Peize Liu
*St. Peter's College*
*University of Oxford*

**Problem Sheet 1**

# B2: Symmetry & Relativity

30 October, 2020

Einstein summation convention is assumed, with indices taking values from 1 to $N$.

**Question 1**

If we have two successive transformations from $u^i = u^i\left(x^1, x^2, ..., x^N\right)$ to $v^i = v^i\left(y^1, y^2, ..., y^N\right)$ and from $v^i$ to $w^i = w^i\left(z^1, z^2, \cdots z^N\right)$, with $i = 1, 2, \cdots N$

$$v^i = \frac{\partial y^i}{\partial x^j} u^j$$

and

$$w^i = \frac{\partial z^i}{\partial y^j} v^j$$

show that we can perform the transformation in one step via

$$w^i = \frac{\partial z^i}{\partial x^j} u^j$$

*Proof.* This is essentially a chain rule:

$$w^i = \frac{\partial z^i}{\partial y^j} v^j = \frac{\partial z^i}{\partial y^j}\frac{\partial y^j}{\partial x^k} u^k = \frac{\partial z^i}{\partial x^k} u^k = \frac{\partial z^i}{\partial x^j} u^j$$

To justify the chain rule in the third equality, let $M$ be the $N$-dimensional (background) manifold. $\varphi = (x^1, ..., x^N)$, $\psi = (y^1, ..., y^N)$ and $\chi = (z^1, ..., z^N)$ are coordinates charts from $U \subseteq M$ to $\mathbb{R}$. Let $r^1, ..., r^N$ be the standard coordinates of $\mathbb{R}^N$. Then

$$\begin{aligned}\frac{\partial z^i}{\partial x^k} &= \frac{\partial(z^i \circ \varphi^{-1})}{\partial r^k} = \frac{\partial\left((z^i \circ \psi^{-1}) \circ (\psi \circ \varphi^{-1})\right)}{\partial r^k} = \frac{\partial(z^i \circ \psi^{-1})}{\partial r^j}\frac{\partial(y^j \circ \varphi^{-1})}{\partial r^k} \qquad \text{(chain rule in } \mathbb{R}^N\text{)}\\ &= \frac{\partial z^i}{\partial y^j}\frac{\partial y^j}{\partial x^k}\end{aligned}$$

□

**Question 2**

If $A^{ij}_k$ is a mixed tensor, $B^{ij}_k$ is another tensor of the same kind, and $\alpha$ and $\beta$ are scalar invariants, show that $\alpha A^{ij}_k + \beta B^{ij}_k$ is yet another tensor of the same kind.

*Proof.* Algebrically, it follows from that the set of all type $(2,1)$ tensors forms a vector space $T^2_1(V) = V \otimes V \otimes V^*$, where $V$ is the tangent space $\mathrm{T}_p M$. If we use the definition given by transformation rule, we can check

$$\alpha \tilde{A}^{ij}_k + \beta \tilde{B}^{ij}_k = \alpha \frac{\partial \tilde{x}^i}{\partial x^\ell}\frac{\partial \tilde{x}^j}{\partial x^m}\frac{\partial x^n}{\partial \tilde{x}^k} A^{ij}_k + \beta \frac{\partial \tilde{x}^i}{\partial x^\ell}\frac{\partial \tilde{x}^j}{\partial x^m}\frac{\partial x^n}{\partial \tilde{x}^k} B^{\ell m}_n = \frac{\partial \tilde{x}^i}{\partial x^\ell}\frac{\partial \tilde{x}^j}{\partial x^m}\frac{\partial x^n}{\partial \tilde{x}^k}\left(\alpha A^{\ell m}_n + \beta B^{\ell m}_n\right) = \overline{\alpha A^{ij}_k + \beta B^{ij}_k}$$

□

**Question 3**

If $A^i_j$ are the components of a mixed tensor, show that $A^i_i$ transforms as a scalar invariant.

*Proof.* $A^i_i$ is the contraction of a type $(1,1)$ tensor $A^i_j$:

$$C^1_1(T) = C^1_1\left(A^i_j \frac{\partial}{\partial x^i} \otimes \mathrm{d}x^j\right) = A^i_j \mathrm{d}x^j\left(\frac{\partial}{\partial x^i}\right) = A^i_j \frac{\partial x^j}{\partial x^i} = A^i_j \delta^j_i = A^i_i$$

The contraction of a type $(1,1)$ tensor is a type $(0,0)$ tensor, which is a scalar. So $A^i_i$ surely transforms as a scalar invariant. □

I think the point of the problem was to show this, rather than assuming it

**Question 4**

Assuming $x$ and $y$ transform as the components of a Euclidean vector, determine which of the following matrices are tensors:

$$A^{ij} = \begin{pmatrix} x^2 & xy \\ xy & y^2 \end{pmatrix}, \quad B^{ij} = \begin{pmatrix} xy & y^2 \\ x^2 & -xy \end{pmatrix}, \quad C^{ij} = \begin{pmatrix} y^2 & xy \\ xy & x^2 \end{pmatrix}$$

[*based on E Butkov, Mathematical Physics*]

*Proof.* *Geometrically speaking, tensors are specified at a fixed point in a manifold. So I think the correct question to ask here is that if the matrices are* ***tensor fields****.*

None of $A, B, C$ looks like a tensor field as they seem not multi-linear with respect to the coordinates. Yet we can verify if the matrices are invariant under orthogonal transformations of $\mathbb{R}^2$ (as done in the lecture recordings).

