Problem sheet 2 Solutions

Joe Keir

October 22, 2019

1. There are many ways to do this problem! Probably the most elegant is using stereographic projection, but another way based on polar coordinates will be tried by at least some students. Both solutions are given below.

Using stereographic projection

First we'll define a chart that covers all of the sphere except for a small circle around the north pole. Then we'll do the same for the south pole, and finally we'll show that the transition functions are smooth.

The first coordinate patch on the sphere is

$$U = \left\{ x^2 + y^2 + z^2 = 1 \, , \, \frac{x^2 + y^2}{(1-z)^2} < R^2 \right\}$$

where $R > 1$ is some constant (we can take R very large if we want).

The map to $\mathbb{R}^2$ is given by

$$\begin{aligned} \phi_U : U &\to \mathbb{R}^2 \\ (x, y, z) &\mapsto \left(\frac{x}{1-z}, \frac{y}{1-z} \right) \end{aligned}$$

The image of U is just the interior of the disc of radius R, i.e. $(X, Y) \in \mathbb{R}^2$ with $X^2 + Y^2 < R^2$. Call this region D_R.

We can check that this is a bijection by exhibiting its inverse:

$$\begin{aligned} \phi_U^{-1} : D_R &\to \mathbb{S}^2 \\ (X, Y) &\mapsto \left(\frac{2X}{1 + X^2 + Y^2} \, , \, \frac{2Y}{1 + X^2 + Y^2} \, , \, \frac{-1 + X^2 + Y^2}{1 + X^2 + Y^2} \right) \end{aligned}$$

We can do exactly the same thing projecting from the south pole. In this case we define

$$U' = \left\{ x^2 + y^2 + z^2 = 1 \, , \, \frac{x^2 + y^2}{(1+z)^2} < R^2 \right\}$$

and

$$\begin{aligned} \phi_{U'} : U' &\to \mathbb{R}^2 \\ (x, y, z) &\mapsto \left(\frac{x}{1+z}, \frac{y}{1+z} \right) \end{aligned}$$

and again we can compute the inverse

$$\begin{aligned} \phi_{U'}^{-1} : D_R &\to \mathbb{S}^2 \\ (X', Y') &\mapsto \left(\frac{2X'}{1 + (X')^2 + (Y')^2} \, , \, \frac{2Y'}{1 + (X')^2 + (Y')^2} \, , \, \frac{1 - (X')^2 - (Y')^2}{1 + (X')^2 + (Y')^2} \right) \end{aligned}$$

Finally, we can compute the transition function $\phi_{U,U'}$. We find that

$$\phi_{U,U'}(X,Y) = \left(\frac{X}{X^2+Y^2}, \frac{Y}{X^2+Y^2}\right)$$

moreover we find that

$$\phi_U\left(U \cap U'\right) = \left\{(X,Y) \in \mathbb{R}^2 \,:\, \frac{1}{R^2} < X^2+Y^2 < R^2\right\}$$

and $\phi_{U,U'}$ is evidently smooth on this set. Note that this function is its own inverse, so $\phi_{U',U} = \phi_{U,U'}$.

Using polar coordinates

An alternative approach is to use polar coordinates. Define the region

$$U = \mathbb{S}^2 \setminus \{y = 0\,,\, x \geq 0\}$$

Then we can define the map

$$\begin{aligned} \phi_U &: U \to \mathbb{R}^2 \\ (x,y,z) &\mapsto (\arccos z\,,\, \arg(x+iy)) \end{aligned}$$

Note that, when $x > 0$ and $y > 0$ then $\arg(x+iy) = \arctan\frac{y}{x}$, but when x and y lie in other quadrants then these two expressions differ.

Note that the image of U is the open set $(0,\pi)\times(0,2\pi)$. The inverse of ϕ_U can be computed:

$$\begin{aligned} \phi_U^{-1} &: (0,\pi)\times(0,2\pi) \to \mathbb{S}^2 \\ (\theta,\phi) &\mapsto (\sin\theta\cos\phi\,,\, \sin\theta\sin\phi\,,\, \cos\theta) \end{aligned}$$

Now we need to define a second region which covers the remaining part of the sphere. Define

$$U' = \mathbb{S}^2 \setminus \{x = 0\,,\, y \leq 0\}$$

and the map

$$\begin{aligned} \phi_{U'} &: U' \to \mathbb{R}^2 \\ (x,y,z) &\mapsto (\arccos x\,,\, \arg(z-iy)) \end{aligned}$$

