

7.1 Let (\mathcal{M}, g) be a Riemannian manifold and $p \in \mathcal{M}$. Let (x^1, \dots, x^n) be normal coordinates centered at p . Show that the components of the metric g satisfy at $p = (0, \dots, 0)$ the following cyclic identity for any $i, j, k, l \in \{1, \dots, n\}$:

$$\partial_i \partial_j g_{kl}|_p + \partial_j \partial_k g_{il}|_p + \partial_k \partial_l g_{ij}|_p = 0.$$

(Hint: recall that the Gauss lemma is equivalent to the statement that, in normal coordinates, $g_{ij}x^j = \delta_{ij}x^j$. Differentiate this relation a few times and evaluate at $(0, \dots, 0)$.)

Manipulating the above formula, show that

$$\partial_i \partial_j g_{kl}|_p = \partial_k \partial_l g_{ij}|_p.$$

Solution. Let (x^1, \dots, x^n) be normal coordinates centered at p . As we have seen in class, the Gauss lemma implies that, in these coordinates, we have at every point in the coordinate chart

$$g_{ij}x^j = \delta_{ij}x^j$$

(this is merely a reformulation of the statement that, for any $v \in \Omega_p \subset T_m \mathcal{M}$ and $w \in T_v T_p \Omega$, we have $g_p(v, w) = g(d\exp_p|_v(v), d\exp_p|_v(w))$; note that, in normal coordinates around p associated to the choice of an orthonormal frame $\{e_i\}_{i=1}^n$ at $T_p \mathcal{M}$, so that $\partial_i|_p = e_i$, the coordinate expression of the exponential map $\exp_p : \Omega_p \subset T_p \mathcal{M} \rightarrow \mathcal{M}$ is the identity map, and the radial vector $v \in T_v T_p \Omega_p$ is mapped to the vector $x^i \partial_i$).

Differentiating the above relation three times (and noting that δ_{ij} is constant in x), we obtain:

$$\partial_k \partial_l \partial_m (g_{ij}x^j) = \partial_k \partial_l \partial_m (\delta_{ij}x^j) \stackrel{\partial_a x^b = \delta_a^b}{\iff} \partial_k \partial_l \partial_m g_{ij} \cdot x^j + \partial_k \partial_l g_{ij} \delta_m^j + \partial_k \partial_m g_{ij} \delta_l^j + \partial_l \partial_m g_{ij} \delta_k^j = 0.$$

Evaluating the above at $p = (0, \dots, 0)$ (where the first term vanishes) and using the fact that $\partial_a \partial_b g_{cd}$ is symmetric in (a, b) and in (c, d) we obtain for all $i, k, l, m \in \{1, \dots, n\}$

$$\partial_k \partial_l g_{mi}|_p + \partial_l \partial_m g_{ki}|_p + \partial_m \partial_k g_{li}|_p = 0.$$

Applying the above formula successively, we have for any $a, b, c, d \in \{1, \dots, n\}$:

$$\begin{aligned} \partial_a \partial_b g_{cd}|_p &= -\partial_c \partial_a g_{bd}|_p - \partial_b \partial_c g_{ad}|_p \\ &= \partial_d \partial_c g_{ba}|_p + \partial_a \partial_d g_{bc}|_p + \partial_d \partial_b g_{ac}|_p + \partial_c \partial_d g_{ab}|_p \\ &= 2\partial_c \partial_d g_{ab}|_p + \partial_a \partial_d g_{bc}|_p + \partial_d \partial_b g_{ac}|_p \\ &= 2\partial_c \partial_d g_{ab}|_p - \partial_a \partial_b g_{dc}|_p \end{aligned}$$

from which we infer after moving the last term on the right hand side to the left hand side:

$$2\partial_a \partial_b g_{cd}|_p = 2\partial_c \partial_d g_{ab}|_p,$$

which is the required identity.

7.2 In this exercise, we will compute the expression in polar coordinates of the three model geometries in 2 dimensions (this was originally part of the last exercise last week).

- As a warm up, express in polar coordinates centered at the origin the flat metric g_E on \mathbb{R}^2 .
- Let $(\mathbb{H}^2, g_{\mathbb{H}})$ be the hyperbolic plane (see Exercise 6.4 for an expression of the metric in the Poincaré disc model, when \mathbb{H}^2 is identified with the interior of the unit disc). Let p be a point in the hyperbolic plane. Compute the metric in polar coordinates around p . *(Hint: Working in the Poincaré disc model, it suffices to only consider the case when p is at the origin, since any point $p \in \mathbb{D}^2$ can be mapped to any other point in \mathbb{D}^2 via an isometry. What are the geodesics in $(\mathbb{D}^2, g_{\mathbb{D}})$ emanating from the origin?)*
- How is the round metric $(\mathbb{S}^2, g_{\mathbb{S}^2})$ expressed in polar coordinates around a point $p \in \mathbb{S}^2$?
- In all of the three Riemannian surfaces (S, g) considered above, compute the volume of the metric ball of radius $r > 0$ centered at any point $p \in (S, g)$ (due to the symmetry of the above spaces, the precise choice of p is irrelevant). Denoting with $B_{(S,g)}[r]$ the corresponding ball, show that

$$\text{Vol}(B_{(\mathbb{S}^2, g_{\mathbb{S}^2})}[r]) < \text{Vol}(B_{(\mathbb{R}^2, g_E)}[r]) < \text{Vol}(B_{(\mathbb{H}^2, g_{\mathbb{H}^2})}[r]).$$

Solution. (a) As we have seen in class, the usual Cartesian coordinates (x, y) on \mathbb{R}^2 are also normal coordinates around $(0, 0)$ for g_E (since the exponential map $\exp_0 : T_0 \mathbb{R}^2 \rightarrow \mathbb{R}^2$ in these coordinates is simply the identity). Therefore, the usual polar coordinates (r, θ) on $\mathbb{R}^2 \setminus \{0\}$, which are defined by the relations $x = r \cos(\theta)$ and $y = r \sin(\theta)$ are also the “Riemannian” polar coordinates. Using the usual transformation formula for the metric $g_E = dx^2 + dy^2$ under changes of coordinates, we have seen that

$$g_E = dr^2 + r^2 d\theta^2.$$

(b) As we saw in the exercises last week, the set of isometries of $(\mathbb{H}^2, g_{\mathbb{H}})$ contains all maps of the form $z \rightarrow \frac{az+b}{cz+d}$, $ad - bc > 0$; therefore, for any $p_1, p_2 \in \mathbb{H}^2$, there exists an isometry $F : (\mathbb{H}^2, g_{\mathbb{H}}) \rightarrow (\mathbb{H}^2, g_{\mathbb{H}})$ such that $F(p_1) = p_2$ (i.e. $(\mathbb{H}^2, g_{\mathbb{H}})$ is *homogeneous*). As a result, the metric $g_{\mathbb{H}}$ expressed in polar coordinates around a point $p \in \mathbb{H}^2$ will have the same form independently of the chosen point p . For this reason, we can choose to work with the point corresponding to the origin in $(\mathbb{D}^2, g_{\mathbb{D}})$.

