Calculus of Variations 5: Geodesics on Surfaces in $\mathbf{R}^3$.

1. As a first application of the calculus of variations in the several functions, single variable setting, we consider geodesics on surfaces in $\mathbf{R}^3$. We'll need a little more multivariable calculus. Denote by $\mathbf{R}^2_{uv}$ and $\mathbf{R}^3_{xyz}$ the uv-plane and the three-dimensional xyz-space. Let X, Y and Z be real-valued functions defined on the uv-plane $\mathbf{R}^2_{uv}$. We set

$$x = X(u, v), \quad y = Y(u, v), \quad \text{and} \quad z = Z(u, v). \tag{1}$$

Let $T : \mathbf{R}^2_{uv} \mapsto \mathbf{R}^3_{xyz}$ by

$$T(u, v) = (X(u, v), Y(u, v), Y(u, v)). \tag{2}$$

The function T will map a blob $\mathcal{B}$ in $\mathbf{R}^2_{uv}$ to a surface $\mathcal{S}$ in $\mathbf{R}^3_{xyz}$. We say that $\mathcal{S}$ is the *image* of $\mathcal{B}$ under T. Define the partial derivatives

$$T_u = (X_u, Y_u, Z_u) \quad \text{and} \quad T_v = (X_v, Y_v, Z_v).$$

Through T, horizontal and vertical lines on the uv-plane are mapped to curves on $\mathcal{S}$. T_u and T_v are the tangent vectors to those curves. We set

$$E = T_u \cdot T_u = X_u^2 + Y_u^2 + Z_u^2,$$

$$F = T_u \cdot T_v = X_u X_v + Y_u Y_v + Z_u Z_v,$$

and

$$G = T_v \cdot T_v = X_v^2 + Y_v^2 + Z_v^2.$$

Note that E, F and G are functions of u and v.

2. **Example**: We set

$$X(u, v) = u, \quad Y(u, v) = v \quad \text{and} \quad Z(u, v) = u^2 + v^2.$$

Then T maps $\mathbf{R}^2_{uv}$ to the paraboloid $z = x^2 + y^2$ in $\mathbf{R}^3_{xyz}$. For this surface,

$$E = 1 + 4u^2, \quad F = 4uv \quad \text{and} \quad G = 1 + 4v^2. \tag{3}$$

3. **Example**: Let $\mathcal{B}$ be the rectangle $0 \le u \le 2\pi$, $0 \le v \le \pi$, and

$$X(u, v) = \cos u \sin v, \quad Y(u, v) = \sin u \sin v \quad \text{and} \quad Z(u, v) = \cos v.$$

Then $T(u, v) = (X(u, v), Y(u, v), Y(u, v))$ maps $\mathcal{B}$ to the sphere $\mathcal{S}$ of radius 1 centered at the origin. In this case,

$$E = \sin^2 v, \quad F = 0 \quad \text{and} \quad G = 1. \tag{4}$$

4. **Example**: Let $\mathcal{B}$ be the strip $0 \le u \le 2\pi$, $-\infty < v < \infty$, and

$$X(u,v) = \cos u, \quad Y(u,v) = \sin u \quad \text{and} \quad Z(u,v) = v. \tag{5}$$

Then $T(u,v) = (X(u,v), Y(u,v), Z(u,v))$ maps $\mathcal{B}$ to the cylinder $\mathcal{S}$ of radius 1 centered on the z. In this case,

$$E = 1, \quad F = 0 \quad \text{and} \quad G = 1. \tag{6}$$

5. Just as the surface $\mathcal{S}$ is the image under T of a blob $\mathcal{B}$, curves on $\mathcal{S}$ are the images under T of curves in $\mathcal{B}$. By taking this view, we are able to reduce calculations about complicated objects (curves on surfaces in space) to calculations about simpler ones (curves on the plane). Suppose for example, that you want to find the geodesics on $\mathcal{S}$ between points P and Q. We state the problem as:

a. Find the shortest curve on $\mathcal{S}$ from P to Q.

Let α and β be points in $\mathcal{B}$ such that

$$T(\alpha) = P \quad \text{and} \quad T(\beta) = Q. \tag{7}$$

If ξ is a curve in $\mathcal{B}$ from α to β, then its image $T(\xi)$ is a curve on $\mathcal{S}$ from P to Q. We can therefore replace (a) with

b. Among all curves in $\mathcal{B}$ from α to β, find the one whose image under T is shortest.

Why consider (b) rather than (a)? Because curves on planes are easier to work with than curves on surfaces. How we formulate (b) mathematically is the subject of the next paragraph.