As before, the image of U' is the open set $(0,\pi)\times(0,2\pi)$. The inverse of $\phi_{U'}$ is

$$\begin{aligned} \phi_{U'}^{-1} &: (0,\pi)\times(0,2\pi) \to \mathbb{S}^2 \\ (\theta',\phi') &\mapsto (\cos\theta'\,,\, -\sin\theta'\sin\phi'\,,\, \sin\theta'\cos\phi') \end{aligned}$$

Finally we can check the transition functions. First, we note that

$$\phi_U\left(U \cap U'\right) = \Big((0,\pi)\times(0,2\pi)\Big) \setminus \left(\left\{\frac{\pi}{2}\right\} \times \left[\frac{\pi}{2}, \frac{3\pi}{2}\right]\right)$$

so this is the domain of the transition function $\phi_{U,U'}$.

This transition function is found to be

$$\phi_{U,U'}(\theta,\phi) = \left(\arccos\left(\sin\theta\cos\phi\right)\,,\, \arg\left(\cos\theta - i\sin\theta\sin\phi\right)\right)$$

Now, arccos is smooth except when its argument is -1 and 1, which occurs when $\theta = \frac{\pi}{2}$ and $\phi = 0$ or π. But these values lie outside the domain of $\phi_{U,U'}$. Similarly, arg is smooth except when its argument lies on the non-negative real line (including the origin), i.e. when θ and ϕ are both multiples of π, and when $0 \le \theta \le \frac{\pi}{2}$. But no such points lie within the domain of $\phi_{U,U'}$.

Checking $\phi_{U',U}$ is very similar.

2.

a) The left hand side is a vector, the right hand a covector.

b) Adding vectors with different indices.

c) Adding tensors of different rank.

d) Left hand side is a vector (in abstract index notation), right hand side is a scalar.

e) The index ν is repeated 3 times.

f) The index μ is repeated 4 times – probably the second pair of μ's is supposed to be some other greek letter.

g) The index μ is repeated twice, but both indices are contravariant (i.e. vector space) indices. Contractions can only be between a contravariant and a covarant index, so probably one of the μ's is supposed to be lowered.

---

3. a) Let $C : [-1, 1] \to \mathcal{M}$ be some arbitrary curve through the point p, with $C(0) = p$. We need to show that the tangent vector to C at p can be expressed as a linear combination of the vectors ∂_a.

Writing the tangent vector to C at p as $\dot{C}$, for any smooth function f we have

$$\begin{aligned}\dot{C}(f) &= \frac{\mathrm{d}}{\mathrm{d}\lambda}(f \circ C)(\lambda)\big|_{\lambda=0} \\ &= \frac{\mathrm{d}}{\mathrm{d}\lambda}(f \circ \phi_U^{-1} \circ \phi_U \circ C)(\lambda)\big|_{\lambda=0} \\ &= \frac{\mathrm{d}}{\mathrm{d}\lambda}(\tilde{f}_U \circ \tilde{C}_U)(\lambda)\big|_{\lambda=0}\end{aligned}$$

where $\tilde{f}_U = f \circ \phi_U^{-1} : \mathbb{R}^n \to \mathbb{R}$, and $\tilde{C}_U = \phi_U \circ C : \mathbb{R} \to \mathbb{R}^n$. Hence, using the chain rule and ordinary multivariate calculus we can write

$$\dot{C}(f) = \mathrm{d}\tilde{f}_U\big|_{\phi_U(p)}\left(\frac{\mathrm{d}}{\mathrm{d}\lambda}\tilde{C}(\lambda)\big|_{\lambda=0}\right)$$

Note that $\frac{\mathrm{d}}{\mathrm{d}\lambda}\tilde{C}(\lambda)\big|_{\lambda=0}$ is just an ordinary vector in $\mathbb{R}^n$.

