Lecture Seven: Theory of Relativity

7.1 Why Riemannian geometry?

A space is said to be Riemannian if the metric interval

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu \tag{1}$$

is *quadratic* in the coordinates.

It is possible to construct other spaces with metrics that are not quadratic in the coordinates, for example,

$$ds^2 = \left(dx^4 + dy^4\right)^{1/2} \tag{2}$$

Mathematicians call spaces with the metric of equation (2) 'Finsler geometries'. Why are we interested in *Riemannian* rather than any other spaces?

The answer is because we are physicists, not mathematicians. We are interested *exclusively* in spaces in which we can set up locally Minkowski coordinates with metric interval

$$ds^2 = \eta_{\mu\nu} d\xi^\mu d\xi^\nu.$$

So in a general coordinate system x^μ, the metric interval is given by equation (1) with

$$g_{\mu\nu} = \eta_{\kappa\lambda} \frac{\partial \xi^\kappa}{\partial x^\mu} \frac{\partial \xi^\lambda}{\partial x^\nu}.$$

It is the principle of equivalence that motivates us to focus exclusively on Riemannian geometry in formulating a theory of gravity.

7.2 Geometric picture of tensors

Suppose that we have a vector A^μ in coordinate system x^μ. Let's form the operator

$$A = A^\mu \frac{\partial}{\partial x^\mu}$$

and apply this operator to a function $f(x^\mu)$

$$Af = \left(A^\mu \frac{\partial}{\partial x^\mu}\right) f = A^\mu \frac{\partial f}{\partial x^\mu}.$$

What happens in another coordinate system x'^μ,

$$\begin{aligned} A'f = A'^\mu \frac{\partial f}{\partial x'^\mu} &= A^\nu \frac{\partial x'^\mu}{\partial x^\nu} \frac{\partial f}{\partial x^\kappa} \frac{\partial x^\kappa}{\partial x'^\mu} \\ &= A^\nu \frac{\partial f}{\partial x^\kappa} \delta^\kappa_\nu = A^\nu \frac{\partial f}{\partial x^\nu}. \end{aligned}$$

Hence

$$A'^{\mu}\frac{\partial}{\partial x'^{\mu}} \equiv A^{\mu}\frac{\partial}{\partial x^{\mu}}$$

and we can think of the $\partial/\partial x^{\mu}$ as forming a basis for all vectors A^{μ}.

For example, imagine the surface of a sphere labelled by spherical polar coordinates coordinates θ and ϕ

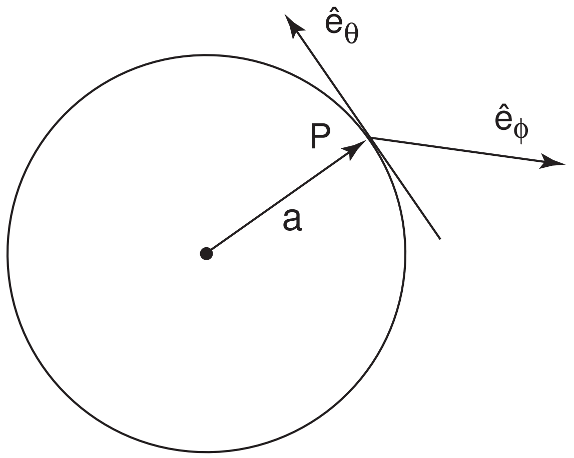

Figure 1:

The metric of the spherical surface in spherical polar coordinates is

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

and so the gradient operator is

$$\boldsymbol{\nabla} = \frac{1}{a}\frac{\partial}{\partial\theta}\hat{\mathbf{e}}_\theta + \frac{1}{a\sin\theta}\frac{\partial}{\partial\phi}\hat{\mathbf{e}}_\phi \tag{3}$$

where $\hat{\mathbf{e}}_\theta$ and $\hat{\mathbf{e}}_\phi$ are (in this example) orthogonal unit vectors as shown in the figure. The gradient operator (3) allows us to create a vector from a scalar function g

$$\boldsymbol{\nabla} g = \frac{1}{a}\frac{\partial g}{\partial\theta}\hat{\mathbf{e}}_\theta + \frac{1}{a\sin\theta}\frac{\partial g}{\partial\phi}\hat{\mathbf{e}}_\phi,$$

and so it follows that we can write any vector in terms of the basis vectors $\hat{\mathbf{e}}_\theta$ and $\hat{\mathbf{e}}_\phi$

$$\mathbf{A} = a_1\hat{\mathbf{e}}_\theta + a_2\hat{\mathbf{e}}_\phi.$$

The derivatives $\partial\,/\partial x^{\nu}$ thus define a basis for an arbitrary vector in the *tangent space* at any given point P as in the following figure:

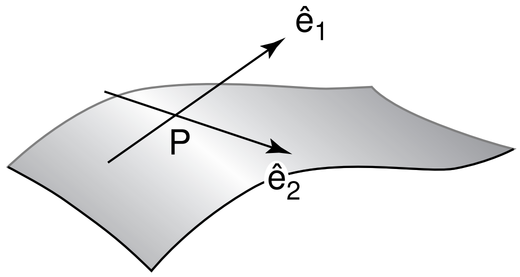

Figure 2:

