v}) - f(\underline{a})}{t}$$

  - if we define:

$$g(t) = f(\underline{a} + t\underline{v})$$

    then equivalently:

$$g'(0) = f'(\underline{a};\underline{v})$$

    since:

$$\begin{aligned} g'(0) &= \lim_{h \to 0} \frac{g(0+h) - g(0)}{h} \\ &= \lim_{h \to 0} \frac{f(\underline{a} + h\underline{v}) - f(\underline{a})}{h} \\ &= f'(\underline{a};\underline{v}) \end{aligned}$$

- **What is the total derivative?**

- consider a function:

$$f : U \to \mathbb{R}, \qquad U \subset \mathbb{R}^n$$

- at a point $\underline{a}$, the **total derivative** is $Df(\underline{a})$
- it is a **linear map**:

$$Df(\underline{a}) : U \to \mathbb{R}$$

- $Df(\underline{a})(\underline{v})$ gives the **directional derivative** of f at $\underline{a}$ in the direction of $\underline{v}$:

$$Df(\underline{a})(\underline{v}) = f'(\underline{a})(\underline{v}) = \frac{d}{dt}\left(f(\underline{a} + t\underline{v})\right)\bigg|_{t=0}$$

- **How can we geometrically interpret the directional derivative?**
  - the curve $\gamma(t) = \underline{a} + t\underline{v}$ represents a line in U, passing through the point $\underline{a}$, in the direction of $\underline{v}$
  - we can think of $\underline{v}$ as giving the **velocity** of a point along the curve
  - similarly, $f(\underline{a} + t\underline{v})$ is a curve in $\mathbb{R}$, which passes through $f(\underline{a})$, and has velocity $Df(\underline{a})(\underline{v})$
  - we can use the total derivative as a **linear approximation** to f at $\underline{a}$:

$$f(\underline{a} + t\underline{v}) = f(\underline{a}) + tDf(\underline{a})(\underline{v}) + o(t)$$

- **What is tangent space?**
  - if $U \subset \mathbb{R}^n$, and $a \in U$, the **tangent space** $T_{\underline{a}}U$ is an **n-dimensional**, **real vector space**
  - for example, $\underline{v} \in T_{\underline{a}}U$, since $\underline{v}$ is tangential to $\underline{a}$ along the curve γ
  - we can think of the **total derivative** as a mapping whose domain is the **tangent space**:

$$Df(\underline{a}) : T_{\underline{a}}(\underline{a}) \to \mathbb{R}$$

- **How are critical points defined, in terms of the total derivative?**
  - a **critical point** will be any point x_0 such that **any** directional derivative is 0 at that point
  - all vectors tangential to x_0 can be found in **tangent space**, so a critical point is **defined** by:

$$Df(x_0)(\underline{v}) = \frac{d}{dt}\left(f(x_0) + t\underline{v})\right)\bigg|_{t=0} = 0, \qquad \forall \underline{v} \in T_{x_0}U$$

All this gives us 3 ingredients which are useful to define a ***critical point*** *for some mapping:*

1. *The* ***point*** *($x_0 \in U \subset \mathbb{R}^n$)*
2. *The* ***mapping*** *($f : U \to \mathbb{R}$)*
3. *The* ***space*** *for varying the point (the* ***tangent space****, $T_{x_0}U$)*

Coming up with these 3 ingredients for functionals will allow us to compute critical points for the arclength functional.

3 Extrema of Functionals

3.1 From SVC to Functionals

- **What is the equivalent of a critical point in SVC for the arclength functional?**
  - the **geodesic** in the space of curves $\mathcal{C}_{\underline{P},\underline{Q}}$
- **What is the equivalent of a function in SVC for the arclength functional?**
  - the functional $S[x]$ itself:
$$S[x] : \mathcal{C}_{\underline{P},\underline{Q}} \to \mathbb{R}$$
- **How can we define a family of curves in $\mathcal{C}_{\underline{P},\underline{Q}}$?**
  - in SVC, we used the **tangent space** to move between points in space
  - in the space $\mathcal{C}_{\underline{P},\underline{Q}}$, we need some way of moving between **functions**
  - for this, we define a **family** of curves, by specifying paths using 2 parameters:
    * s: defines a curve within the space
    * t: defines a point within the curve

