Physics 690/482-02 Spring, 2005 Josh Erlich
Notes: Geometry I

We will start by reviewing Euclidean geometry in a language suitable for generalizing to non-Euclidean geometries.

Consider a two-dimensional spatial manifold, *i.e.* no time coordinate. A point in a two-dimensional manifold is specified by two numbers (x^1, x^2). In Cartesian coordinates, a differential length element along a path is given by,

$$\begin{aligned} ds^2 &= (dx^1)^2 + (dx^2)^2 \\ &= (dx^1, dx^2) \begin{pmatrix} 1 & \\ & 1 \end{pmatrix} \begin{pmatrix} dx^1 \\ dx^2 \end{pmatrix} \\ &= \sum_{i,j=1}^{2} dx^i \, \delta_{ij} \, dx^j \\ &\equiv dx^i \, \delta_{ij} \, dx^j. \end{aligned}$$

The identity matrix δ_{ij} is called the **metric** of Euclidean space in Cartesian coordinates. In the last line of the equation above we introduced **Einstein's summation convention:** When an index $(i, j, \ldots)$ appears twice in a monomial it is automatically summed over. From now on, if you ever want to repeat an index in a monomial and *not* sum over its values, you must state your intent explicitly. (The rule does not apply to sums of expressions, *e.g.* $A^i + B^i$, but it does apply for products, *e.g.* $A_i \, B^i = \sum_i A_i \, B^i$.

Notice also that each index which was summed over appeared once in a lower position and once in an upper position. There will be a reason later to distinguish between upper and lower indices. For now we will always write a repeated index that is summed over once as a superscript and once as a subscript in that expression. The rules of index conventions are:

- An index which appears twice in a monomial is summed over, and should appear once as a superscript and once as a subscript. Such an index is called a "dummy index, and can be called anything." For example, $A_i \, B^i = A_j \, B^j = \sum_i A_i \, B^i$.
- The position (superscript or subscript) of a free index (one that is not summed

over) in an equation should be consistent throughout the equation. Hence, $A^i + B^i = C^i$ is okay as an equation, bu $A_i + B^i = C^i$ is not.

- No index should ever appear more than twice in a monomial. Hence, A^i_{ii} has no meaning.

With these rules, any equation involving products of vectors (for example x^i) and the metric g_{ij} will be consistent (covariant) if the coordinate system is changed, as we shall explain in a little bit.

The length of a path is given by,

$$s = \int ds.$$

The meaning of the right hand side is as follows: parametrize the path by two functions $f^i(\tau)$ such that $x^i = f^i(\tau)$, where τ is a parameter that varies monotonically from some value τ_1 to another value τ_2 along the path. Then,

$$\begin{aligned} ds &= d\tau \sqrt{\left(\frac{dx^1}{d\tau}\right)^2 + \left(\frac{dx^2}{d\tau}\right)^2} \\ &= d\tau \sqrt{\frac{dx^i}{d\tau}\,\delta_{ij}\,\frac{dx^j}{d\tau}} \\ &\equiv d\tau \sqrt{\dot{x}^i\,\delta_{ij}\,\dot{x}^j} \\ &\equiv d\tau \sqrt{\dot{x}^i\,\dot{x}_i}. \end{aligned}$$

Here we have defined $\dot{x}^i \equiv dx^i/d\tau$, and we have also defined $\dot{x}_i \equiv \delta_{ij}\,\dot{x}^j$. We will in general use the metric to lower indices in this way. The purpose of all this shorthand notation is simplicity, but if you are not yet comfortable with these manipulations then feel free to use more ink than is necessary.

In 3-dimensions, the line element is $ds^2 = \sum_{i=1}^3 (dx^i)^2 = dx^i\,dx_i$, and in d-dimensions it is $ds^2 = \sum_{i=1}^d (dx^i)^2 = dx^i\,dx_i$. Notice that in our shorthand notation, the repeated index is summed from 1 to d in general.

Coordinate transformations:

We will now repeat this discussion in other coordinate systems. In 2D polar coordinates,

$$x^1 = r\cos\theta$$

$$x^2 = r\sin\theta,$$

we can rewrite the line element as follows:

$$\begin{aligned}
ds^2 &= (dx^1)^2 + (dx^2)^2 \\
&= \left(\frac{\partial x^1}{\partial r}\,dr + \frac{\partial x^1}{\partial \theta}\,d\theta\right)^2 + \left(\frac{\partial x^2}{\partial r}\,dr + \frac{\partial x^2}{\partial \theta}\,d\theta\right)^2 \\
&= (\cos\theta\,dr - r\sin\theta\,d\theta)^2 + (\sin\theta\,dr + r\cos\theta\,d\theta)^2 \\
&= dr^2\,(\cos^2\theta + \sin^2\theta) + dr\,d\theta\,(-r\cos\theta\,\sin\theta + r\cos\theta\,\sin\theta) \\
&\quad +d\theta^2\,(r^2\sin^2\theta + r^2\cos^2\theta) \\
&= dr^2 + r^2\,d\theta^2 \\
&= (dr, d\theta)\begin{pmatrix} 1 & \\ & r^2 \end{pmatrix}\begin{pmatrix} dr \\ d\theta \end{pmatrix}.
\end{aligned}$$

The matrix diag(1,r^2) in the last line is the metric in polar coordinates.