Let us use the notation (x, y) and $(\bar{r}, \bar{\theta})$ for the standard Cartesian and radial coordinates, respectively, on \mathbb{R}^2 (so that $\bar{r}^2 = x^2 + y^2$ and $\tan \bar{\theta} = \frac{y}{x}$). In the (x, y) coordinate system, the tangent vectors $e_1 = \frac{\partial}{\partial x} \Big|_p$ and $e_2 = \frac{\partial}{\partial y} \Big|_p$ constitute an orthonormal basis of $T_p \mathbb{D}^2$ with respect to $g_{\mathbb{D}}|_p$ (since $(g_{\mathbb{D}})_{ij}|_p = \delta_{ij}$). Therefore, we can use the coordinates on $T_p \mathbb{D}^2$ with respect to (e_1, e_2) to construct a normal coordinate system in a neighborhood of $p = (0, 0)$ in $(\mathbb{D}^2, g_{\mathbb{D}})$ via the map \exp_p ; we will use the notation (x^1, x^2) for this coordinate system and (r, θ) for the associated polar coordinates (so that $r^2 = (x^1)^2 + (x^2)^2$ and $\tan \bar{\theta} = \frac{x^2}{x^1}$). Notice that, since $e_1 = \frac{\partial}{\partial x} \Big|_p$ and $e_2 = \frac{\partial}{\partial y} \Big|_p$, we have

$$\frac{\partial}{\partial x^1} \Big|_p = \frac{\partial}{\partial x} \Big|_p \quad \text{and} \quad \frac{\partial}{\partial x^2} \Big|_p = \frac{\partial}{\partial y} \Big|_p. \quad (1)$$

Moreover, in the (r, θ) coordinate system, the curves $\theta = \text{const}$ correspond to geodesic rays emanating from p . Recall that, as we saw in class, the metric $g_{\mathbb{D}}$ in polar coordinates takes the form

$$g_{\mathbb{D}} = dr^2 + (b(r, \theta))^2 d\theta^2,$$

with $\lim_{r \rightarrow 0} b(r, \theta) = 0$ and $\lim_{r \rightarrow 0} \frac{b(r, \theta)}{r} = 1$. Our aim is to express r, θ as functions of the background coordinates $\bar{r}, \bar{\theta}$ on $\mathbb{D}^2 \subset \mathbb{R}^2$ and compute $b(r, \theta)$. To this end, we want to make use of the fact that $(\mathbb{D}^2, g_{\mathbb{D}})$ expressed in the $(\bar{r}, \bar{\theta})$ coordinate system is rotationally symmetric to infer that $\theta = \bar{\theta}$ and that r and b are functions only of \bar{r} (and not of θ). Even though this statement should be intuitively clear, let us try to set up this argument in detail.

It is easy to verify that the geodesics of $(\mathbb{D}^2, g_{\mathbb{D}})$ emanating from the origin are straight line segments in \mathbb{D}^2 . Therefore, the curves $\{\theta = \text{const}\}$ are the same as the curves $\{\bar{\theta} = \text{const}\}$, i.e. $\theta = \theta(\bar{r}, \bar{\theta})$ is a function *only* of $\bar{\theta}$. We will now show that this implies that $\theta = \bar{\theta}$: The condition (1) implies that the Jacobian matrix of the transformation matrix $(x, y) \rightarrow (\bar{x}, \bar{y})$ satisfies

$$\begin{bmatrix} \partial_x x^1 & \partial_x x^2 \\ \partial_y x^1 & \partial_y x^2 \end{bmatrix} \xrightarrow{(x,y) \rightarrow (0,0)} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Thus, using the fact that $\theta = \text{Arctan}(\frac{x^2}{x^1})$ and $\bar{\theta} = \text{Arctan}(\frac{y}{x})$, we infer that

$$\lim_{\bar{r} \rightarrow 0} \frac{\theta}{\bar{\theta}} = 1.$$

The fact that $\theta = \theta(\bar{\theta})$ then implies that

$$\theta = \bar{\theta}.$$

We will now seek an expression for $r = r(\bar{r}, \bar{\theta})$. Recall that the point q in $(\mathbb{D}^2, g_{\mathbb{D}}$ corresponding to the polar coordinate pair (r, θ) is simply

$$q = \exp_p (r \cos \theta e_1 + r \sin \theta e_2).$$

In particular, for any $\rho > 0$, the set $\{r = \rho\}$ in \mathbb{D} is the images under \exp_p of the set $S_\rho^{(p)} = \{v = (v^1, v^2) \in T_p \mathbb{D}^2 : (v^1)^2 + (v^2)^2 = \rho^2\}$ (where (v^1, v^2) are the coordinates of v in the orthonormal basis $\{e_1, e_2\} = \{\partial_x|_p, \partial_y|_p\}$). The following observation is crucial: In the $(\bar{r}, \bar{\theta})$ coordinate system, the metric $g_{\mathbb{D}}$ takes the form

$$g_{\mathbb{D}} = \frac{4}{(1 - \bar{r}^2)^2} (d\bar{r}^2 + \bar{r}^2 d\bar{\theta}), \quad (2)$$