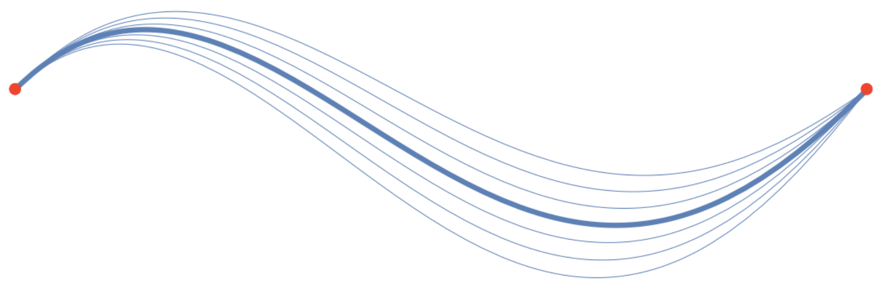


Figure 1: A family of curves $x(s,t)$. In bold is the "origin" curve, $x(0,t)$. For each value of s we have a different regular curve.

  - for this particular case, we set:
$$x(s,0) = \underline{P} \qquad x(s,1) = \underline{Q}, \qquad \forall s$$
- **What is a variation?**
  - a mapping of the form:
$$\varepsilon(t) : [0,1] \to \mathbb{R}^2$$
  - the **variation** is defined by:
$$\varepsilon(t) = \left.\frac{\partial x(s,t)}{\partial s}\right|_{s=0}$$
- **How does the variation compare to the tangent vectors?**
  - in SVC, we used the tangent vectors to a point to move between points
  - for functionals, we use the **variation** to move between curves:

– in fact, close to $s = 0$, we have:

$$x(s,t) = x(t) + s\varepsilon(t) + o(s)$$

Strictly speaking, for every fixed t, $\varepsilon(t) \in T_{x(t)}\mathbb{R}^2$; that is, it is a ***tangent vector*** *to $\mathbb{R}^2$ at $x(t)$.*
Moreover the endpoint conditions are:

$$\varepsilon(0) = 0 \in T_{\underline{P}}\mathbb{R}^2$$

$$\varepsilon(1) = 0 \in T_{\underline{Q}}\mathbb{R}^2$$

However, we can (and will) identify all the tangent spaces with $\mathbb{R}^2$ by translating them to the origin in R^2 and this is why we have written ε as a map $\varepsilon : [0,1]ß\mathbb{R}^2$ $\varepsilon(0) = \varepsilon(1) = 0$.

- **What is an endpoint-fixed variation?**
  – a variation with pre-defined endpoints
  – in the case of our functional, we impose:

$$\varepsilon(0) = \varepsilon(1) = 0$$

  – this ensures that $x(s,0) = \underline{P}$ and $x(s,1) = \underline{Q}$, since:

$$x(s,0) = x(0) + s\varepsilon(0) + o(s) = \underline{P} + s\varepsilon(0) + o(s) = \underline{P}$$

$$x(s,1) = x(1) + s\varepsilon(1) + o(s) = \underline{Q} + s\varepsilon(1) + o(s) = \underline{Q}$$

- **How can we determine a critical point to the arclength functional?**
  – in SVC, a **critical point** required:

$$\left.\frac{d}{dt}\left(f(x_0) + t\underline{v}\right)\right)\Big|_{t=0} = 0, \qquad \forall \underline{v} \in T_{x_0}U$$

  – analogously for functionals:

$$\left.\frac{d}{ds}\left(S[x + s\varepsilon]\right)\right|_{s=0} = 0, \qquad \textit{for all endpoint-fixed variations } \varepsilon(t)$$

3.2 Lemma: The Fundamental Lemma of the Calculus of Variations

Before solving the arclength problem, we present a very important - in fact, fundamental - lemma.