A contravariant vector lives in the tangent space defined by $\hat{\mathbf{e}}_1$ and $\hat{\mathbf{e}}_2$, *not* in the manifold of real space. If space is flat, however, the tangent space coincides with the real space as in the example shown below:

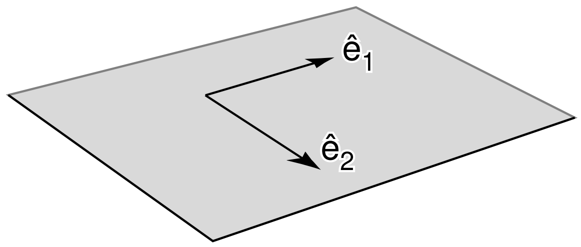

Figure 3:

Covariant vectors transform like the basis vectors $\partial\ /\partial x^{\mu}$, which are tangent vectors in the real space. These examples provide a good way to think about contravariant and covariant vectors. Covariant vectors transform like the basis vectors defined by the coordinates. Contravariant vectors live in the tangent space defined by these basis vectors.

7.3 Geodesics

Suppose that we have two points A and B joined by a path defined by a parameter p that varies along the curve

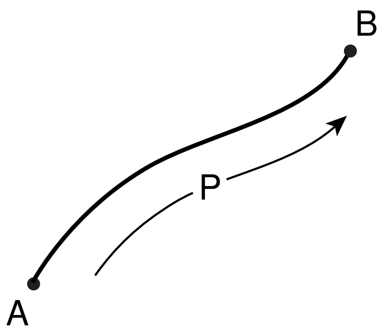

Figure 4:

The length of the path is

$$\mathrm{L}_{AB} = \int_A^B ds = \int_A^B \frac{ds}{dp}\, dp.$$

The metric interval is

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu,$$

hence

$$\mathrm{L}_{AB} = \int_A^B \left| g_{\mu\nu} \frac{dx^\mu}{dp} \frac{dx^\nu}{dp} \right|^{1/2} dp.$$

Now, vary the path $x^\mu(p) \to x^\mu(p) + \delta x^\mu(p)$ keeping the endpoints A and B fixed

$$\delta \mathrm{L}_{AB} \;=\; \int_A^B \left| g_{\mu\nu} \frac{dx^\mu}{dp} \frac{dx^\nu}{dp} \right|^{-1/2} \left\{ \frac{\partial g_{\mu\nu}}{\partial x^\lambda} \delta x^\lambda \frac{dx^\mu}{dp} \frac{dx^\nu}{dp} + g_{\mu\nu} \frac{d\delta x^\mu}{dp} \frac{dx^\nu}{dp} + g_{\mu\nu} \frac{dx^\mu}{dp} \frac{d\delta x^\nu}{dp} \right\} dp.$$

But

$$\left| g_{\mu\nu} \frac{dx^\mu}{dp} \frac{dx^\nu}{dp} \right|^{1/2} = \frac{ds}{dp}$$

so substituting in the equation for $\delta \mathrm{L}_{AB}$

$$\begin{aligned} \delta \mathrm{L}_{AB} &= \int_A^B \frac{dp}{ds} \left\{ \frac{\partial g_{\mu\nu}}{dx^\lambda} \delta x^\lambda \frac{dx^\mu}{dp} \frac{dx^\nu}{dp} + 2 g_{\mu\nu} \frac{dx^\mu}{dp} \frac{d\delta x^\nu}{dp} \right\} dp \\ &= \int_A^B \left\{ \frac{\partial g_{\mu\nu}}{dx^\lambda} \delta x^\lambda \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} + 2 g_{\mu\nu} \frac{dx^\mu}{ds} \frac{d\delta x^\nu}{ds} \right\} ds \end{aligned}$$

integrate this term by parts

$$\int_A^B 2 g_{\mu\nu} \frac{dx^\mu}{ds} \frac{d\delta x^\nu}{ds}\, ds \;=\; 2 g_{\mu\nu} \frac{dx^\mu}{ds} \delta x^\nu \Big|_A^B - \int \left[2 \frac{\partial g_{\mu\nu}}{\partial x^\kappa} \frac{dx^\kappa}{ds} \frac{dx^\mu}{ds} \delta x^\nu + 2 g_{\mu\nu} \frac{d^2 x^\mu}{ds^2} \delta x^\nu \right].$$

The first term on the rhs of this equation is zero because the endpoints of the path are constrained to be fixed ($\delta x^\nu = 0$), hence

$$\delta \mathrm{L}_{AB} = \int_A^B \left[\frac{\partial g_{\mu\nu}}{\partial x^\lambda} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} - 2 g_{\mu\lambda} \frac{d^2 x^\mu}{ds^2} - 2 \frac{\partial g_{\mu\lambda}}{\partial x^\kappa} \frac{dx^\kappa}{ds} \frac{dx^\mu}{ds} \right] \delta x^\lambda ds. \tag{4}$$