i.e. the coefficients of the metric are *independent* of $\bar{\theta}$, hence the rotations $\Phi_\lambda : (\bar{r}, \bar{\theta}) \rightarrow (\bar{r}, \bar{\theta} + \lambda)$ are *isometries* for $g_{\mathbb{D}}$. Using the fact that isometries map geodesics to geodesics (see Ex. 6.1), and $\Phi_*|_p$ maps S_ρ to S_ρ , we infer that, for any $\rho > 0$, the set $\{r = \rho\}$ is invariant under the rotations Φ_λ , $\lambda \in \mathbb{R}$. Since these rotations also leave the circles $\{\bar{r} = \text{const}\}$ invariant, we infer that the curves $\{r = \text{const}\}$ and $\{\bar{r} = \text{const}\}$ are the same, i.e. r is a function only of \bar{r} . Therefore, since $r = r(\bar{r})$ and $\theta = \bar{\theta}$, in the (r, θ) coordinate system the isometries Φ_λ also take the form $(r, \theta) \rightarrow (r, \theta + \lambda)$; we deduce that, in the polar (r, θ) coordinate system, the coefficients of $g_{\mathbb{D}}$ should be independent of θ ,

i.e. that b is a function only of r . Thus, we have the following expressions for $g_{\mathbb{D}}$ in the coordinate systems (r, θ) and $(\bar{r}, \bar{\theta}) = (\bar{r}, \theta)$:

$$g_{\mathbb{D}} = dr^2 + (b(r))^2 d\theta^2 = \left(\frac{dr}{d\bar{r}}\right)^2 d\bar{r}^2 + (b(r))^2 d\theta^2$$

and, in view of (2):

$$g_{\mathbb{D}} = \frac{4}{(1 - \bar{r}^2)^2} (d\bar{r}^2 + \bar{r}^2 d\theta^2).$$

We therefore infer that

$$\frac{dr}{d\bar{r}} = \frac{2}{1 - \bar{r}^2} \quad \text{and} \quad b(r(\bar{r})) = \frac{2\bar{r}}{1 - \bar{r}^2}$$

from which we obtain

$$r(\bar{r}) = \log\left(\frac{1 + \bar{r}}{1 - \bar{r}}\right) \quad \text{and} \quad b(r) = \sinh(r).$$

Thus, in polar coordinates (r, θ) around $p = (0, 0)$, $g_{\mathbb{D}}$ takes the form:

$$g_{\mathbb{D}} = dr^2 + (\sinh r)^2 d\theta^2.$$

Notice that $(r, \theta) \in (0, +\infty) \times [0, 2\pi)$ covers all of $\mathbb{D}^2 \setminus 0$.

(c) As in the case of the hyperbolic plane, the round sphere $(\mathbb{S}^2, g_{\mathbb{S}^2})$ is homogenous and, therefore, the metric expressed in polar coordinates around a point $p \in \mathbb{S}^2$ will have the same form independently of the choice of p ; we can therefore choose p to be the north pole N . Recall that, in stereographic coordinates from N (which parametrize $\mathbb{S}^2 \setminus S$ by points on the plane \mathbb{R}^2 , see Ex. 2.3), the round metric $g_{\mathbb{S}^2}$ takes the form

$$g_{\mathbb{S}^2} = \frac{4}{(1 + x^2 + y^2)^2} (dx^2 + dy^2)$$

(with $(x, y) = (0, 0)$ corresponding to p and $x^2 + y^2 \rightarrow +\infty$ corresponding to N). In particular, switching to radial coordinates $(\bar{r}, \bar{\theta})$ on \mathbb{R}^2 , we have

$$g_{\mathbb{S}^2} = \frac{4}{(1 + \bar{r}^2)^2} (d\bar{r}^2 + \bar{r}^2 d\bar{\theta}^2). \quad (3)$$

We immediately notice that geodesics emanating from p correspond, in the above coordinate system, to straight lines $\bar{\theta} = \text{const}$ and that the metric $g_{\mathbb{S}^2}$ is invariant under rotations $(\bar{r}, \bar{\theta}) \rightarrow (\bar{r}, \bar{\theta} + \lambda)$. Therefore, arguing exactly as in the case of the hyperbolic plane, we infer that the polar coordinate system (r, θ) around p satisfies $\theta = \bar{\theta}$ and $r = r(\bar{r})$ and that $b(r, \theta)$ is a function of r only, i.e.

$$g_{\mathbb{S}^2} = dr^2 + (b(r))^2 d\theta^2.$$

Comparing the above expression with (3), we deduce that

$$\frac{dr}{d\bar{r}} = \frac{2}{1 + \bar{r}^2} \quad \text{and} \quad b(r(\bar{r})) = \frac{2\bar{r}}{1 + \bar{r}^2},$$

i.e. that

$$r(\bar{r}) = 2 \arctan \bar{r} \quad \text{and} \quad b(r) = \sin(r).$$

Thus,

$$g_{\mathbb{S}^2} = dr^2 + \sin^2 r d\theta^2$$

and $(r, \theta) \in (0, \pi) \times [0, 2\pi)$ covers $\mathbb{S}^2 \setminus \{N, S\}$

Remark. Notice the analogy with the corresponding expression for the hyperbolic metric.

(d) Using the fact that, on any Riemannian manifold (\mathcal{M}, g) , the Riemannian volume of a domain $\mathcal{U} \subset \mathcal{M}$ is given by

$$\text{Vol}(\mathcal{U}) = \int_{\mathcal{U}} d\text{Vol}_g$$

with $d\text{Vol}_g = \sqrt{\det(g)} dx^1 \dots dx^n$ (in local coordinates), we can readily compute using the formulas for the metrics in polar coordinates computed earlier that, for any $\rho > 0$:

$$\begin{aligned} \text{Vol}(B_{(\mathbb{R}^2, g_E)}[\rho]) &= \int_{B_{(\mathbb{R}^2, g_E)}[\rho]} d\text{Vol}_{g_E} = \int_0^\rho \int_0^{2\pi} \sqrt{\det(g_E)} d\theta dr = \int_0^\rho \int_0^{2\pi} r d\theta dr = \pi \rho^2, \\ \text{Vol}(B_{(\mathbb{H}^2, g_{\mathbb{H}^2})}[\rho]) &= \int_{B_{(\mathbb{H}^2, g_{\mathbb{H}^2})}[\rho]} d\text{Vol}_{\mathbb{H}^2} = \int_0^\rho \int_0^{2\pi} \sqrt{\det(g_{\mathbb{H}^2})} d\theta dr = \int_0^\rho \int_0^{2\pi} \sinh(r) d\theta dr = 2\pi(\cosh(\rho) - 1). \end{aligned}$$