Let:

$$f : [0,1] \to \mathbb{R}^n$$

*be a **continuous** function which obeys:*

$$\int_0^1 \langle f(t), h(t) \rangle \, dt = 0$$

for all:

$$h : [0,1] \to \mathbb{R}^n$$

$$h(0) = h(1) = 0$$

where $h \in C^\infty$ (that is, h is infinitely continuously differentiable).
Then, $f \equiv 0$.
(Theorem 2.1)

Proof. We begin by proving this for the simpler case $n = 1$.

Let $f : [0,1] \to \mathbb{R}$ be a continuous function obeying:

$$\int_0^1 f(t)h(t)dt = 0$$

for all $h : [0,1] \to \mathbb{R}$ with $h \in C^\infty$ and $h(0) = h(1) = 0$. We claim that $f \equiv 0$.
Assume $\exists t_0 \in (0,1)$ such that $f(t_0) \neq 0$ (by continuity, if f is non-zero at the endpoints, it must be non-zero in their neighbourhood, so we can just consider $(0,1)$). We can also assume that $f(t_0) > 0$ (otherwise just apply same argument with $-f$). Furthermore, by continuity $\exists c \in (a,b) \subset (0,1)$ such that:

$$f(t) > c, \qquad \forall t \in (a,b)$$

Now, assume there exists a C^∞, non-negative function $h : [0,1] \to \mathbb{R}$, with $h(0) = h(1) = 0$, such that:

1. $h(t) = 0, \qquad \forall t \notin (a,b)$
2. $\int_0^1 h(t)dt > 0$

If such a function exists, then:

$$\begin{aligned}\int_0^1 f(t)h(t)dt &= \int_a^b f(t)h(t)dt \\ &> c\int_a^b h(t)dt \\ &> 0\end{aligned}$$

But this is a contradiction, since we claimed that the integral was 0. Hence, no such t_0 must exist, so $f(t) = 0, \forall t \in [0,1]$ as required.
We now consider th emultidimensional case. Let $f : [0,1] \to \mathbb{R}^n$, and assume that:

$$\int_0^1 \langle f(t), h(t) \rangle \, dt = 0$$

for all C^∞ functions $h : [0,1] \to \mathbb{R}^n$, with $h(0) = h(1) = 0$.

Again, assume that $\exists t_0 \in (0,1)$ such that $f(t_0) \neq \underline{0}$. Then, at least one component of f must be non-zero. Without loss of generality, assume $f^1(t_0) \neq 0$. By the work for $n = 1$, if there exists a function $h^1 : [0,1] \to \mathbb{R}$ with $h^1(0) = h^1(1) = 0$, then:

$$\int_0^1 f^1(t)h^1(t)dt \neq 0$$

Now, define:

$$h : [0,1] \to \mathbb{R}^n$$

such that:

$$h(t) = \begin{pmatrix} h^1(t) \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

h is smooth, and satisfies $h(0) = h(1) = \underline{0}$. Furthermore,

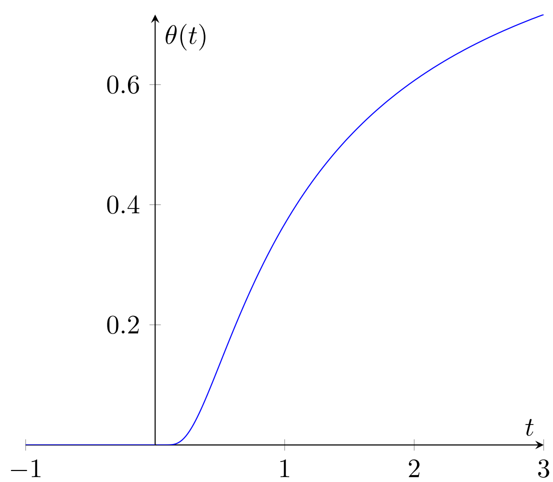
$$\int_0^1 \langle f(t), h(t) \rangle \, dt = \int_0^1 f^1(t)h^1(t)dt \neq 0$$

so no such t_0 must exist, and $f = \underline{0}, \forall t \in [0,1]$, as required.