Next, we note that the vector $\frac{\mathrm{d}}{\mathrm{d}\lambda}\tilde{\gamma}_a(\lambda)$ is (at all values of λ) simply the a-th unit vector in $\mathbb{R}^n$. Hence, if we choose the parameter λ appropriately, we can expand

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\tilde{C}(\lambda)\big|_{\lambda=0} = \dot{C}^a \frac{\mathrm{d}}{\mathrm{d}\lambda}\tilde{\gamma}_a(\lambda)\big|_{\lambda=0}$$

where $\dot{C}^a$ are some real numbers (remember the summation convention!). Using the linearity of $\mathrm{d}\tilde{f}_U$ we find that

$$\begin{aligned}
\dot{C}(f) &= \mathrm{d}\tilde{f}_U\big|_{\phi_U(p)} \left(\dot{C}^a \frac{\mathrm{d}}{\mathrm{d}\lambda} \tilde{\gamma}_a(\lambda)\big|_{\lambda=0} \right) \\
&= \dot{C}^a \mathrm{d}\tilde{f}_U\big|_{\phi_U(p)} \left(\frac{\mathrm{d}}{\mathrm{d}\lambda} \tilde{\gamma}_a(\lambda)\big|_{\lambda=0} \right) \\
&= \dot{C}^a \frac{\mathrm{d}}{\mathrm{d}\lambda} \left(\tilde{f}_U \circ \tilde{\gamma}_a \right)(\lambda)\big|_{\lambda=0} \\
&= \dot{C}^a \frac{\mathrm{d}}{\mathrm{d}\lambda} \left(f \circ \gamma_a \right)(\lambda)\big|_{\lambda=0} \\
&= \dot{C}^a \partial_a(f)
\end{aligned}$$

so indeed the vector $\dot{C}$ can be expressed as a linear combination of the vectors ∂_a. Since $\dot{C}$ was arbitrary, the vectors $\{\partial_a\}$ span the tangent space at p.

b) The easiest way to show that the covectors $\{\mathrm{d}x^a\}$ form a basis for the cotangent space $T_p^*(\mathcal{M})$ is to show that they form the dual basis to the basis $\{\partial_a\}$ for the tangent space $T_p(\mathcal{M})$ (alternatively, we could expand an arbitrary covector in terms of this basis). We can compute

$$\begin{aligned}
\mathrm{d}x^b\left(\partial_a\right) &= \partial_a x^b \\
&= \frac{\mathrm{d}}{\mathrm{d}\lambda}\left(x^b \circ \gamma_a\right) \\
&= \frac{\mathrm{d}}{\mathrm{d}\lambda}\left(x^b \circ \tilde{\gamma}_a\right) \\
&= \delta^b_a
\end{aligned}$$

where in the third line we use a (common) abuse of notation to write x^a both for the standard coordinates on $\mathbb{R}^n$, and for ϕ_U^{-1} applied to these coordinate functions. To prove the last line, we note that, from the definitions of $\tilde{\gamma}_a$,

$$x^b(\tilde{\gamma}_a(\lambda)) = \begin{cases} \lambda + \text{constant} & \text{if } a = b \\ \text{constant} & \text{if } a \neq b \end{cases}$$

From this calculation we see that $\{\mathrm{d}x^a\}$ is the dual basis to the basis $\{\partial_a\}$, and hence it spans the cotangent space.

c) i) Using the chain rule

$$\begin{aligned}
\partial_a &= \frac{\partial}{\partial x^a}\bigg|_{(x^b, b\neq a)} \\
&= \left(\frac{\partial y^c}{\partial x^a}\bigg|_{(x^b, b\neq a)} \right) \frac{\partial}{\partial y^c}\bigg|_{(y^d, d\neq c)} \\
&= \left(\frac{\partial y^c}{\partial x^a}\bigg|_{(x^b, b\neq a)} \right) \partial'_c \\
&= \begin{cases} \partial'_a + \frac{\partial f}{\partial x^a}\partial'_n & a = 0, \ldots, n-1 \\ \frac{\partial f}{\partial x^n}\partial'_n & a = n \end{cases}
\end{aligned}$$

and so

$$\begin{aligned}
\partial'_n &= \left(\frac{\partial f}{\partial x^n}\right)^{-1} \partial_n \\
\partial'_a &= \partial_a - \left(\frac{\partial f}{\partial x^n}\right)^{-1} \left(\frac{\partial f}{\partial x^a}\right) \partial_n \qquad a = 0, \ldots, n-1
\end{aligned}$$

ii) Again, using the chain rule, we can calculate

$$
\begin{aligned}
\mathrm{d}y^a &= \left(\left. \frac{\mathrm{d}y^a}{\mathrm{d}x^b} \right|_{(x^c, c\neq b)} \right) \mathrm{d}x^b \\
&= \begin{cases} \delta^a_b \mathrm{d}x^b = \mathrm{d}x^a & a = 0, \dots n-1 \\ \left(\frac{\partial f}{\partial x^b} \right) \mathrm{d}x^b & a = n \end{cases}
\end{aligned}
$$

---

4. a)

For H_f to define a tensor at a point p it must define a multilinear (in this case bilinear) map on $(T_p(\mathcal{M}))^2$. Since H_f is symmetric it suffices to check linearity in one argument.