Similarly, for any $\rho \in (0, \pi)$ (note that π is the radius of injectivity of $(\mathbb{S}^2, g_{\mathbb{S}^2})$), we have

$$\text{Vol}(B_{(\mathbb{S}^2, g_{\mathbb{S}^2})}[\rho]) = \int_{B_{(\mathbb{S}^2, g_{\mathbb{S}^2})}[\rho]} d\text{Vol}_{\mathbb{S}^2} = \int_0^\rho \int_0^{2\pi} \sqrt{\det(g_{\mathbb{S}^2})} d\theta dr = \int_0^\rho \int_0^{2\pi} \sin(r) d\theta dr = 2\pi(1 - \cos(\rho))$$

while, for $\rho \geq \pi$, we have

$$B_{(\mathbb{S}^2, g_{\mathbb{S}^2})}[\rho] = B_{(\mathbb{S}^2, g_{\mathbb{S}^2})}[\pi] = \mathbb{S}^2.$$

Since, for $0 < r \leq \pi$, we have

$$1 - \cos(r) < \frac{r^2}{2} < \cosh(r) - 1,$$

we infer that, for any $\rho > 0$:

$$\text{Vol}(B_{(\mathbb{S}^2, g_{\mathbb{S}^2})}[\rho]) < \text{Vol}(B_{(\mathbb{R}^2, g_E)}[\rho]) < \text{Vol}(B_{(\mathbb{H}^2, g_{\mathbb{H}^2})}[\rho]).$$

7.3 The Euler–Lagrange equations. Let $\Omega \subset \mathbb{R}^n$ be an open domain and let $(t, x, v) \rightarrow \mathcal{L}(t, x, v) \in \mathbb{R}$ be a smooth function for $(t, x, v) \in [a, b] \times \Omega \times \mathbb{R}^n$. For any smooth map $f : [a, b] \rightarrow \Omega$, we will define its action with respect to \mathcal{L} by the relation

$$S_{\mathcal{L}}[f] \doteq \int_a^b \mathcal{L}(t, f(t), \frac{df}{dt}(t)) dt.$$

Let $F : (-\delta, \delta) \times [a, b] \rightarrow \Omega$ be a smooth variation of f , i.e. a smooth function satisfying

$$F(0, \cdot) = f(\cdot).$$

Show that

$$\begin{aligned} \frac{d}{ds} S_{\mathcal{L}}[F(s, \cdot)] \Big|_{s=0} &= \left[\frac{\partial F^i}{\partial s}(0, t) \cdot \partial_{v^i} \mathcal{L}(t, f(t), f'(t)) \right]_{t=a}^b \\ &\quad + \int_a^b \frac{\partial F^i}{\partial s}(0, t) \left(\partial_{x^i} \mathcal{L}(t, f(t), f'(t)) - \frac{d}{dt} (\partial_{v^i} \mathcal{L}(t, f(t), f'(t))) \right) dt, \end{aligned} \quad (4)$$

where $\partial_{x^i} \mathcal{L}$ and $\partial_{v^i} \mathcal{L}$ denote the corresponding partial derivative of the function

$$\mathcal{L} = \mathcal{L}(t; x^1, \dots, x^n; v^1, \dots, v^n)$$

with respect to the Cartesian coordinates x^i and v^i on Ω and \mathbb{R}^n , respectively (*Hint: After applying the $\frac{\partial}{\partial s}$ derivative inside the integral defining $S_{\mathcal{L}}$, perform an integration by parts on the term $\partial_s \partial_t F(s, t)$.*)

Deduce that if $f : [a, b] \rightarrow \Omega$ is a stationary point of $S_{\mathcal{L}}$ under all variations that fix the endpoints $t = a, b$ (i.e. $F(s, a) = f(a)$ and $F(s, b) = f(b)$), then f satisfies the *Euler–Lagrange equations*:

$$\partial_{x^i} \mathcal{L}(t, f(t), f'(t)) - \frac{d}{dt} (\partial_{v^i} \mathcal{L}(t, f(t), f'(t))) = 0.$$

Remark. In classical mechanics, $f : [a, b] \rightarrow \Omega$ can be thought of as the trajectory of a particle moving in the domain Ω for time $t \in [a, b]$. In this case, we can define \mathcal{L} to be the *Lagrangian* of the particle; in the case when the particle moves under the influence of a conservative force (i.e. one which can be written as minus the gradient of a potential), the Lagrangian takes the form of the difference between the kinetic and potential energy of the particle:

$$\mathcal{L}(t, x, v) = \frac{1}{2} mv^2 - U(x).$$

The functional $S_{\mathcal{L}}$ is called the *action* of the trajectory f . An equivalent way of formulating Newtonian mechanics is by assuming the *principle of least action*: The particle moves along a trajectory for which the action is stationary among all paths between $f(a)$ and $f(b)$. You can verify that, in the case of a conservative force, the Euler–Lagrange equations are the standard Newtonian equations of motion for the particle:

$$m \frac{d^2 f^i}{dt^2}(t) = -\partial_i U \circ f(t).$$

Solution. We can directly calculate using the chain rule formula for differentiation:

$$\frac{d}{ds} S_{\mathcal{L}}[F(s, \cdot)] = \frac{d}{ds} \int_a^b \mathcal{L}(t, F(s, t), \partial_t F(s, t)) dt$$