We now consider the construction of the smooth function $h(t)$ for the case $n = 1$.
Consider:

$$\theta(t) = \begin{cases} e^{-\frac{1}{t}}, & t > 0 \\ 0, & t \leq 0 \end{cases}$$

θ is infinitely differentiable, even at $t = 0$



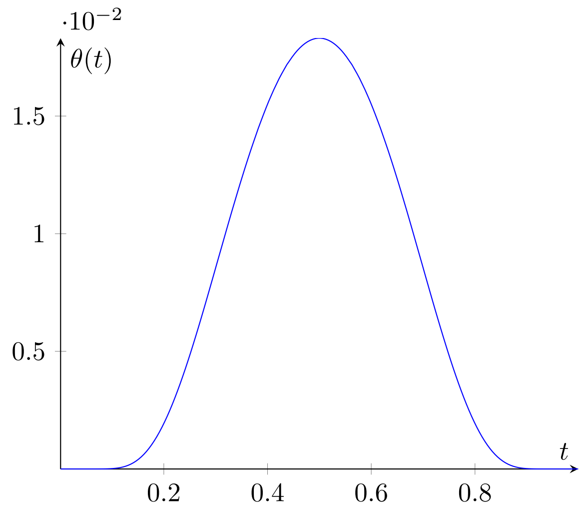
Now define a new function:

$$\phi(t) = \theta(t)\theta(1-t)$$

Notice that:

- this is a product of smooth functions, so it is smooth
- $\theta(t)$ enforces that for $t \leq 0$, $\phi(t) = 0$

- $\theta(1-t)$ enforces that for $t \geq 1$, $\phi(t) = 0$



We can then rescale ϕ, such that it vanishes outside any interval (a, b), not just $(0, 1)$:

$$\phi_{a,b}(t) = \phi\left(\frac{t-a}{b-a}\right)$$

Picking $0 < a < b < 1$, let:

$$u(t) = \frac{t-a}{b-a} \implies \frac{du}{dt} = \frac{t}{b-a}$$

Then we have:

$$\begin{aligned}\int_a^b \phi_{a,b}(t)dt &= \int_{u(a)}^{u(b)} \phi(u)(b.a)du \\ &= (b-a)\int_0^1 \phi(t)dt\end{aligned}$$

From the graph, we can see that $\int_0^1 \phi(t)dt > 0$, so it follows that:

$$\int_a^b \phi_{a,b}(t)dt > 0$$

Hence, taking $h(t) = \phi_{a,b}$ satisfies the restrictions placed.

□

3.3 Solving the Arclength Problem

The geodesic for the arclength functional is obtained by finding x such that:

$$\frac{d}{ds}\left(S[x+s\varepsilon]\right)\Big|_{s=0} = 0$$

In terms of the functional,

$$S[x+s\varepsilon] = \int_0^1 \|\dot{x}(t) + s\dot{\varepsilon}(t)\|dt$$

Using the inner product, we can write:

$$\begin{aligned}\|\dot{x}+s\dot{\varepsilon}\| &= \sqrt{\langle \dot{x}+s\dot{\varepsilon}, \dot{x}+s\dot{\varepsilon}\rangle} \\ &= \sqrt{\langle \dot{x},\dot{x}\rangle + 2s\langle \dot{x},\dot{\varepsilon}\rangle + s^2\langle \dot{\varepsilon},\dot{\varepsilon}\rangle}\end{aligned}$$