We can generalize this to an arbitrary set of coordinates. Let's say that the metric in some coordinate system (x^1, x^2) is g_{ij}. Then in a different coordinate system $(\tilde{x}^1, \tilde{x}^2)$ such that $x^1 = f^1(\tilde{x}^1, \tilde{x}^2)$, $x^2 = f^2(\tilde{x}^1, \tilde{x}^2)$,

$$\begin{aligned}
ds^2 &= dx^i\,g_{ij}\,dx^j \\
&= \left(\frac{\partial x^i}{\partial \tilde{x}^m}\,d\tilde{x}^m\right)\,g_{ij}\,\left(\frac{\partial x^j}{\partial \tilde{x}^n}\,d\tilde{x}^n\right) \\
&= d\tilde{x}^m\,\left[\frac{\partial x^i}{\partial \tilde{x}^m}\,\frac{\partial x^j}{\partial \tilde{x}^n}\,g_{ij}\right]\,d\tilde{x}^n \\
&\equiv d\tilde{x}^m\,\tilde{g}_{mn}\,d\tilde{x}^n.
\end{aligned}$$

In the new coordinate system the metric has become,

$$\tilde{g}_{mn} = \frac{\partial x^i}{\partial \tilde{x}^m}\,\frac{\partial x^j}{\partial \tilde{x}^n}\,g_{ij}.$$

We can invert this expression using the chain rule, namely

$$\frac{\partial x^i}{\partial \tilde{x}^m}\frac{\partial \tilde{x}^m}{\partial x^j} = \delta^i_j,$$

where as usual the repeated index n is summed over and δ^i_j is the Kronecker delta, which takes the values 1 if $k = j$ and 0 otherwise.

Multiplying the expression for $\tilde{g}_{mn}$ above by $\frac{\partial \tilde{x}^m}{\partial x^k}\frac{\partial \tilde{x}^n}{\partial x^l}$ and summing over the repeated indices yields,

$$\begin{aligned}\frac{\partial \tilde{x}^m}{\partial x^k}\frac{\partial \tilde{x}^n}{\partial x^l}\tilde{g}_{mn} &= \frac{\partial \tilde{x}^m}{\partial x^k}\frac{\partial x^i}{\partial \tilde{x}^m}\,\frac{\partial \tilde{x}^n}{\partial x^l}\frac{\partial x^j}{\partial \tilde{x}^n}\,g_{ij} \\ &= \delta^i_k\,\delta^j_l\,g_{ij} = g_{kl}.\end{aligned}$$

In other words,

$$g_{kl} = \frac{\partial \tilde{x}^m}{\partial x^k}\frac{\partial \tilde{x}^n}{\partial x^l}\,\tilde{g}_{mn}.$$

A matrix that transforms this way under coordinate transformations is called a **rank-two covariant tensor**. In general, a rank-n covariant tensor has n indices

and transforms as,

$$\widetilde{T}_{m_1 m_2 \cdots m_n} = \frac{\partial x^{i_1}}{\partial \tilde{x}^{m_1}}\,\frac{\partial x^{i_2}}{\partial \tilde{x}^{m_2}}\cdots\frac{\partial x^{i_n}}{\partial \tilde{x}^{m_n}}\,T_{i_1 i_2 \cdots i_n},$$

or, inverting the epression as we did for the metric,

$$T_{i_1 i_2 \cdots i_n} = \frac{\partial \tilde{x}^{m_1}}{\partial x^{i_1}}\frac{\partial \tilde{x}^{m_2}}{\partial x^{i_2}}\cdots\frac{\partial \tilde{x}^{m_n}}{\partial x^{i_n}}\,\widetilde{T}_{m_1 m_2 \cdots m_n}.$$

Notice that this is a generalization of the transformation rule of the metric, and each lower index is contracted with a factor like $\frac{\partial \tilde{x}^m}{\partial x^i}\,\widetilde{T}_{\ldots m \ldots}$ upon a coordinate transformation.