$$\begin{aligned}
&= \int_a^b \frac{\partial}{\partial s} (\mathcal{L}(t, F(s, t), \partial_t F(s, t))) dt \\
&= \int_a^b \left(\frac{\partial \mathcal{L}(t, x, v)}{\partial x^i} \Big|_{(x,v)=(F(s,t),\partial_t F(s,t))} \cdot \frac{\partial F^i}{\partial s}(s, t) \right. \\
&\quad \left. + \frac{\partial \mathcal{L}(t, x, v)}{\partial v^i} \Big|_{(x,v)=(F(s,t),\partial_t F(s,t))} \cdot \frac{\partial^2 F^i}{\partial s \partial t}(s, t) \right) dt \\
&= \int_a^b \left(\frac{\partial \mathcal{L}(t, x, v)}{\partial x^i} \Big|_{(x,v)=(F(s,t),\partial_t F(s,t))} \cdot \frac{\partial F^i}{\partial s}(s, t) \right. \\
&\quad \left. + \frac{\partial \mathcal{L}(t, x, v)}{\partial v^i} \Big|_{(x,v)=(F(s,t),\partial_t F(s,t))} \cdot \frac{\partial}{\partial t} \left(\frac{\partial F^i}{\partial s}(s, t) \right) \right) dt \\
&= \int_a^b \left[\frac{\partial \mathcal{L}(t, x, v)}{\partial x^i} \Big|_{(x,v)=(F(s,t),\partial_t F(s,t))} \cdot \frac{\partial F^i}{\partial s}(s, t) \right. \\
&\quad \left. + \frac{\partial}{\partial t} \left\{ \frac{\partial \mathcal{L}(t, x, v)}{\partial v^i} \Big|_{(x,v)=(F(s,t),\partial_t F(s,t))} \cdot \frac{\partial F^i}{\partial s}(s, t) \right\} \right. \\
&\quad \left. - \frac{\partial}{\partial t} \left(\frac{\partial \mathcal{L}(t, x, v)}{\partial v^i} \Big|_{(x,v)=(F(s,t),\partial_t F(s,t))} \right) \cdot \frac{\partial F^i}{\partial s}(s, t) \right] dt \\
&= \int_a^b \left[\frac{\partial \mathcal{L}(t, x, v)}{\partial x^i} \Big|_{(x,v)=(F(s,t),\partial_t F(s,t))} \cdot \frac{\partial F^i}{\partial s}(s, t) \right. \\
&\quad \left. - \frac{\partial}{\partial t} \left(\frac{\partial \mathcal{L}(t, x, v)}{\partial v^i} \Big|_{(x,v)=(F(s,t),\partial_t F(s,t))} \right) \cdot \frac{\partial F^i}{\partial s}(s, t) \right] dt \\
&\quad + \left[\frac{\partial \mathcal{L}(t, x, v)}{\partial v^i} \Big|_{(x,v)=(F(s,t),\partial_t F(s,t))} \cdot \frac{\partial F^i}{\partial s}(s, t) \right]_{t=a}^b.
\end{aligned}$$

Setting $s = 0$ above and using our assumption that $F(0, t) = f(t)$, we obtain the desired relation:

$$\begin{aligned}
\frac{d}{ds} S_{\mathcal{L}}[F(s, \cdot)] \Big|_{s=0} &= \int_a^b \left[\partial_{x^i} \mathcal{L}(t, f(t), f'(t)) - \frac{d}{dt} (\partial_{v^i} \mathcal{L}(t, f(t), f'(t))) \right] \cdot \frac{\partial F^i}{\partial s}(s, t) \Big|_{s=0} dt \quad (5) \\
&\quad + \left[\partial_{v^i} \mathcal{L}(t, f(t), f'(t)) \cdot \frac{\partial F^i}{\partial s}(s, t) \Big|_{s=0} \right]_{t=a}^b.
\end{aligned}$$

Suppose, now, that $f : [a, b] \rightarrow \Omega$ satisfies

$$\frac{d}{ds} S_{\mathcal{L}}[F(s, \cdot)] \Big|_{s=0} = 0 \quad (6)$$

for all variations $F : (-\delta, \delta) \times [a, b] \rightarrow \Omega$ fixing the endpoints, i.e. satisfying $F(s, a) = f(a)$ and $F(s, b) = f(b)$ for all $s \in (-\delta, \delta)$. Let $\chi : [a, b] \rightarrow [0, +\infty)$ be a smooth function such that $\chi(a) =$

$\chi(b) = 0$ and $\chi(t) > 0$ for $t \in (a, b)$. Let us choose a variation $F : (-\delta, \delta) \times [a, b] \rightarrow \Omega$ of f with the following properties:

1. $F(s, a) = f(a)$ and $F(s, b) = f(b)$ for all $s \in (-\delta, \delta)$,
2. $\partial_s F(s, t) \Big|_{s=0} = \chi(t) \left(\partial_{x^i} \mathcal{L}(t, f(t), f'(t)) - \frac{d}{dt} (\partial_{v^i} \mathcal{L}(t, f(t), f'(t))) \right)$.

Note that such a variation exists: Since f is smooth and $f([a, b])$ is a compact subset of the open set Ω , there exists a $\rho > 0$ such that

$$\text{dist}(f(t), \partial\Omega) > \rho \quad \text{for all } t \in [a, b].$$

Therefore, choosing

$$0 < \delta < \frac{\rho}{\max_{t \in [a, b]} \left[\chi(t) \left(\partial_{x^i} \mathcal{L}(t, f(t), f'(t)) - \frac{d}{dt} (\partial_{v^i} \mathcal{L}(t, f(t), f'(t))) \right) \right]},$$

the map $F : (-\delta, \delta) \times [a, b] \rightarrow \mathbb{R}^n$ defined by

$$F(s, t) = f(t) + s\chi(t) \left(\partial_{x^i} \mathcal{L}(t, f(t), f'(t)) - \frac{d}{dt} (\partial_{v^i} \mathcal{L}(t, f(t), f'(t))) \right)$$

stays inside Ω and satisfies properties 1 and 2.

Using F as above in (6) and applying the formula (5) (noting that

$$\left[\partial_{v^i} \mathcal{L}(t, f(t), f'(t)) \cdot \frac{\partial F^i}{\partial s}(s, t) \Big|_{s=0} \right]_{t=a}^b = 0$$

since $F(s, a) = f(a)$ and $F(s, b) = f(b)$), we obtain

$$\int_a^b \chi(t) \left[\partial_{x^i} \mathcal{L}(t, f(t), f'(t)) - \frac{d}{dt} (\partial_{v^i} \mathcal{L}(t, f(t), f'(t))) \right]^2 dt = 0.$$

Since $\chi|_{(a,b)} > 0$ and $\partial_{x^i} \mathcal{L}(t, f(t), f'(t)) - \frac{d}{dt} (\partial_{v^i} \mathcal{L}(t, f(t), f'(t)))$ is continuous on $[a, b]$, we deduce that

$$\partial_{x^i} \mathcal{L}(t, f(t), f'(t)) - \frac{d}{dt} (\partial_{v^i} \mathcal{L}(t, f(t), f'(t))) = 0 \quad \text{for all } t \in [a, b].$$