For vector fields X, Y, Z and we can compute

$$
\begin{aligned}
H_f(X, Y+Z) &= \frac{1}{2} X(Y(f) + Z(f)) + \frac{1}{2}\Big(Y(X(f)) + Z(X(f)) \Big) \\
&= \frac{1}{2} X(Y(f)) + \frac{1}{2} Y(X(f)) + \frac{1}{2} X(Z(f)) + \frac{1}{2} Z(X(f)) \\
&= H_f(X, Y) + H_f(X, Z)
\end{aligned}
$$

consistent with linearity. However, for a scalar field a, we find

$$
\begin{aligned}
H_f(X, aY) &= \frac{1}{2} X\left(aY(f)\right) + \frac{1}{2} aY\left(X(f)\right) \\
&= a\Big(\frac{1}{2} X(Y(f)) + \frac{1}{2} Y(X(f)) \Big) + \frac{1}{2} X(a) Y(f)
\end{aligned}
$$

If H_f were multilinear then this second term would be absent. So we see that H_f does not generally define a tensor field.

If H_f is to define a tensor at some point p, then we must have

$$
X(a)Y(f) = (X(a))\, \mathrm{d}f(Y) = 0
$$

for *all* scalar fields a and vector fields X and Y. This is only possible if $\mathrm{d}f = 0$. Recalling that, in any local coordinates, the components of $\mathrm{d}f$ are $(\mathrm{d}f)_a = \partial_a f$, we see that the condition for H_f to define a tensor at a point p is that f has a local extremum at p.

b) Now suppose that $\mathrm{d}f\big|_p = 0$, so f is locally extremised at p. Then we can calculate in components

$$
\begin{aligned}
H_f(X, Y) &= \frac{1}{2} X^a \partial_a \left(Y^b \partial_b f \right) + \frac{1}{2} Y^a \partial_a \left(X^b \partial_b f \right) \\
&= X^a Y^b \partial_a \partial_b f + \frac{1}{2} X^a (\partial_a Y^b) \partial_b f + \frac{1}{2} Y^a (\partial_a X^b) \partial_b f \\
&= X^a Y^b \partial_a \partial_b f \\
&= X^a Y^b (H_f)_{ab}
\end{aligned}
$$

so

$$
(H_f)_{ab} = \partial_a \partial_b f = \frac{\partial^2 f}{\partial x^a \partial x^b}
$$

---

5. a) (and **bonus**): given an arbitrary tensor $T^{\mu\nu}$, we can express it in terms of the coordinate induced derivatives ∂_a (see problem 3) as

$$\begin{aligned} T^{\mu\nu} &= T^{ab}(\partial_a)^\mu(\partial_b)^\nu \\ &= \left(T^{ab}(\partial_a)^\mu\right)(\partial_b)^\nu \\ &= \sum_{b=0}^{n-1}\left(T^{ab}(\partial_a)^\mu\right)(\partial_b)^\nu \end{aligned}$$

which is a sum over n products of pairs of vectors as promised.

b) Suppose we work in coordinates: then the components $T^{ab} = X^a Y^b$ can be represented as a matrix. Let the first index label the row and the second index label the column of this matrix. Then it is easy to see that the rows and columns of this matrix are all proportional to one another: for example, the column vectors are given by $X^a Y^b$ for *fixed* b, so each column is proportional to the vector X^a. Similarly, each row is proportional to the vector Y^b.