Any orthogonal matrix $O \in \mathrm{O}(2)$ is of the form

$$O = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \qquad \text{or} \qquad O = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} \qquad (\theta \in [0, 2\pi])$$

Consider the change of variables

$$\begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} = \begin{pmatrix} \cos\theta & \mp\sin\theta \\ \sin\theta & \pm\cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

which implies that

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ \mp\sin\theta & \pm\cos\theta \end{pmatrix} \begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} = \begin{pmatrix} \tilde{x}\cos\theta + \tilde{y}\sin\theta \\ \pm(-\tilde{x}\sin\theta + \tilde{y}\cos\theta) \end{pmatrix}$$

Let $\widetilde{A}(\tilde{x}, \tilde{y}) = A(x, y)$. Then

$$\begin{aligned} \widetilde{A} &= \begin{pmatrix} (\tilde{x}\cos\theta + \tilde{y}\sin\theta)^2 & \pm(\tilde{x}\cos\theta + \tilde{y}\sin\theta)(-\tilde{x}\sin\theta + \tilde{y}\cos\theta) \\ \pm(\tilde{x}\cos\theta + \tilde{y}\sin\theta)(-\tilde{x}\sin\theta + \tilde{y}\cos\theta) & (-\tilde{x}\sin\theta + \tilde{y}\cos\theta)^2 \end{pmatrix} \\ &= \begin{pmatrix} \tilde{x}^2\cos^2\theta + \tilde{y}^2\sin^2\theta + \tilde{x}\tilde{y}\sin 2\theta & \pm\left(\dfrac{\tilde{y}^2 - \tilde{x}^2}{2}\sin 2\theta + \tilde{x}\tilde{y}\cos 2\theta\right) \\ \pm\left(\dfrac{\tilde{y}^2 - \tilde{x}^2}{2}\sin 2\theta + \tilde{x}\tilde{y}\cos 2\theta\right) & \tilde{x}^2\sin^2\theta + \tilde{y}^2\cos^2\theta - \tilde{x}\tilde{y}\sin 2\theta \end{pmatrix} \end{aligned}$$

On the other hand, after change of basis

$$O^{\mathrm{T}}AO = \begin{pmatrix} \cos\theta & \sin\theta \\ \mp\sin\theta & \pm\cos\theta \end{pmatrix} \begin{pmatrix} x^2 & xy \\ xy & y^2 \end{pmatrix} \begin{pmatrix} \cos\theta & \mp\sin\theta \\ \sin\theta & \pm\cos\theta \end{pmatrix} = \begin{pmatrix} \tilde{x}^2\cos^2\theta + \tilde{y}^2\sin^2\theta + \tilde{x}\tilde{y}\sin 2\theta & \pm\left(\dfrac{\tilde{y}^2 - \tilde{x}^2}{2}\sin 2\theta + \tilde{x}\tilde{y}\cos 2\theta\right) \\ \pm\left(\dfrac{\tilde{y}^2 - \tilde{x}^2}{2}\sin 2\theta + \tilde{x}\tilde{y}\cos 2\theta\right) & \tilde{x}^2\sin^2\theta + \tilde{y}^2\cos^2\theta - \tilde{x}\tilde{y}\sin 2\theta \end{pmatrix}$$

Hence $\widetilde{A} = O^{\mathrm{T}}AO$. $A$ is invariant under orthogonal transformations.

I would not say that A is inv. Rather, it transforms but it does so in a covariant way

For matrix $B$,

$$\widetilde{B} = \begin{pmatrix} \pm\left(\dfrac{\tilde{y}^2 - \tilde{x}^2}{2}\sin 2\theta + \tilde{x}\tilde{y}\cos 2\theta\right) & \tilde{x}^2\sin^2\theta + \tilde{y}^2\cos^2\theta - \tilde{x}\tilde{y}\sin 2\theta \\ \tilde{x}^2\cos^2\theta + \tilde{y}^2\sin^2\theta + \tilde{x}\tilde{y}\sin 2\theta & \mp\left(\dfrac{\tilde{y}^2 - \tilde{x}^2}{2}\sin 2\theta + \tilde{x}\tilde{y}\cos 2\theta\right) \end{pmatrix}$$

$$O^{\mathrm{T}}BO = \begin{pmatrix} \cos\theta & \sin\theta \\ \mp\sin\theta & \pm\cos\theta \end{pmatrix} \begin{pmatrix} xy & y^2 \\ x^2 & -xy \end{pmatrix} \begin{pmatrix} \cos\theta & \mp\sin\theta \\ \sin\theta & \pm\cos\theta \end{pmatrix} = \begin{pmatrix} \left(\dfrac{\tilde{x}^2 + \tilde{y}^2}{2}\sin 2\theta + \tilde{x}\tilde{y}\cos 2\theta\right) & \pm\left(\tilde{x}^2\cos^2\theta - \tilde{y}^2\sin^2\theta + \tilde{x}\tilde{y}\sin 2\theta\right) \\ \pm\left(\tilde{x}^2\sin^2\theta - \tilde{y}^2\cos^2\theta - \tilde{x}\tilde{y}\sin 2\theta\right) & \left(-\dfrac{\tilde{x}^2 + \tilde{y}^2}{2}\sin 2\theta - \tilde{x}\tilde{y}\cos 2\theta\right) \end{pmatrix}$$