***7.4 Geodesics as stationary points of the energy functional.** We will now extend the formalism of the previous exercise to the realm of manifolds. Let (\mathcal{M}, g) be a smooth Riemannian manifold. Let $\mathcal{L} : T\mathcal{M} \rightarrow \mathbb{R}$ be a smooth function; for any $x \in \mathcal{M}$ and $v \in T_x \mathcal{M}$, we will denote with $\mathcal{L}(x, v)$ the value of \mathcal{L} at $(x, v) \in T\mathcal{M}$. For any smooth curve $\gamma : [a, b] \rightarrow \mathcal{M}$, we will define the action of γ with respect to \mathcal{L} by

$$S_{\mathcal{L}}[\gamma] \doteq \int_a^b \mathcal{L}[\gamma(t), \dot{\gamma}(t)] dt.$$

(a) Let $\gamma : [a, b] \rightarrow \mathcal{M}$ be a given curve and $\phi : (-\delta, \delta) \times [a, b] \rightarrow \mathcal{M}$ be a smooth variation of γ which is entirely contained in a coordinate chart (x^1, \dots, x^n) on \mathcal{M} ; we will denote with $\frac{\partial \phi_s}{\partial s}|_{s=0}$ the variation field along γ (as we did in class). Show that (4) also holds in this case, i.e.

$$\begin{aligned} \frac{d}{ds} S_{\mathcal{L}}[\phi_s]|_{s=0} &= \left[\partial_{v^i} \mathcal{L}(\gamma(t), \dot{\gamma}(t)) \cdot \frac{\partial \phi_s^i}{\partial s}|_{s=0}(t) \right]_{t=a}^b \\ &\quad + \int_a^b \left(\partial_{x^i} \mathcal{L}(\gamma(t), \dot{\gamma}(t)) - \frac{d}{dt} (\partial_{v^i} \mathcal{L}(\gamma(t), \dot{\gamma}(t))) \right) \cdot \frac{\partial \phi_s^i}{\partial s}|_{s=0}(t) dt, \end{aligned}$$

where, in the local coordinates $(x^1, \dots, x^n; v^1, \dots, v^n)$ on $T\mathcal{M}$ associated to (x^1, \dots, x^n) (recall that $v^i(V) = dx^i(V)$ for any $V \in \Gamma(\mathcal{M})$), $\partial_{x^i} \mathcal{L}$ and $\partial_{v^i} \mathcal{L}$ denote the partial derivatives of $\mathcal{L}(x^1, \dots, x^n; v^1, \dots, v^n)$ with respect to the corresponding variables.

Moreover, if γ is a stationary point for $S_{\mathcal{L}}$ for all variations ϕ_s with $\phi_s(a) = \gamma(a)$ and $\phi_s(b) = \gamma(b)$, then

$$\partial_{x^i} \mathcal{L}(\gamma(t), \dot{\gamma}(t)) - \frac{d}{dt} (\partial_{v^i} \mathcal{L}(\gamma(t), \dot{\gamma}(t))) = 0, \quad i = 1, \dots, n.$$

(b) Let us now examine the case when

$$\mathcal{L}(x, v) = \frac{1}{2} g|_x(v, v) \quad \text{for } x \in \mathcal{M}, v \in T_x \mathcal{M}$$

(this can be thought of as an extension of the Newtonian function for the kinetic energy in the setting of Riemannian manifolds). In this case, the action $S_{\mathcal{L}}$ is known as the *energy* functional (which we also saw in Ex. 2.1):

$$\mathcal{E}[\gamma] = \int_a^b g(\dot{\gamma}(t), \dot{\gamma}(t)) dt.$$

Show that if ϕ_s is a variation of a smooth curve $\gamma : [a, b] \rightarrow \mathcal{M}$, not necessarily contained in a single coordinate chart, then

$$\frac{d}{ds} \mathcal{E}[\phi_s]|_{s=0} = \left\langle \frac{\partial \phi_s}{\partial s}|_{s=0}, \dot{\gamma} \right\rangle_g|_{t=a}^b - \int_a^b \left\langle \frac{\partial \phi_s}{\partial s}|_{s=0}, \nabla_{\dot{\gamma}} \dot{\gamma} \right\rangle_g dt.$$

This is known as the 1st variation formula for the energy. (*Hint: Break up the variation into smaller intervals in t such that each one is contained inside a single coordinate chart.*) Deduce that if γ is a stationary point for the energy under all variations which fix the endpoints $\gamma(a), \gamma(b)$, then γ is a geodesic.

Remark. In contrast to the case of stationary curves for the length functional (which are reparametrizations of geodesics, not necessarily with constant speed), a reparametrization of a geodesic is not a stationary point for $\mathcal{E}[\gamma]$. Thus, $\mathcal{E}[\gamma]$ can be used to single out “properly parametrized” geodesics via a minimization process.

(c) For any $p \in \mathcal{M}$, let $\sigma : [a, b] \rightarrow \Omega_p \subset T_p \mathcal{M}$ be a smooth curve (Ω_p is the domain of definition of the exponential map \exp_p). Show that

$$\frac{d}{ds} (\|\sigma(s)\|_{g|_p}^2) = \langle d\exp_p|_{\sigma(s)} \dot{\sigma}(s), d\exp_p|_{\sigma(s)} \sigma(s) \rangle_g,$$

where we view $\sigma(s)$ both as point in Ω_p and a vector in $T_{\sigma(s)} \Omega_p$ (namely as the tangent vector of the line $t \rightarrow \sigma(s)t$ at $t = 1$). *Hint: What is the energy of the geodesic $t \rightarrow \exp_p(\sigma(s)t)$, $t \in [0, 1]$?* Deduce from the above formula the statement of the lemma of Gauss.

Solution. (a) The proof of this part is essentially Exercise 7.3.

(b) Let $a = t_0 < t_1 < \dots < t_{k-1} < t_k = b$ be a partition of $[a, b]$ such that, for any $l = 0, \dots, k-1$, the curve $\{\gamma(t) : t \in [t_l, t_{l+1}]\}$, lies inside the domain \mathcal{U}_l of a coordinate chart (x^1, \dots, x^n) on \mathcal{M} . The continuity of $\phi_s(t)$ in (s, t) then implies that, by possibly choosing a smaller value of $\delta > 0$ if necessary, we have $\{\phi_s(t) : t \in [t_l, t_{l+1}]\} \subset \mathcal{U}_l$.