From these considerations we can see that the determinant of the matrix T^{ab} must vanish. So, a sufficient condition guaranteeing that the tensor $T^{\mu\nu}$ *cannot* be written as a product of a pair of vectors is that the determinant of the matrix of its components T^{ab} is nonzero.

c) Consider the linear map

$$\begin{aligned} \tilde{T} : T_p^*(\mathcal{M}) &\to T_p^(\mathcal{M}) \\ \eta &\mapsto T(\eta, \cdot) \end{aligned}$$

I claim that

$$\operatorname{rank}\tilde{T} \le 1 \qquad \Longleftrightarrow \qquad \exists \text{ vector fields } X, Y \text{ s.t. } T^{\mu\nu} = X^\mu Y^\nu$$

First let's do the $\Longleftarrow$ direction. If $T^{\mu\nu} = X^\mu Y^\nu$ then

$$\tilde{T}(\eta) = \eta(X)Y$$

If $X = 0$ or $Y = 0$ then this is just the trivial map, with 0-dimensional image. Otherwise, the image of $T_p^*(\mathcal{M})$ is the 1-dimensional space

$$\operatorname{Im}\tilde{T} = \{Z \in \mathbb{R}^n \ : \ Z = aY, \ a \in \mathbb{R}\}$$

Next let's do the $\Longrightarrow$ direction. First we have the trivial case: if the dimension of the image of $\tilde{T}$ is zero, then clearly $T = 0$ and we can write $T^{\mu\nu} = X^\mu Y^\nu$ where $X = 0$ and Y is any vector.

Now let's do the non-trivial case, where the dimension of the image of $\tilde{T}$ is one. In this case, there is some nonzero vector Y such that

$$\operatorname{Im}\tilde{T} = \{Z \in \mathbb{R}^n \ : \ Z = aY, \ a \in \mathbb{R}\}$$

That is, all vectors in the image of $\tilde{T}$ are proportional to Y. Moreover, for any covector η we have

$$\tilde{T}(\eta) = f(\eta)Y$$

for some function $f : T_p^*(\mathcal{M}) \to \mathbb{R}$. Since $\tilde{T}$ is linear, f is also linear. By the usual identification of the double dual of a vector space with the original vector space, we can identify f with a vector X, so that

$$\tilde{T}(\eta) = \eta(X)Y$$

from which it follows that

$$T^{\mu\nu} = X^\mu Y^\nu$$

6. a) In rectangular coordinates, the components of the Minkowski metric are diag $(-1, 1, 1, 1)$ and so all derivatives vanish, and hence so do the Christoffel symbols.

b) i) Since $x = \rho \sin\phi$ we have $\mathrm{d}x = (\sin\phi)\mathrm{d}\rho + (\rho\cos\phi)\mathrm{d}\phi$. Similarly, $y = \rho\cos\phi$ so $\mathrm{d}y = (\cos\phi)\mathrm{d}\rho - (\rho\sin\phi)\mathrm{d}\phi$. Hence

$$\begin{aligned} g &= -\mathrm{d}t^2 + \left(\sin^2\phi \mathrm{d}\rho^2 + 2\rho\sin\phi\cos\phi \mathrm{d}\rho\mathrm{d}\phi + \rho^2\cos^2\phi \mathrm{d}\phi^2\right) \\ &\quad + \left(\cos^2\phi \mathrm{d}\rho^2 - 2\rho\sin\phi\cos\phi \mathrm{d}\rho\mathrm{d}\phi + \rho^2\sin^2\phi \mathrm{d}\phi^2\right) + \mathrm{d}z^2 \\ &= -\mathrm{d}t^2 + \mathrm{d}\rho^2 + \rho^2\mathrm{d}\phi^2 + \mathrm{d}z^2 \end{aligned}$$

ii) We can calculate

$$\begin{aligned} \Gamma^t_{ab} &= \frac{1}{2}(g^{-1})^{tc}\left(\partial_a g_{bc} + \partial_b g_{ac} - \partial_c g_{ab}\right) \\ &= -\frac{1}{2}\left(\partial_a g_{bt} + \partial_b g_{at} - \partial_t g_{ab}\right) \end{aligned}$$

using the fact that $(g^{-1})^{ta} = -1$ if $a = t$, and 0 otherwise. We see that all derivatives of g_{at} vanish for any a, since $g_{at} = -1$ or 0. Finally, we see that all the metric components are independent of t, so the final term vanishes. Hence $\Gamma^t_{ab} = 0$ for any a, b.