Hence $\widetilde{B} = O^{\mathrm{T}}BO$. $B$ is not invariant under orthogonal transformations.

For matrix $C$,

$$\widetilde{C} = \begin{pmatrix} \widetilde{x}^2 \sin^2\theta + \widetilde{y}^2 \cos^2\theta - \widetilde{x}\widetilde{y}\sin 2\theta & \pm\left(\dfrac{\widetilde{y}^2 - \widetilde{x}^2}{2}\sin 2\theta + \widetilde{x}\widetilde{y}\cos 2\theta\right) \\ \pm\left(\dfrac{\widetilde{y}^2 - \widetilde{x}^2}{2}\sin 2\theta + \widetilde{x}\widetilde{y}\cos 2\theta\right) & \widetilde{x}^2 \cos^2\theta + \widetilde{y}^2 \sin^2\theta + \widetilde{x}\widetilde{y}\sin 2\theta \end{pmatrix}$$

$$O^{\mathrm{T}}CO = \begin{pmatrix} \cos\theta & \sin\theta \\ \mp\sin\theta & \pm\cos\theta \end{pmatrix}\begin{pmatrix} y^2 & xy \\ xy & x^2 \end{pmatrix}\begin{pmatrix} \cos\theta & \mp\sin\theta \\ \sin\theta & \pm\cos\theta \end{pmatrix} = \begin{pmatrix} \widetilde{x}^2 \cos^2\theta + \widetilde{y}^2 \sin^2\theta + \widetilde{x}\widetilde{y}\sin 2\theta & \pm\left(\dfrac{\widetilde{x}^2 - \widetilde{y}^2}{2}\sin 2\theta + \widetilde{x}\widetilde{y}\cos 2\theta\right) \\ \pm\left(\dfrac{\widetilde{x}^2 - \widetilde{y}^2}{2}\sin 2\theta + \widetilde{x}\widetilde{y}\cos 2\theta\right) & \widetilde{x}^2 \sin^2\theta + \widetilde{y}^2 \cos^2\theta - \widetilde{x}\widetilde{y}\sin 2\theta \end{pmatrix}$$

Hence $\widetilde{C} = O^{\mathrm{T}}CO$. $C$ is not invariant under orthogonal transformations. □

**Remark.** We observe that $A = x^i \dfrac{\partial}{\partial x^i} \otimes x^j \dfrac{\partial}{\partial x^j}$ is a tensor product of two vector fields. It makes me wonder if $A$ is still a tensor field in the most general sense...

What is the most general sense?

**Question 5**

Show that if the components of a contravariant vector vanish in one coordinate system, they vanish in all coordinate systems. What can be said of two contravariant vectors whose components are equal in one coordinate system?

*Proof.* If $A^i = 0$ for $i \in \{1, \dots, N\}$. Then it represents a zero tangent vector $A^i \dfrac{\partial}{\partial x^i} = 0 \in \mathrm{T}_p M$. Hence it is zero in any coordinate charts.

Suppose that $A^i = k \neq 0$ for $i \in \{1, \dots, N\}$. Consider a new coordinate chart $(y^1, \dots, y^N)$, where $y^1 = 2x^1$ and $y^i = x^i$ for $i \in \{1, \dots, N\}$. Then $A^i \dfrac{\partial}{\partial x^i} = \widetilde{A}^j \dfrac{\partial}{\partial y^j}$ implies that $(\widetilde{A}^1, \dots, \widetilde{A}^N) = (k/2, k, \dots, k)$. So after transformation the components are no longer equal. □

this is not what the problem asked

You should have picked 2 vectors $A^i = B^i$ and showed $\widetilde{A}^i = \widetilde{B}^i$

**Question 6**

Let $A_{ij}$ be a skew-symmetric tensor with $A_{ij} = -A_{ji}$, and $S_{ij}$ a symmetric tensor with $S_{ij} = S_{ji}$. Show that the symmetry properties are preserved in coordinate transformations. Also show that the quantities with raised indices, $A^{ij}$ and $S^{ij}$, possess the same properties. From this, show that $A^{ij}S_{ij} = 0$ and $A_{ij}S^{ij} = 0$.