Expressed in the associated coordinate chart $(x^1, \dots, x^n; v^1, \dots, v^n)$ on $T\mathcal{U}_l \simeq \mathcal{U}_l \times \mathbb{R}^n$ (where, for any $p \in \mathcal{U}_l$ and $\xi \in T_p \mathcal{M}$, we have $v^i(\xi) = dx^i|_p(\xi)$), the function \mathcal{L} takes the form

$$\mathcal{L}(x, v) = \frac{1}{2} g_{ab}(x) v^a v^b.$$

Note that $\partial_{v^i} \mathcal{L}(x, v) = g_{ib} v^b$ and, therefore,

$$\begin{aligned} \partial_{x^i} \mathcal{L}(\gamma(t), \dot{\gamma}(t)) - \frac{d}{dt} (\partial_{v^i} \mathcal{L}(\gamma(t), \dot{\gamma}(t))) \\ &= \frac{1}{2} \partial_i g_{ab}|_{\gamma(t)} \dot{\gamma}^a(t) \dot{\gamma}^b(t) - \frac{d}{dt} (g_{ib}|_{\gamma(t)} \dot{\gamma}^b(t)) \\ &= \frac{1}{2} \partial_i g_{ab}|_{\gamma(t)} \dot{\gamma}^a(t) \dot{\gamma}^b(t) - \left(\frac{d}{dt} (g_{ib}|_{\gamma(t)} \dot{\gamma}^b(t)) + g_{ib}|_{\gamma(t)} \ddot{\gamma}^b(t) \right) \\ &= \frac{1}{2} \partial_i g_{ab}|_{\gamma(t)} \dot{\gamma}^a(t) \dot{\gamma}^b(t) - (\partial_j g_{ib}|_{\gamma(t)} \dot{\gamma}^j(t) \dot{\gamma}^b(t) + g_{ib}|_{\gamma(t)} \ddot{\gamma}^b(t)) \\ &= - \left(g_{ib}|_{\gamma(t)} \ddot{\gamma}^b(t) + \frac{1}{2} \left(2\partial_j g_{ib}|_{\gamma(t)} \dot{\gamma}^j(t) \dot{\gamma}^b(t) - \partial_i g_{ab}|_{\gamma(t)} \dot{\gamma}^a(t) \dot{\gamma}^b(t) \right) \right) \\ &= - \left(g_{ib}|_{\gamma(t)} \ddot{\gamma}^b(t) + \frac{1}{2} \left(\partial_j g_{ib}|_{\gamma(t)} \dot{\gamma}^j(t) \dot{\gamma}^b(t) + \partial_b g_{ij}|_{\gamma(t)} \dot{\gamma}^j(t) \dot{\gamma}^b(t) - \partial_i g_{ab}|_{\gamma(t)} \dot{\gamma}^a(t) \dot{\gamma}^b(t) \right) \right). \end{aligned}$$

Thus, using the fact that

$$g_{ik} \Gamma_{jb}^k = \frac{1}{2} \left(\partial_j g_{ib} + \partial_b g_{ij} - \partial_i g_{jb} \right),$$

the above calculation yields:

$$\partial_{x^i} \mathcal{L}(\gamma(t), \dot{\gamma}(t)) - \frac{d}{dt} (\partial_{v^i} \mathcal{L}(\gamma(t), \dot{\gamma}(t))) \tag{7}$$

$$\begin{aligned}
&= - \left(g_{ik}|_{\gamma(t)} \ddot{\gamma}^k(t) + g_{ik}|_{\gamma(t)} \Gamma_{jb}^k|_{\gamma(t)} \dot{\gamma}^j(t) \dot{\gamma}^b(t) \right) \\
&= -g_{ik}|_{\gamma(t)} (\nabla_{\dot{\gamma}(t)} \dot{\gamma}(t))^k.
\end{aligned}$$

Denoting

$$\mathcal{E}^{(l)}[\phi_s] = \int_{t_l}^{t_{l+1}} \mathcal{L}(\phi_s(t), \dot{\phi}_s(t)) dt,$$

we obtain after applying part (a) of this exercise and using (7):

$$\begin{aligned}
\frac{d}{ds} \mathcal{E}^{(l)}[\phi_s]|_{s=0} &= \left[g_{ij}|_{\gamma(t)} \frac{\partial \phi_s^i}{\partial s}(0, t) \dot{\gamma}^j(t) \right] \Big|_{t=t_l}^{t_{l+1}} - \int_{t_l}^{t_{l+1}} g_{ij}|_{\gamma(t)} \frac{\partial \phi_s^i}{\partial s}(0, t) (\nabla_{\dot{\gamma}(t)} \dot{\gamma}(t))^j dt \\
&= \left\langle \frac{\partial \phi_s}{\partial s}|_{s=0}, \dot{\gamma} \right\rangle_g \Big|_{t=t_l}^{t_{l+1}} - \int_{t_l}^{t_{l+1}} \left\langle \frac{\partial \phi_s}{\partial s}|_{s=0}, \nabla_{\dot{\gamma}} \dot{\gamma} \right\rangle_g dt.
\end{aligned} \tag{8}$$

Summing over $l = 0, \dots, k-1$ and using the fact that $\mathcal{E}[\phi_s] = \sum_{l=0}^{k-1} \mathcal{E}^{(l)}[\phi_s]$ (noting also the cancellations occurring in the sum for the first terms in the right hand side of (8)), we obtain the desired identity:

$$\begin{aligned}
\frac{d}{ds} \mathcal{E}[\phi_s]|_{s=0} &= \sum_{l=0}^{k-1} \left\langle \frac{\partial \phi_s}{\partial s}|_{s=0}, \dot{\gamma} \right\rangle_g \Big|_{t=t_l}^{t_{l+1}} - \sum_{l=0}^{k-1} \int_{t_l}^{t_{l+1}} \left\langle \frac{\partial \phi_s}{\partial s}|_{s=0}, \nabla_{\dot{\gamma}} \dot{\gamma} \right\rangle_g dt \\
&= \left\langle \frac{\partial \phi_s}{\partial s}|_{s=0}, \dot{\gamma} \right\rangle_g \Big|_{t=t_0}^{t_k} - \int_{t_0}^{t_k} \left\langle \frac{\partial \phi_s}{\partial s}|_{s=0}, \nabla_{\dot{\gamma}} \dot{\gamma} \right\rangle_g dt \\
&= \left\langle \frac{\partial \phi_s}{\partial s}|_{s=0}, \dot{\gamma} \right\rangle_g \Big|_{t=a}^b - \int_a^b \left\langle \frac{\partial \phi_s}{\partial s}|_{s=0}, \nabla_{\dot{\gamma}} \dot{\gamma} \right\rangle_g dt.
\end{aligned} \tag{9}$$