The other components can be calculated following the same kind of steps.

c) i) We have

$$\begin{aligned} x &= r\sin\theta\cos\phi \\ \Rightarrow \mathrm{d}x &= (\sin\theta\cos\phi)\mathrm{d}r + (r\cos\theta\cos\phi)\mathrm{d}\theta - (r\sin\theta\sin\phi)\mathrm{d}\phi \\ y &= r\sin\theta\sin\phi \\ \Rightarrow \mathrm{d}y &= (\sin\theta\sin\phi)\mathrm{d}r + (r\cos\theta\sin\phi)\mathrm{d}\theta + (r\sin\theta\cos\phi)\mathrm{d}\phi \\ z &= r\cos\theta \\ \Rightarrow \mathrm{d}z &= (\cos\theta)\mathrm{d}r - (r\sin\theta)\mathrm{d}\theta \end{aligned}$$

A little bit of algebra then gives the expression for the metric.

ii) These are fairly tedious calculations, but as an example

$$\begin{aligned} \Gamma^\theta_{ab} &= \frac{1}{2}(g^{-1})^{\theta c}\left(\partial_a g_{bc} + \partial_b g_{ac} - \partial_c g_{ab}\right) \\ &= \frac{1}{2}r^{-2}\left(\partial_a g_{b\theta} + \partial_b g_{a\theta} - \partial_\theta g_{ab}\right) \end{aligned}$$

now we see that $\partial_a g_{b\theta}$ is nonzero only if $a = r$ and $b = \theta$, in which case we find

$$\Gamma^\theta_{r\theta} = \frac{1}{2}r^{-2}\cdot 2r = r^{-1}$$

Similarly, $\partial_b g_{a\theta}$ is nonzero only if $a = \theta$ and $b = r$, in which case we find that $\Gamma^\theta_{\theta r} = r^{-1}$.

Finally, we note that $\partial_\theta g_{ab}$ is nonzero only if $a = b = \phi$, in which case we find

$$\Gamma^\theta_{\phi\phi} = -\frac{1}{2}r^{-2}\cdot 2r^2\sin\theta\cos\theta = -\sin\theta\cos\theta$$

The other computations are all similar.

7. a) Since Y is parallel-transported along the integral curves of X, we have

$$\nabla_X Y = 0$$

writing this in components:

$$\begin{aligned}\nabla_X Y &= X^a \nabla_a (Y^b \partial_b) \\ &= X^a (\partial_a Y^b)\partial_b + X^a Y^b \Gamma^c_{ab} \partial_c \\ &= \left(X^b \partial_b Y^a + X^b Y^c \Gamma^a_{bc}\right) \partial_a\end{aligned}$$

so if $\nabla_X Y = 0$ then

$$X^b \partial_b Y^a + \left(X^b \Gamma^a_{bc}\right) Y^c = 0$$

Finally we note that, since X is tangent to a curve parametrised by s, we have $X^b \partial_b = \frac{\mathrm{d}}{\mathrm{d}s}$.

b) First note that

$$\begin{aligned}Xg(A,B) &= X^a \partial_a \left(g_{bc} A^b B^c\right) \\ &= X^a \nabla_a \left(g_{bc} A^b B^c\right) \\ &= g_{bc} \left(X^a (\nabla_a A^b) B^c + A^b X^a (\nabla_a B^c)\right) \\ &= g(\nabla_X A, B) + g(A, \nabla_X B)\end{aligned}$$

Now we can compute

$$\frac{\mathrm{d}}{\mathrm{d}s} g(X,X) = X\left(g(X,X)\right) = g(\nabla_X X, X) + g(X, \nabla_X X) = 2g(\nabla_X X, X) = 0$$

Similarly

$$\frac{\mathrm{d}}{\mathrm{d}s} g(Y,Y) = 2g(\nabla_X Y, Y) = 0$$

and

$$\frac{\mathrm{d}}{\mathrm{d}s} g(X,Y) = g(\nabla_X X, Y) + g(X, \nabla_X Y) = 0$$

c) As above, we can compute

$$\frac{\mathrm{d}}{\mathrm{d}s} g(Y,Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z) = 0$$

d) We have $K(X) = K_\mu X^\mu$, so

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}s} K(X) &= X^\mu \nabla_\mu \left(K_\nu X^\nu\right) \\ &= X^\mu X^\nu \nabla_\mu K_\nu + K_\mu (\nabla_X X)^\mu \\ &= \frac{1}{2}\left(X^\mu X^\nu + X^\nu X^\mu\right) \nabla_\mu K_\nu \\ &= \frac{1}{2} X^\mu X^\nu \left(\nabla_\mu K_\nu + \nabla_\nu K_\mu\right) \\ &= 0\end{aligned}$$