*Proof.* Consider the type $(0,2)$ skew-symmetric tensor

$$A_{ij}\mathrm{d}x^i \wedge \mathrm{d}x^j = \widetilde{A}_{ij}\mathrm{d}\widetilde{x}^i \wedge \mathrm{d}\widetilde{x}^j$$

We have

$$\widetilde{A}_{ji} = \frac{\partial x^\ell}{\partial \widetilde{x}^j}\frac{\partial x^k}{\partial \widetilde{x}^i}A_{\ell k} = -\frac{\partial x^\ell}{\partial \widetilde{x}^j}\frac{\partial x^k}{\partial \widetilde{x}^i}A_{k\ell} = -\widetilde{A}_{ij}$$

Consider the type $(0,2)$ symmetric tensor

$$S_{ij}\mathrm{d}x^i \otimes \mathrm{d}x^j = \widetilde{S}_{ij}\mathrm{d}\widetilde{x}^i \otimes \mathrm{d}\widetilde{x}^j$$

We have

$$\widetilde{S}_{ji} = \frac{\partial x^\ell}{\partial \widetilde{x}^j}\frac{\partial x^k}{\partial \widetilde{x}^i}S_{\ell k} = \frac{\partial x^\ell}{\partial \widetilde{x}^j}\frac{\partial x^k}{\partial \widetilde{x}^i}S_{k\ell} = -\widetilde{S}_{ij}$$

Consider the type $(2,0)$ skew-symmetric tensor

$$A^{ij}\frac{\partial}{\partial x^i} \wedge \frac{\partial}{\partial x^j} = \widetilde{A}^{ij}\frac{\partial}{\partial \widetilde{x}^i} \wedge \frac{\partial}{\partial \widetilde{x}^j}$$

We have

$$\widetilde{A}^{ji} = \frac{\partial \widetilde{x}^j}{\partial x^\ell}\frac{\partial \widetilde{x}^i}{\partial x^k}A^{\ell k} = -\frac{\partial \widetilde{x}^j}{\partial x^\ell}\frac{\partial \widetilde{x}^i}{\partial x^k}A^{k\ell} = -\widetilde{A}^{ij}$$

Consider the type $(2,0)$ symmetric tensor

$$S^{ij}\frac{\partial}{\partial x^i} \otimes \frac{\partial}{\partial x^j} = \widetilde{S}^{ij}\frac{\partial}{\partial \widetilde{x}^i} \otimes \frac{\partial}{\partial \widetilde{x}^j}$$

We have

$$\widetilde{S}^{ji} = \frac{\partial \widetilde{x}^j}{\partial x^\ell}\frac{\partial \widetilde{x}^i}{\partial x^k} S^{\ell k} = \frac{\partial \widetilde{x}^j}{\partial x^\ell}\frac{\partial \widetilde{x}^i}{\partial x^k} S^{k\ell} = \widetilde{S}^{ij}$$

If $A$ is skew-symmetric and $S$ is symmetric, then

$$A^{ij}S_{ij} = -A^{ji}S_{ij} = -A^{ji}S_{ji} = -A^{ij}S_{ij} \qquad A_{ij}S^{ij} = -A_{ji}S^{ij} = -A_{ji}S^{ji} = -A_{ij}S^{ij}$$

Hence $A^{ij}S_{ij} = 0$ and $A_{ij}S^{ij} = 0$. □

**Question 7**

Let $C^{k\ell} = A^{ijk}B^\ell_{ij}$ be a rank-2 contravariant tensor given by contracting the $N^3$ functions $A^{ijk}$ with the tensor $B^\ell_{mn}$, which is symmetric in the $mn$ indices but otherwise arbitrary, i.e., $B^\ell_{mn} = B^\ell_{nm}$. Show that $A^{ijk} + A^{jik}$ is a rank-3 contravariant tensor. Give reasons why the same is not true for $A^{ijk}$ or $A^{jik}$ separately.

*Proof.* First we prove the following lemma:

*Suppose that $A^{ijk}B^\ell_{ij} = 0$ for any type (1,2) tensor $B$ symmetric in the two lower indices. Then $A^{ijk} + A^{jik} = 0$.*

For $a, b \in \{1, \dots, N\}$, consider a tensor $B$ where $B^\ell_{ij} = 1$ whenever $(i, j) = (a, b)$ and $(b, a)$, and $B^\ell_{ij} = 0$ otherwise. Then we have $A^{ijk}B^\ell_{ij} = A^{abk} + A^{bak} = 0$. We deduce that $A^{ijk} + A^{jik} = 0$ for all $i, j, k \in \{1, \dots, N\}$.