Hence:

$$\begin{aligned}\frac{d}{ds}\left(S[x+s\varepsilon]\right) &= \frac{d}{ds}\int_0^1 \sqrt{\langle \dot{x},\dot{x}\rangle + 2s\langle \dot{x},\dot{\varepsilon}\rangle + s^2\langle \dot{\varepsilon},\dot{\varepsilon}\rangle}dt \\ &= \int_0^1 \frac{d}{ds}\left(\sqrt{\langle \dot{x},\dot{x}\rangle + 2s\langle \dot{x},\dot{\varepsilon}\rangle + s^2\langle \dot{\varepsilon},\dot{\varepsilon}\rangle}\right)dt \\ &= \int_0^1 \frac{2\langle \dot{x},\dot{\varepsilon}\rangle + 2s\langle \dot{\varepsilon},\dot{\varepsilon}\rangle}{2\|\dot{x}+s\dot{\varepsilon}\|}dt \\ &= \int_0^1 \frac{\langle \dot{x}+s\dot{\varepsilon},\dot{\varepsilon}\rangle}{\|\dot{x}+s\dot{\varepsilon}\|}dt\end{aligned}$$

Evaluating at $s=0$:

$$\frac{d}{ds}\left(S[x+s\varepsilon]\right)\Big|_{s=0} = \int_0^1 \frac{\langle \dot{x},\dot{\varepsilon}\rangle}{\|\dot{x}\|}dt = \int_0^1 \left\langle \frac{\dot{x}}{\|\dot{x}\|},\dot{\varepsilon}\right\rangle dt$$

Integration by Parts for the Dot Product

Say you have a function:

$$f(t) = \langle u(t), v(t)\rangle$$

then:

$$\dot{f}(t) = \langle \dot{u}(t), v(t)\rangle + \langle u(t), \dot{v}(t)\rangle$$

If we then have:

$$\begin{aligned}\int_a^b \langle u(t),\dot{v}(t)\rangle\,dt &= \int_a^b (\dot{f}(t) - \langle \dot{u}(t), v(t)\rangle)dt \\ &= [\langle u(t), v(t)\rangle]_a^b - \int_a^b \langle \dot{u}(t), v(t)\rangle)dt\end{aligned}$$

We don't know anything about $\dot{\varepsilon}$, so we can use **integration by parts** to get rid of it:

$$\begin{aligned}\frac{d}{ds}\left(S[x+s\varepsilon]\right)\Big|_{s=0} &= \int_0^1 \left\langle \frac{\dot{x}}{\|\dot{x}\|},\dot{\varepsilon}\right\rangle dt \\ &= \left[\left\langle \frac{\dot{x}}{\|\dot{x}\|},\varepsilon\right\rangle\right]_0^1 - \int_0^1 \left\langle \frac{d}{dt}\left(\frac{\dot{x}}{\|\dot{x}\|}\right),\varepsilon\right\rangle)dt \\ &= -\int_0^1 \left\langle \frac{d}{dt}\left(\frac{\dot{x}}{\|\dot{x}\|}\right),\varepsilon\right\rangle)dt, \qquad \textit{(since } \varepsilon(0)=\varepsilon(1)=0\textit{)}\end{aligned}$$

But then, if we require that:

$$\frac{d}{ds}\left(S[x+s\varepsilon]\right)\Big|_{s=0} = 0 \implies -\int_0^1 \left\langle \frac{d}{dt}\left(\frac{\dot{x}}{\|\dot{x}\|}\right),\varepsilon\right\rangle)dt = 0$$

by the Fundamental Lemma of the Calculus of Variations, it must be the case that:

$$\frac{d}{dt}\left(\frac{\dot{x}}{\|\dot{x}\|}\right) = 0$$

In other words, the velocity vector of x must be constant, so in particular, $x(t)$ must just be a straight line, as expected.

3.4 Exercises

1. **Show that for variations ε proportional to the tangent of the original curve $x(t)$, the expression:**

$$\int_0^1 \left\langle \frac{d}{dt}\left(\frac{\dot{x}}{\|\dot{x}\|}\right), \varepsilon \right\rangle dt = 0$$

   interpret the result.

2. **Generalise the preceding discussion to paths in $\mathbb{R}^n$ between any 2 distinct points.**

3. **What is the shortest path in the plane from the origin to the line $x = 1$? Solve this problem using the variational calculus, but notice that variations are not necessarily fixed at one endpoint.**