In the case when γ is a stationary point of $\mathcal{E}[\phi_s]$ for all variations ϕ_s with $\phi_s(a) = \gamma(a)$ and $\phi_s(b) = \gamma(b)$, by choosing a variation ϕ_s such that

$$\frac{\partial \phi_s}{\partial s}(0, t) = \chi(t) \nabla_{\dot{\gamma}} \dot{\gamma}|_{\gamma(t)}$$

(where $\chi : [a, b] \rightarrow [0, +\infty)$ is chosen as in Exercise 7.1, so that $\chi(a) = \chi(b) = 0$ and $\chi(t) > 0$ for $t \in (a, b)$) we obtain

$$0 = \frac{d}{ds} \mathcal{E}[\phi_s]|_{s=0} = - \int_a^b \chi(t) \langle \nabla_{\dot{\gamma}} \dot{\gamma}, \nabla_{\dot{\gamma}} \dot{\gamma} \rangle_g dt$$

from which we infer (since $\langle \nabla_{\dot{\gamma}} \dot{\gamma} \rangle_g \geq 0$ and $\nabla_{\dot{\gamma}} \dot{\gamma}|_{\gamma(t)}$ is continuous in t) that $\nabla_{\dot{\gamma}} \dot{\gamma} \equiv 0$. The construction of a variation ϕ_s as above can be done similarly as in Exercise 7.1 (by constructing $\{\phi_s(t) : t \in [t_l, t_1]\}$ for each successive $l = 0, \dots, k-1$ in the coordinate chart \mathcal{U}_l so that the points $\phi_s(t_l) \in \mathcal{M}$ coincide with the value obtained from the construction in the previous coordinate chart \mathcal{U}_{l-1}).

(c) Let us consider the family of curves $\phi_s : [0, 1] \rightarrow \mathcal{M}$, $s \in [a, b]$, defined by the relation

$$\phi_s(t) = \exp_p(\sigma(s)t).$$

Note that for any value of s , ϕ_s is a geodesic of (\mathcal{M}, g) ; moreover, we have

$$\phi_s(0) = p \quad \text{for all } s \in [a, b].$$

Thus, for any $s_0 \in (a, b)$, applying the variation formula (9) for the family of curves ϕ_s (considering them as variations of the geodesic ϕ_{s_0}), we obtain:

$$\frac{d}{ds} \mathcal{E}[\phi_s]|_{s=s_0} = \left\langle \frac{\partial \phi_s}{\partial s}|_{s=s_0}, \dot{\phi}_{s_0} \right\rangle_g|_{t=0} - \int_0^1 \left\langle \frac{\partial \phi_s}{\partial s}|_{s=s_0}, \nabla_{\dot{\phi}_{s_0}} \dot{\phi}_{s_0} \right\rangle_g dt.$$

Since $\frac{\partial \phi_s}{\partial s}|_{t=0} = 0$ and ϕ_{s_0} is a geodesic, we deduce

$$\begin{aligned} \frac{d}{ds} \mathcal{E}[\phi_s]|_{s=s_0} &= \left\langle \frac{\partial \phi_s}{\partial s}(s_0, 1), \dot{\phi}_{s_0}(1) \right\rangle_g \\ &= \left\langle \left(\frac{d}{ds} \exp_p(\sigma(s)) \right)|_{s=s_0}, \frac{d}{dt} \exp_p(\sigma(s_0)t)|_{t=1} \right\rangle_g. \end{aligned} \quad (10)$$

Since $\phi_s(t)$ is a geodesic, $\|\dot{\phi}_s(t)\|$ is constant in $t \in [0, 1]$. Therefore,

$$\mathcal{E}[\phi_s] = \int_0^1 g(\dot{\phi}_s(t), \dot{\phi}_s(t)) dt = \int_0^1 \|\dot{\phi}_s(t)\|^2 dt = \|\dot{\phi}_s(0)\|^2 = \left\| \frac{d}{dt} (\exp_p(\sigma(s)t))|_{t=0} \right\|^2.$$

Using the fact that $d \exp_p|_{v=0} = \text{Id}$, we therefore have $\frac{d}{dt} (\exp_p(\sigma(s)t))|_{t=0} = \sigma(s)$ and, thus,

$$\mathcal{E}[\phi_s] = \|\sigma(s)\|_{g_p}^2.$$

Moreover, we calculate (using the general formula $\frac{d}{ds}(F \circ \gamma(s)) = dF|_{\gamma(s)} \dot{\gamma}(s)$ for any curve $\gamma : I \rightarrow \mathcal{N}_1$ and any smooth map $F : \mathcal{N}_1 \rightarrow \mathcal{N}_2$):

$$\frac{d}{ds} \exp_p(\sigma(s))|_{s=s_0} = d \exp_p|_{\sigma(s)} \dot{\sigma}(s)$$

and

$$\frac{d}{dt} \exp_p(\sigma(s_0)t)|_{t=1} = d \exp_p|_{\sigma(s_0)} \sigma(s_0).$$

Substituting the above relations in (10), we finally obtain that, for any s_0 :

$$\frac{d}{ds} (\|\sigma(s)\|_{g_p}^2)|_{s=s_0} = \left\langle d \exp_p|_{\sigma(s_0)} \dot{\sigma}(s_0), d \exp_p|_{\sigma(s_0)} \sigma(s_0) \right\rangle_g.$$

Let $v, \xi \in T_p \mathcal{M}$ be tangent vectors so that $v \perp \xi$. Using the above formula for the curves $\sigma_1(s) = vs$ and $\sigma_2(s) = \cos(s)v + \sin(s)\xi$ for $s_0 = 0$, we obtain the statement of the lemma o Gauss, namely

$$\langle d \exp_p|_v v, d \exp_p|_v v \rangle_g = \langle v, v \rangle_g, \quad \langle d \exp_p|_v \xi, d \exp_p|_v v \rangle_g = 0.$$