Now we return to the problem. Using the fact that $C$ is a type $(2,0)$ tensor and that $B$ is a type $(1,2)$ tensor, we have the transformation rule

$$\widetilde{A}^{ijk}\frac{\partial \widetilde{x}^\ell}{\partial x^n}\frac{\partial x^p}{\partial \widetilde{x}^i}\frac{\partial x^q}{\partial \widetilde{x}^j} B^n_{pq} = \widetilde{A}^{ijk}\widetilde{B}^\ell_{ij} = \widetilde{C}^{k\ell} = \frac{\partial \widetilde{x}^k}{\partial x^m}\frac{\partial \widetilde{x}^\ell}{\partial x^n} C^{mn} = \frac{\partial \widetilde{x}^k}{\partial x^m}\frac{\partial \widetilde{x}^\ell}{\partial x^n} A^{pqm}B^n_{pq}$$

Since the coordinate transformation is invertible, $\partial \widetilde{x}^\ell / \partial x^n \neq 0$. Hence we have

$$B^n_{pq}\left(\frac{\partial x^p}{\partial \widetilde{x}^i}\frac{\partial x^q}{\partial \widetilde{x}^j}\widetilde{A}^{ijk} - \frac{\partial \widetilde{x}^k}{\partial x^m}A^{pqm}\right) = 0 \tag{$*$}$$

This holds for all type $(1,2)$ tensors symmetric in the lower indices. By the lemma above we have

$$\frac{\partial x^p}{\partial \widetilde{x}^i}\frac{\partial x^q}{\partial \widetilde{x}^j}\widetilde{A}^{ijk} - \frac{\partial \widetilde{x}^k}{\partial x^m}A^{pqm} + \frac{\partial x^q}{\partial \widetilde{x}^i}\frac{\partial x^p}{\partial \widetilde{x}^j}\widetilde{A}^{ijk} - \frac{\partial \widetilde{x}^k}{\partial x^m}A^{qpm} = 0$$

By relabeling the indices and combining we obtain

$$\widetilde{A}^{ijk} + \widetilde{A}^{jik} = \frac{\partial \widetilde{x}^k}{\partial x^m}\frac{\partial \widetilde{x}^i}{\partial x^p}\frac{\partial \widetilde{x}^j}{\partial x^q}\left(A^{pqm} + A^{qpm}\right)$$

Hence $A^{ijk} + A^{jik}$ transforms like a type $(3,0)$ tensor.

The conclusion does not apply to $A^{ijk}$. Let $C$ be a non-zero anti-symmetric type $(3,0)$ tensor. Then $C^{ijk} + C^{jik} = 0$ and hence $C^{ijk}B^\ell_{ij} = 0$ for any type $(1,2)$ tensor $B$ symmetric in the two lower indices. The equation $(*)$ is satisfied by

$$\frac{\partial x^p}{\partial \widetilde{x}^i}\frac{\partial x^q}{\partial \widetilde{x}^j}\widetilde{A}^{ijk} - \frac{\partial \widetilde{x}^k}{\partial x^m}A^{pqm} = C^{pqk}$$

Then

$$\widetilde{A}^{ijk} = \frac{\partial \widetilde{x}^k}{\partial x^m}\frac{\partial \widetilde{x}^i}{\partial x^p}\frac{\partial \widetilde{x}^j}{\partial x^q}A^{pqm} + \frac{\partial \widetilde{x}^i}{\partial x^p}\frac{\partial \widetilde{x}^j}{\partial x^q}C^{pqk}$$

$A^{ijk}$ does not transform like a type $(3,0)$ tensor for a suitably chosen anti-symmetric $C$. □

**Question 8**

In this problem, we will consider a transformation from Cartesian to polar coordinate systems in two Euclidean dimensions. Let $x^1 = x$ and $x^2 = y$ for the Cartesian system, and $\hat{x}^1 = r$ and $\hat{x}^2 = \theta$ for the polar, with the transformation

$$x^1 = r\cos\theta = \hat{x}^1 \cos\hat{x}^2$$
$$x^2 = r\sin\theta = \hat{x}^1 \sin\hat{x}^2$$

The metric for the Cartesian system is $g_{ij} = \delta_{ij}$. Derive the metric tensor $\hat{g}_{ij}$ for the polar coordinate system, its reciprocal $\hat{g}^{ij}$, and the covariant polar components $\hat{x}_1$ and $\hat{x}_2$ in terms of $r$ and $\theta$. Why might it not be appropriate to calculate a length from the origin to a point specified by finite values of $r$ and $\theta$ using these covariant components? Show that the components of the metrics $g_{ij}$ and $\hat{g}_{ij}$ do not change under rotations of the coordinate system through a fixed angle $\alpha$ around the origin.

*Proof.* Note the the metric is a type $(0,2)$ tensor. It transforms via

$$\widehat{g}_{ij} = \frac{\partial x^k}{\partial \widehat{x}^i}\frac{\partial x^\ell}{\partial \widehat{x}^j} g_{k\ell} = \frac{\partial x^k}{\partial \widehat{x}^i}\frac{\partial x^\ell}{\partial \widehat{x}^j}\delta_{k\ell} = \frac{\partial x^k}{\partial \widehat{x}^i}\frac{\partial x^k}{\partial \widehat{x}^j}$$

Since $x = r\cos\theta$ and $y = r\sin\theta$, the Jacobian matrix of the transformation:

$$\frac{\partial(x,y)}{\partial(r,\theta)} = \begin{pmatrix} \partial x/\partial r & \partial x/\partial\theta \\ \partial y/\partial r & \partial y/\partial\theta \end{pmatrix} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$$