8. a)

Using the normal formula for changing coordinates

$$
\begin{aligned}
[X,Y]^{a'} &= \frac{\partial y^{a'}}{\partial x^a}[X,Y]^a \\
&= \frac{\partial y^{a'}}{\partial x^a}\left(X(Y^a) - Y(X^a)\right) \\
&= X\left(\frac{\partial y^{a'}}{\partial x^a}Y^a\right) - X^b Y^a \frac{\partial^2 y^{a'}}{\partial x^a \partial x^b} - Y\left(\frac{\partial y^{a'}}{\partial x^a}X^a\right) + Y^b X^a \frac{\partial^2 y^{a'}}{\partial x^a \partial x^b} \\
&= X(Y^{a'}) - Y(X^{a'}) \\
&= X^{b'}\frac{\partial Y^{a'}}{\partial y^{b'}} - Y^{b'}\frac{\partial X^{a'}}{\partial y^{b'}}
\end{aligned}
$$

b) i) We need to check that $\mathrm{d}\eta$ defines a multilinear (in this case, bilinear) map from pairs of vector fields to the real line. For vector fields X, Y and Z and a scalar field a, we can check linearity in the first argument:

$$
\begin{aligned}
\mathrm{d}\eta(aX+Y,Z) &= (aX+Y)\left(\eta(Z)\right) - Z\left(\eta(aX+Y)\right) - \eta\left([aX+Y,Z]\right) \\
&= (aX+Y)\left(\eta(Z)\right) - Z\left(a\eta(X)+\eta(Y)\right) - \eta\left(a[X,Z] - Z(a)X + [Y,Z]\right) \\
&= aX\left(\eta(Z)\right) + Y\left(\eta(Z)\right) - Z(a)\eta(X) - aZ\left(\eta(X)\right) - Z\left(\eta(Y)\right) - a\eta([X,Z]) \\
&\quad + Z(a)\eta(X) - \eta([Y,Z]) \\
&= a\mathrm{d}\eta(X,Z) + \mathrm{d}\eta(Y,Z)
\end{aligned}
$$

so we have linearity in the first argument.

Next we note that $\mathrm{d}\eta$ is antisymmetric in its two arguments. Hence, since it it linear in its first argument, it is also linear in the second argument, and therefore it defines a tensor field.

ii) Calculating in components

$$
\begin{aligned}
\mathrm{d}\eta(X,Y) &= X^a\partial_a\left(\eta_b Y^b\right) - Y^a\partial_a\left(\eta_b X^b\right) - \eta_a\left(X^b\partial_b Y^a - Y^b\partial_b X^a\right) \\
&= X^aY^b\left(\partial_a\eta_b - \partial_b\eta_a\right) + \eta_b\left(X^a\partial_a Y^b - Y^a\partial_a X^b\right) - \eta_a\left(X^b\partial_b Y^a - Y^b\partial_b X^a\right) \\
&= X^aY^b\left(\partial_a\eta_b - \partial_b\eta_a\right) \\
&= X^aY^b\left(\mathrm{d}\eta\right)_{ab}
\end{aligned}
$$

and so

$$(\mathrm{d}\eta)_{ab} = \partial_a\eta_b - \partial_b\eta_a$$

iii) There are at least two ways to do this question. One way is to use normal coordinates: in normal coordinates at some point p, derivatives and partial derivatives are the same thing, so at p

$$(\mathrm{d}\eta)_{ab} = \nabla_a\eta_b - \nabla_b\eta_a$$

We have already shown that $\mathrm{d}\eta$ is a tensor field, and the right hand side of the equation above is also evidently the components of the tensor field

$$\nabla_\mu\eta_\nu - \nabla_\nu\eta_\mu$$

Since the components (in normal coordinates) of these two tensor fields agree at p, their components at p must agree in *all* coordinate systems, i.e. the two tensor fields must be equal at p. But p was arbitrary, so this argument establishes the equality of the two tensor fields everywhere.