6. Let ξ be a curve in $\mathcal{B}$ from α to β. In other words, ξ is a function mapping some interval $[a,b]$ to $\mathcal{B}$ by

$$\xi(t) = (u(t), v(t)). \tag{8}$$

where $u(t)$ and $v(t)$ are smooth, real-valued functions on $[a,b]$ with

$$\xi(a) = (u(a), v(a)) = \alpha \quad \text{and} \quad \xi(b) = (u(b), v(b)) = \beta. \tag{9}$$

As t increases from a to b, $\xi(t)$ traces a curve in $\mathcal{B}$ from α to β. Its image $T(\xi(t))$ traces a curve on $\mathcal{S}$ from P to Q. Let s be arclength along the latter curve.

$$\begin{aligned} ds^2 &= dx^2 + dy^2 + dz^2 \\ &= (X_u du + X_v dv)^2 + (Y_u du + Y_v dv)^2 + (Z_u du + Z_v dv)^2 \\ &= E\,du^2 + 2F\,du dv + G\,dv^2 \\ &= \left(E\,\dot{u}^2 + 2F\,\dot{u}\dot{v} + G\,\dot{v}^2\right) dt^2. \end{aligned} \tag{10}$$

Therefore,

$$ds = \sqrt{E\,\dot{u}^2 + 2F\,\dot{u}\dot{v} + G\,\dot{v}^2}\,dt. \tag{11}$$

It follows that the length of the curve traced on $\mathcal{S}$ by $T(\xi(t))$ is

$$J(\xi) = J(u, v) = \int_a^b \sqrt{E\,\dot{u}^2 + 2F\,\dot{u}\dot{v} + G\,\dot{v}^2}\,dt. \tag{12}$$

Remember that E, F and G depend on u and v, so the Lagrangian is

$$L(\xi, \dot{\xi}) = \sqrt{E\,\dot{u}^2 + 2F\,\dot{u}\dot{v} + G\,\dot{v}^2}.$$

Remember also that J acts on ξ by returning the length of its *image* on $\mathcal{S}$. So version (b) requires that we minimize J over a domain $\mathcal{D}$ of smooth curves ξ which satisfy the boundary conditions (7).

7. We go about the minimization in the usual way. The class $\mathcal{A}$ of admissible variations consists of smooth curves $\gamma(t) = (\varrho(t), \sigma(t))$ in $\mathcal{B}$ that vanish at a and b. The extremals of J are the functions $\xi \in \mathcal{D}$ for which

$$\delta J(\xi, \gamma) = 0, \tag{13}$$

for all γ in $\mathcal{A}$. The condition (13) leads to the system of Euler equations,

$$\begin{cases} \dfrac{\partial L}{\partial u} - \dfrac{d}{dt}\dfrac{\partial L}{\partial \dot{u}} = 0, \\[2ex] \dfrac{\partial L}{\partial v} - \dfrac{d}{dt}\dfrac{\partial L}{\partial \dot{v}} = 0. \end{cases} \tag{14}$$

We have to solve the system (14) subject to the boundary conditions (7).

8. **Note**: The specific choice of interval $[a, b]$ is a matter of convience. If $\xi(t)$ traces a curve on $\mathcal{B}$ for $a \le t \le b$, with $\xi(a) = \alpha$ and $\xi(b) = \beta$, then

$$\xi_1(t) = \xi(tb + (1-t)a), \quad 0 \le t \le 1,$$

traces the same curve. We may thus replace any curve ξ with ξ_1, and the interval $[a, b]$ with $[0, 1]$.

9. **Example**: Let $\mathcal{S}$ be the cylinder of radius 1 about the z-axis, as given by (5). In this case the Lagrangian is

$$L = \sqrt{\dot{u}^2 + \dot{v}^2},$$

Since L does not depend on u and v, $L_{\dot{u}}$ and $L_{\dot{v}}$ are first integrals. So for an extremal $\xi(t) = (u(t), v(t))$,

$$L_{\dot{u}} = \frac{\dot{u}}{L} = C_1 \quad \text{and} \quad L_{\dot{v}} = \frac{\dot{v}}{L} = C_2, \tag{15}$$

where C_1 and C_2 are constants. We can use this pair of equations to eliminate the parameter t.

$$\frac{dv}{du} = \frac{\dot{v}}{\dot{u}} = \frac{C_2}{C_1} = \text{constant}. \tag{16}$$

It follows that the extremals of J are line segments in the strip $\mathcal{B}$. From (5), you can see that if the extremal is a vertical line segment, its image on $\mathcal{S}$ will also be a vertical line segment. If the extremal is a horizontal line segment, its image will be an arc on the cylinder, parallel to the xy-plane. Finally, if the extremal lies along the line $v = Au + B$, with $A \neq 0$, its image on $\mathcal{S}$ will be will be a section of the spiral

$$x = \cos u, \quad y = \sin u, \quad z = Au + B. \tag{17}$$

a. Let $P = (1, 0, 0)$ and $Q = (-1, 0, 2)$. Which smooth curve from P to Q on $\mathcal{S}$ is shortest? Since $T(0, 0) = P$ and $T(\pi, 2) = Q$, the extermal is the line segment from $(0, 0)$ to $(\pi, 2)$ in the uv-plane. The image of this segment under T is the section of the spiral

$$x = \cos u, \quad y = \sin u, \quad z = (2/\pi)u, \tag{18}$$

with $0 \leq u \leq \pi$. This is plausibly the geodesic from P to Q.

b. Let $P = (1, 0, 0)$ and $Q = (1, 0, 3)$. Since $T(0, 0) = P$ and $T(0, 3) = Q$, the extermal is the line segment from $(0, 0)$ to $(0, 3)$ in the uv-plane. The image of this segment under T is the vertical line segment

$$x = 1, \quad y = 0, \quad z = v, \tag{19}$$

with $0 \leq v \leq 3$. This is clearly the geodesic from P to Q.