Then

$$\begin{aligned} \widehat{g}_{11} &= \frac{\partial x}{\partial r}\frac{\partial x}{\partial r} + \frac{\partial y}{\partial r}\frac{\partial y}{\partial r} = 1 \\ \widehat{g}_{12} = \widehat{g}_{21} &= \frac{\partial x}{\partial r}\frac{\partial x}{\partial \theta} + \frac{\partial y}{\partial r}\frac{\partial y}{\partial \theta} = 0 \\ \widehat{g}_{22} &= \frac{\partial x}{\partial \theta}\frac{\partial x}{\partial \theta} + \frac{\partial y}{\partial \theta}\frac{\partial y}{\partial \theta} = r^2 \end{aligned}$$

Hence the metric tensor in polar coordinates is

$$g = \widehat{g}_{ij}\mathrm{d}\widehat{x}^i \otimes \mathrm{d}\widehat{x}^j = \mathrm{d}r\otimes\mathrm{d}r + r^2\mathrm{d}\theta\otimes\mathrm{d}\theta$$

Let $G$ be the matrix whose components are $g_{ij}$. Then the reciprocal is $g^{ij} = (G^{-1})^{\mathrm{T}}_{ij}$, where

$$G = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix} \qquad (G^{-1})^{\mathrm{T}} = \begin{pmatrix} 1 & 0 \\ 0 & 1/r^2 \end{pmatrix}$$

Then the reciprocal of the metric tensor $g$ is

$$\frac{\partial}{\partial r}\otimes\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial}{\partial\theta}\otimes\frac{\partial}{\partial\theta}$$

The covariant components $\widehat{x}^i$ are obtained by contracting the tensor product of $g$ and $\widehat{x}^j\dfrac{\partial}{\partial x^j}$:

$$\widehat{x}_i = g_{ij}x^j \implies \widehat{x}_1 = g_{11}r = r,\ \widehat{x}_2 = g_{22}\theta = r^2\theta$$

Note that the Jacobian vanishes at the origin, so the coordinate transformation is not invertible. So we cannot use $\widehat{g}_{ij}$ to calculate the distance from origin to a fixed point.

Let $A$ be the rotation matrix:

$$A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \qquad A^{-1} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}$$

Since $\widetilde{x}^i = A^i_j x^j$, we have $\dfrac{\partial x^i}{\partial \widetilde{x}^j} = \left(A^{-1}\right)^i_j$. Then

$$\widetilde{g}_{ij} = \frac{\partial x^k}{\partial \widetilde{x}^i}\frac{\partial x^\ell}{\partial \widetilde{x}^j} g_{k\ell} \implies \widetilde{G} = (A^{-1})^{\mathrm{T}} G A^{-1} = A G A^{-1}$$

In Cartesian coordinates, $G = I$. So after rotation $\widetilde{G} = AIA^{-1} = I$ so the components of the metric tensor are invariant.

In polar coordinates, the rotation is given by $r \mapsto r$ and $\theta \mapsto \theta + \alpha$. So the rotation matrix (in fact the Jacobian matrix) $A = I$. So after rotation $\widetilde{G} = G$. The components of the metric tensor are invariant. □

## Some Mathematical Stuff

In the lectures, contravariant and covariant vectors are defined via coordinate transformations. Here I would like to present a more geometric way of defining vectors and tensors.

**Definition. Smooth Manifolds.** Suppose that $M$ is a *Hausdorff, second countable* topological space. $X$ is called a smooth manifold or simply a manifold, if there exists a family $\mathscr{A} = \{(U_i, \varphi_i) : i \in I\}$, where:

1. $\{U_i : i \in I\}$ is an open cover of $M$;
2. $\varphi_i : U_i \to \mathbb{R}^n$ is an homeomorphism onto its image;
3. for $U_i \cap U_j \neq \varnothing$, the **transition map** $\varphi_j \circ \varphi_i^{-1}|_{\varphi_i(U_i \cap U_j)} : \varphi_i(U_i \cap U_j) \to \varphi_j(U_i \cap U_j)$ is bijecitve, $\mathrm{C}^\infty$, and has a $\mathrm{C}^\infty$ inverse. (This is called the **compatibility** condition.)

$\mathscr{A}$ is called an **atlas** and the pairs $(U_i, \varphi_i)$ are called **coordinate charts**. $n$ is the dimension of the manifold $M$.

If the altas $\mathscr{A}$ is maximal in the sense that every coordinate chart that is compatible with one in $\mathscr{A}$ is contained in $\mathscr{A}$, then we say that $\mathscr{A}$ defines a **differentiable structure** on $M$.

Here we only concern the local properties of manifolds. Let $M$ be an $n$-dimensional manifold and $p \in M$. Let $(U, \varphi) = (U; x^1, \dots, x^n)$ be a coordinate chart with $p \in U$. Let $(r^1, \dots, r^n)$ be the standard coordinates of $\mathbb{R}^n$. Then we have $n$ real-valued functions $x^i = r^i \circ \varphi : U \to \mathbb{R}$.