An alternative approach is to use the torsion-free property of the Levi-Civita connection: in arbitrary coordinates we can write

$$\nabla_a\eta_b - \nabla_b\eta_a = \partial_a\eta_b - \partial_b\eta_a - \Gamma^c_{ab}\eta_c + \Gamma^c_{ba}\eta_c = (\mathrm{d}\eta)_{ab}$$

using the symmetry of the Christoffel symbols.

iv) If $\eta = \mathrm{d}f$ then we can compute

$$\begin{aligned}(\mathrm{dd}f)(X,Y) &= X(Y(f)) - Y(X(f)) - \mathrm{d}f\,([X,Y]) \\ &= X(Y(f)) - Y(X(f)) - [X,Y](f) \\ &= 0\end{aligned}$$

Since this holds for arbitrary X, Y, we have $\mathrm{dd}f = 0$.

Alternatively we could work in coordinates:

$$(\mathrm{dd}f)_{ab} = \partial_a\partial_b f - \partial_b\partial_a f = 0$$

c) i) If we had made the other choice of orientation then, instead of $(\mathrm{d}\eta)_{01}$ in the integrand we would have $(\mathrm{d}\eta)_{10}$. But, since $\mathrm{d}\eta$ is antisymmetric, this differs only by a minus sign.

ii) If we work in the coordinates (y^0, y^1) then the integrand is

$$\begin{aligned}(\mathrm{d}\eta\big|_\Sigma)'_{01} &= (\mathrm{d}\eta|_\Sigma)\left(\frac{\partial}{\partial y^0}, \frac{\partial}{\partial y^1}\right) \\ &= (\mathrm{d}\eta|_\Sigma)\left(\frac{\partial x^a}{\partial y^0}\frac{\partial}{\partial x^a}\ ,\ \frac{\partial x^b}{\partial y^1}\frac{\partial}{\partial x^b}\right)\end{aligned}$$

where derivatives with respect the y coordinates are taken while holding the other y coordinates constant, and derivatives w.r.t. the x coordinates are taken holding the other x coordinates constant. Continuing the calculation:

$$\begin{aligned}(\mathrm{d}\eta\big|_\Sigma)'_{01} &= \left(\frac{\partial x^a}{\partial y^0}\right)\left(\frac{\partial x^b}{\partial y^1}\right)(\mathrm{d}\eta\big|_\Sigma)\left(\frac{\partial}{\partial x^a}, \frac{\partial}{\partial x^b}\right) \\ &= \left(\frac{\partial x^a}{\partial y^0}\right)\left(\frac{\partial x^b}{\partial y^1}\right)(\mathrm{d}\eta\big|_\Sigma)_{ab} \\ &= \left(\frac{\partial x^0}{\partial y^0}\right)\left(\frac{\partial x^1}{\partial y^1}\right)(\mathrm{d}\eta\big|_\Sigma)_{01} + \left(\frac{\partial x^1}{\partial y^0}\right)\left(\frac{\partial x^0}{\partial y^1}\right)(\mathrm{d}\eta\big|_\Sigma)_{10} \\ &= \left(\left(\frac{\partial x^0}{\partial y^0}\right)\left(\frac{\partial x^1}{\partial y^1}\right) - \left(\frac{\partial x^1}{\partial y^0}\right)\left(\frac{\partial x^0}{\partial y^1}\right)\right)(\mathrm{d}\eta\big|_\Sigma)_{01} \\ &= |J|(\mathrm{d}\eta\big|_\Sigma)_{01}\end{aligned}$$

iii) We have

$$\begin{aligned}\int_\Sigma \mathrm{d}\eta &= \int_{x^0=0}^1\int_{x^1=0}^1 (\mathrm{d}\eta)_{01}\mathrm{d}x^0\mathrm{d}x^1 \\ &= \int_{x^0=0}^1\int_{x^1=0}^1 (\partial_0\eta_1 - \partial_1\eta_0)\mathrm{d}x^0\mathrm{d}x^1 \\ &= \int_{x_1=0}^1 \left(\eta_1\mathrm{d}x^1\big|_{x^0=0} - \eta_1\mathrm{d}x^1\big|_{x^0=1}\right) + \int_{x_0=0}^1 \left(\eta_0\mathrm{d}x^0\big|_{x^1=1} - \eta_0\mathrm{d}x^0\big|_{x^1=0}\right) \\ &= \int_{\partial\Sigma}\eta\big|_{\partial\Sigma} = \int_{\partial\Sigma}\eta\end{aligned}$$