Recall the definition of the affine connection, which I will write in the form,

$$2 g_{\rho\nu} \Gamma^\rho_{\mu\lambda} = \left[\frac{\partial g_{\mu\nu}}{\partial x^\lambda} + \frac{\partial g_{\lambda\nu}}{\partial x^\mu} - \frac{\partial g_{\mu\lambda}}{\partial x^\nu} \right]$$

then after some tedious manipulation of indices we can rewrite equation (4) as

$$\delta \mathcal{L}_{AB} = -2 \int_A^B \left[\frac{d^2 x^\rho}{ds^2} + \Gamma^\rho_{\nu\mu} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} \right] g_{\rho\nu} \delta x^\lambda\, ds,$$

and for this to be stationary for any arbitrary δx^λ,

$$\frac{d^2 x^\rho}{ds^2} + \Gamma^\rho_{\nu\mu} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} = 0.$$

We have recovered the equations of motion of a freely falling particle that we derived in Section 6.1, with s as an 'affine' parameter. (An affine parameter in differential geometry is a special parameter that characterises a curve in such a way that a tangent vector is propagated onto itself as the parameter is varied. If you don't understand this definition, then you can regard an affine parameter as any parameter related to s by $s' = \alpha s + \beta$, where α and β are constants.)

How is the derivation of the equations of motion in this section related to the derivation of Section 6? There we derived the equations of motion (or *geodesic equations*) by assuming that we could construct a locally Euclidean coordinate system ξ^μ in which a particle moves in a straight line. What we have now proved is that motion along these paths minimises the coordinate distance between two points on the manifold. These paths are called *geodesics*.

The derivation of Section 6 therefore has the following geometrical interpretation. In a Riemanian manifold we can construct locally Euclidean coordinates in which the $\Gamma^\rho_{\kappa\sigma}$'s vanish. Geodesics connect these Euclidean frames along a path in such a way that the coordinate distance between two points is minimised. (The Γ's therefore allow you to "connect" an affine parameter along a geodesic path in an arbitrary coordinate system – this is why the Γ's are called "affine connections").

7.4 Relation to the equations of motion in classical mechanics

My guess is that you have met much of this before in courses on classical mechanics, though perhaps in a different mathematical language.

Define a set of generalised coordinates q^i. These coordinates define a space with a metric

$$ds^2 = g_{ij} dq^i dq^j.$$

In classical mechanics this space is called the *configuration space* of the system.

We can form a Lagrangian from the kinetic and potential energies,

$$\begin{aligned} T &= \frac{1}{2} g_{ij} \dot{q}^i \dot{q}^j, \\ V &= V(q^i) \qquad \left(\text{Forces,} \quad F_i = \frac{\partial V}{\partial q^i} \right). \end{aligned}$$

The Lagrangian

$$\mathcal{L} = T - V$$

of the system obeys Lagrange's equations of motion,

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}^i}\right) - \frac{\partial \mathcal{L}}{\partial q^i} = 0,$$

which we can derive from an action principle.

Define the action S as

$$S = \int_{t_1}^{t_2} \mathcal{L} dt$$

and vary the action along a path keeping the endpoints fixed,

$$\delta S = \int_{t_1}^{t_2} \delta \mathcal{L} dt = \int_{t_1}^{t_2} \left[\frac{\partial \mathcal{L}}{\partial q^i} \delta q^i + \frac{\partial \mathcal{L}}{\partial \dot{q}^i} \delta \dot{q}^i \right] dt.$$

Integrating the last term by parts

$$\delta S = \int_{t_1}^{t_2} \delta \mathcal{L} dt = \int_{t_1}^{t_2} \left[\frac{\partial \mathcal{L}}{\partial q^i} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}^i} \right) \right] \delta q^i dt = 0,$$

and this must vanish for all δq^i. So, we derive the usual Lagrangian equation of motion,

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}^i} \right) - \frac{\partial \mathcal{L}}{\partial q^i} = 0.$$

All of this should be familiar to you. Less familiar, perhaps, is how the equations of motion look if we write them out in full:

$$\ddot{q}^i + \Gamma^i_{jk} \dot{q}^i \dot{q}^k = \mathrm{F}^i.$$

These are just our geodesic equations of motion with a force term on the right hand side. In this case, the Γ^i_{jk} are the metric connections of the *configuration space*. If the forces vanish, then Lagrange's equations say that 'freely falling' particles move along geodesics in the configuration space.

As an aside, in GR the Lagrangian of a free particle is

$$\mathcal{L} = m_0 (g_{\mu\nu} \dot{x}^\mu \dot{x}^\nu)^{1/2}, \qquad \dot{x}^\mu = \frac{dx^\mu}{d\tau}.$$

I leave it as an exercise for you to figure out how this reduces to the classical Lagrangian.