Let $\mathrm{C}^\infty_p$ denotes the set of smooth functions from $U$ to $\mathbb{R}$. Let $\gamma : [-\varepsilon, \varepsilon] \to M$ be a $\mathrm{C}^1$ path on $M$ with $\gamma(0) = p$. For $f \in \mathrm{C}^\infty_p$, $f \circ \gamma : [-\varepsilon, \varepsilon] \to \mathbb{R}$ is $\mathrm{C}^1$ and we can differentiate at 0:

$$(f\circ\gamma)'(0) = \frac{\mathrm{d}}{\mathrm{d}t}(f\circ\gamma)(t)\bigg|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t}(f\circ\varphi^{-1}\circ\varphi\circ\gamma)(t)\bigg|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t}(f\circ\varphi^{-1})\left(c^1(t),\dots,c^n(t)\right)\bigg|_{t=0} = \left(c^{i\prime}(0)\,\frac{\partial}{\partial r^i}\bigg|_{\varphi(p)}\right)(f\circ\varphi^{-1})$$

where $\left(c^1(t), \dots, c^n(t)\right) = \varphi \circ \gamma(t)$.

Let $\dfrac{\partial f}{\partial x^i}\bigg|_p$ denotes $\dfrac{\partial}{\partial r^i}\bigg|_{\varphi(p)}(f \circ \varphi^{-1})$. In order words, $\dfrac{\partial}{\partial x^i}\bigg|_p$ is the pullback of $\dfrac{\partial}{\partial r^i}\bigg|_{\varphi(p)}$ by $\varphi^{-1}$.

We observe that the path $\gamma$ induces the linear functional $f \mapsto (f \circ \gamma)'(0)$, which spans a (dual) vector space. We can take $\left\{\dfrac{\partial}{\partial x^1}\bigg|_p, \dots, \dfrac{\partial}{\partial x^n}\bigg|_p\right\}$ as a basis. Hence we have identified the path $\gamma$ with a vector $(c^1, \dots, c^n)$ in $\mathbb{R}^n$.

**Definition. Tangent Vectors.** Let $\gamma : [-\varepsilon, \varepsilon] \to M$ be a $\mathrm{C}^1$ path on $M$ with $\gamma(0) = p$. Then $v_\gamma : C^\infty_p \to \mathbb{R}$, $f \mapsto (f \circ \gamma)'(0)$ is called a tangent vector of $M$ at $p$.

**Definition. Tangent Spaces.** The set of tangent vectors at $p$ is called the tangent space at $p$ and is denoted by $\mathrm{T}_pM$. $\mathrm{T}_pM$ is a $n$-dimensional vector space with a basis $\left\{\dfrac{\partial}{\partial x^1}\bigg|_p, \dots, \dfrac{\partial}{\partial x^n}\bigg|_p\right\}$.

We shall show that the tangent vectors are exactly the **contravariant vectors** defined in the lectures:

Suppose that $(U; x^1, \dots, x^n)$ and $(\widetilde{U}, \widetilde{x}^1, \dots, \widetilde{x}^n)$ are two coordinate charts and $p \in U \cap \widetilde{U}$. Then from the discussions above we know that $\mathrm{T}_pM$ has two bases $\left\{\dfrac{\partial}{\partial x^1}\bigg|_p, \dots, \dfrac{\partial}{\partial x^n}\bigg|_p\right\}$ and $\left\{\dfrac{\partial}{\partial \widetilde{x}^1}\bigg|_p, \dots, \dfrac{\partial}{\partial \widetilde{x}^n}\bigg|_p\right\}$. Consider a tangent vector

$$v = v^i \frac{\partial}{\partial x^i}\bigg|_p = \widetilde{v}^i \frac{\partial}{\partial \widetilde{x}^i}\bigg|_p \in \mathrm{T}_pM$$

For any $f \in C_p^\infty$,

$$v(f) = v^i \left.\frac{\partial f}{\partial x^i}\right|_p = v^i \left.\frac{\partial f \circ \varphi^{-1}}{\partial r^i}\right|_{\varphi(p)} = v^i \left.\frac{\partial (f \circ \widetilde{\varphi}^{-1}) \circ (\widetilde{\varphi} \circ \varphi^{-1})}{\partial r^i}\right|_{\varphi(p)} = v^i \left.\frac{\partial f \circ \widetilde{\varphi}^{-1}}{\partial r^j}\right|_{\widetilde{\varphi}(p)} \left.\frac{\partial \left(\widetilde{\varphi} \circ \varphi^{-1}\right)^j}{\partial r^i}\right|_{\varphi(p)} = v^i \left.\frac{\partial f}{\partial \widetilde{x}^j}\frac{\partial \widetilde{x}^j}{\partial x^i}\right|_p = \widetilde{v}^j \left.\frac{\partial f}{\partial \widetilde{x}^j}\right|_p$$

Hence the coordinates are transformed via

$$\widetilde{v}^j = \left.\frac{\partial \widetilde{x}^j}{\partial x^i}\right|_p v^i$$