On the other hand, recall that the coordinate differentials transformed as,

$$dx^i = \frac{\partial x^i}{\partial \tilde{x}^m}\,d\tilde{x}^m.$$

Notice that compared to the transformation of a vector with a lower (covariant) index, the positions of x and $\tilde{x}$ in the derivative are reversed. An object that transforms like dx^i is called a **contravariant vector**, to distinguish it from a covariant vector like $dx_i \equiv g_{ij}\,dx^j$. *This is the reason we have been distinguishing upper from lower indices – tensors with upper indices and lower indices transform differently under coordinate transformations.*

Generalizing the transformation of a contravariant vector, we can similarly define the transformation properties of a tensor with n upper indices (*i.e.* a **rank-n**

contravariant tensor),

$$T^{i_1 i_2 \cdots i_n} = \frac{\partial x^{i_1}}{\partial \tilde{x}^{m_1}} \frac{\partial x^{i_2}}{\partial \tilde{x}^{m_2}} \cdots \frac{\partial x^{i_n}}{\partial \tilde{x}^{m_n}} \widetilde{T}^{m_1 m_2 \cdots m_n}.$$

A vector is a rank-1 tensor, and a scalar is a rank-0 tensor. By contracting contravariant and covariant indices, *i.e.* by setting the upper and lower indices equal and summing over them, we get tensors that transform as though those indices were removed. For example, consider $T^k_{ij} = g_{ij} V^k$ for some contravariant vector V^k. Under a coordinate transformation,

$$T^k_{ij} = \frac{\partial x^k}{\partial \tilde{x}^l} \frac{\partial \tilde{x}^m}{\partial x^i} \frac{\partial \tilde{x}^n}{\partial x^j} \widetilde{T}^l_{mn}.$$

Now consider $T^j_{ij} = \sum_j g_{ij} V^j$. The tensor with two indices contracted transforms as,

$$\begin{aligned} T^j_{ij} &= \frac{\partial x^j}{\partial \tilde{x}^l} \frac{\partial \tilde{x}^n}{\partial x^j} \frac{\partial \tilde{x}^m}{\partial x^i} \widetilde{T}^l_{mn} \\ &= \delta^n_l \frac{\partial \tilde{x}^m}{\partial x^i} \widetilde{T}^l_{mn} \\ &= \frac{\partial \tilde{x}^m}{\partial x^i} \widetilde{T}^n_{mn}. \end{aligned}$$

This is how a covariant vector transforms, so T^j_{ij} transforms as a tensor without the indices j. It is therefore consistent with our rules for transforming indices to define $V_i \equiv g_{ij} V^j$. This is why we **lower indices with the metric**. Similarly, we can raise indices with the inverse metric: $V^i \equiv g^{ij} V_j$.

The metric or its inverse can also be used to form tensors of lower rank. For example, consider a tensor A_{ijk}, and **contract it with the metric** g^{ij}, *i.e.* form the product $g^{ij} A_{ijk}$ and sum over i and j. As in the previous example, the contracted tensor transforms a covariant vector. Contractions with the metric are called **traces**. Be careful to note that these are not ordinary matrix traces unless the metric is the identity δ_{ij}.

Note also that in our notation, $g^j_i = g_{ik} g^{kj} = \delta^j_i$, so the metric with one lower and one upper index is always the Kronecker delta, in any geometry. (This is true also in spacetime.)

Other useful coordinate systems:

Spherical coordinates:

$$\begin{aligned} x^1 &= r\,\cos\phi\,\sin\theta \\ x^2 &= r\,\sin\phi\,\sin\theta \\ x^3 &= r\,\cos\theta \end{aligned}$$

$$ds^2 = dr^2 + r^2\,d\theta^2 + r^2\,\sin^2\theta\,d\phi^2$$

Cylindrical coordinates:

$$\begin{aligned} x^1 &= r\,\cos\theta \\ x^2 &= r\,\sin\theta \\ x^3 &= z \end{aligned}$$

$$ds^2 = dz^2 + dr^2 + r^2 d\theta^2$$

Areas and Volumes

In 2D, the area of a region is obtained by integrating an **area element**, which we will also call the two-dimensional **volume element**, over that region. In Cartesian coordinates the area element is $dx^1\,dx^2$. In polar coordinates it is $r\,dr\,d\theta$.

Similarly, in 3D the volume element in Cartesian coordinates is $dx^1\,dx^2\,dx^3$; in spherical coordinates it is $r^2\,\sin\theta\,dr\,d\theta\,d\phi$; in cylindrical coordinates it is $r\,dr\,d\theta\,dz$.

Notice that each of these expressions can be written in terms of the metric as, $\sqrt{\det|g_{kl}|}\,\prod_{i=1}^{d} d\tilde{x}^i$, where $\tilde{x}^i$ are the relevant coordinates. This expression gives the volume element in any dimension d in an arbitrary coordinate system. This volume element is invariant under a change of coordinates. In other words, the volume of a region so defined does not depend on the choice of coordinates. You can see this by considering the Jacobian that appears in the transformation $d^d x \to \det|\partial x^k/\partial\tilde{x}^l|\,d^d\tilde{x}$, and comparing with the transformation of $g \equiv \det|g_{ij}| = (\det|\partial\tilde{x}^i/\partial x^j|)^2\,\det|\tilde{g}_{kl}|$.

A word of caution: The symbol g is commonly used to denote the determinant of the metric g_{ij}, not its trace (which is $g_{ij}\,g^{ij} = d$ in d-dimensions).