For $f \in C_p^\infty$, $f$ naturally induces a dual vector of a tangent vector: $\mathrm{d}f\big|_p : \mathrm{T}_p M \to \mathbb{R}$, $\mathrm{d}f\big|_p(v) := v(f)$. In this way we can identify the dual space of the tangent space:

**Definition. Cotangent Spaces.** Let $\mathrm{T}_p M$ be the tangent space at $p \in M$. The cotangent space $\mathrm{T}_p^* M$ is the dual space of $\mathrm{T}_p M$. $\left\{\mathrm{d}x^1\big|_p, \ldots, \mathrm{d}x^n\big|_p\right\}$ is a basis of $\mathrm{T}_p^* M$, and is the dual basis of $\left\{ \left.\frac{\partial}{\partial x^1}\right|_p, \ldots, \left.\frac{\partial}{\partial x^n}\right|_p \right\}$. A **cotangent vector** or a **covariant vector** is an element in the cotangent space.

$$\mathrm{d}x^i\big|_p \left( \left.\frac{\partial}{\partial x^j}\right|_p \right) = \left.\frac{\partial x^i}{\partial x^j}\right|_p = \delta^i_j$$

Coordinate transformations of cotangent vectors: Consider a cotangent vector

$$\omega = \omega_i \mathrm{d}x^i\big|_p = \widetilde{\omega}_i \mathrm{d}\widetilde{x}^i\big|_p$$

For any tangent vector $v = v^i \left.\frac{\partial}{\partial x^i}\right|_p = \widetilde{v}^i \left.\frac{\partial}{\partial \widetilde{x}^i}\right|_p \in \mathrm{T}_p M$, we have

$$\omega(v) = \omega_i v^i = \widetilde{\omega}_j \widetilde{v}^j$$

Hence the coordinates are transformed via

$$\widetilde{\omega}_j = \left.\frac{\partial x^j}{\partial \widetilde{x}^i}\right|_p \omega_i$$

**Definition. Tensor Product.** Suppose that $V_1, \ldots, V_n$ are vector spaces over a field $F$. The tensor product space $V_1 \otimes \cdots \otimes V_n$ is a vector space satisfying the following universal property:

There exists a multilinear map $\varphi : V_1 \times \cdots \times V_n \to V_1 \otimes \cdots \otimes V_n$ such that for any $F$-vector space $W$ and multilinear map $\sigma : V_1 \times \cdots \times V_n \to W$ such that there exists a unique *linear map* $\widetilde{\sigma} : V_1 \otimes \cdots \otimes V_n \to W$ such that $\sigma = \widetilde{\sigma} \circ \varphi$.

**Definition. Tensors.** Let $\mathrm{T}_p M$ and $\mathrm{T}_p^* M$ be the tangent and cotangent spaces at $p \in M$. A tensor product space of type $(r, s)$ at $p$ is

$$T^r_s(p) = \underbrace{\mathrm{T}_p M \otimes \cdots \otimes \mathrm{T}_p M}_{r \text{ times}} \otimes \underbrace{\mathrm{T}_p^* M \otimes \cdots \otimes \mathrm{T}_p^* M}_{s \text{ times}}$$

A tensor of type $(r, s)$ (rank $r$ contravariant and rank $s$ covariant) is an element of $T^r_s(p)$.

Naturally a basis of $T^r_s(p)$ is the set of vectors

$$\frac{\partial}{\partial x^{i_1}} \otimes \cdots \otimes \frac{\partial}{\partial x^{i_r}} \otimes \mathrm{d}x^{j_1} \otimes \cdots \otimes \mathrm{d}x^{j_s}$$

where $i_1, \ldots, i_r, j_1, \ldots, j_s \in \{1, \ldots, n\}$.

For a tensor $T \in T^r_s(p)$:

$$T = T^{i_1 \cdots i_r}_{j_1 \cdots j_s} \frac{\partial}{\partial x^{i_1}} \otimes \cdots \otimes \frac{\partial}{\partial x^{i_r}} \otimes \mathrm{d}x^{j_1} \otimes \cdots \otimes \mathrm{d}x^{j_s} = \widetilde{T}^{i_1 \cdots i_r}_{j_1 \cdots j_s} \frac{\partial}{\partial \widetilde{x}^{i_1}} \otimes \cdots \otimes \frac{\partial}{\partial \widetilde{x}^{i_r}} \otimes \mathrm{d}\widetilde{x}^{j_1} \otimes \cdots \otimes \mathrm{d}\widetilde{x}^{j_s}$$

The coordinates are transformed via

$$\widetilde{T}^{k_1 \cdots k_r}_{\ell_1 \cdots \ell_s} = \frac{\partial \widetilde{x}^{k_1}}{\partial x^{i_1}} \cdots \frac{\partial \widetilde{x}^{k_r}}{\partial x^{i_r}} \frac{\partial x^{j_1}}{\partial \widetilde{x}^{\ell_1}} \cdots \frac{\partial x^{j_s}}{\partial \widetilde{x}^{\ell_s}} T^{i_1 \cdots i_r}_{j_1 \cdots j_s}$$