Massachusetts Institute of Technology
Department of Physics
8.962 Spring 2024

Lecture 7
Introduction to the principle of equivalence; transport on a curved manifold

7.1 Freely falling frames

A key foundational concept for us in special relativity was the notion of an inertial reference frame, a frame in which objects are unaccelerated if no forces act on them. We appear to lose this idea once we consider gravity: *everything* accelerates due to the gravitational force. However, focusing for the moment on the non-relativistic motion of ordinary matter, the gravitational force $\mathbf{F}_g$ that everything feels is proportional to its mass. As such, the *acceleration* imparted by gravity is universal: all objects in the same "gravitational field" experience the same acceleration. This means that *in the absence of non-gravitational forces, objects maintain their* **relative** *velocities.*

These concepts allow us to define a reference frame that captures the essence of the IRF: the *freely falling frame*, or FFF, is the frame that "falls" under gravity at the same rate as all objects within it. If we define all of our bodies' kinematics relative to the acceleration of the FFF, then the effect of gravity is essentially canceled out. Since all objects maintain their relative velocities, the FFF ends up playing a role when gravity is important much like the role that the IRF played in special relativity.

A corollary of these ideas is a rule known as the *Weak Equivalence Principle*, or WEP: *Over sufficiently small regions, the motion of freely falling bodies due to gravity cannot be distinguished from uniform acceleration.* The restriction to "sufficiently small regions" is because in any real situation, gravity will change over a large region. These changes are known as *tides*, and they set the size of the small region over which the WEP is effective. The WEP tells us that the freely falling frames are in fact exactly the local Lorentz frames that we worked out in the previous lecture.

Is this enough for us? Can we essentially promote the IRFs of special relativity to FFFs and develop a description of gravity that works with relativity? Because of tides, the answer is no. FFFs are not the same at all points, and gravity is non uniform. In special relativity, we could define an IRF and use it to describe all of spacetime; a FFF will only cover some finite portion of the spacetime manifold. The acceleration of a FFF near the surface of the Earth is very different from the acceleration of a FFF in geostationary orbit.

We can cancel out the local acceleration by changing frames, but we cannot transform away tides; that is a lesson of our derivation of local Lorentz frames, which showed that we can only put our spacetime into a special-relativity-like representation over a finite region. Tides cause trajectories which start out parallel to one another in spacetime to become non-parallel. This means that motion in spacetime violates one of Euclid's postulates. It's worth recalling that, in retrospect, we learned that this postulate of Euclid (the "parallelism" axiom) only applies when our manifold is "flat" (in a manner to be made more precise later). Tides mean that the geometry of spacetime must be curved.

Suppose we go into a FFF, within which the spacetime metric has the form $g_{\bar\mu\bar\nu} = \eta_{\bar\mu\bar\nu} + \mathcal{O}[(\delta x)^2(\partial^2 g)]$. In that frame, we can define another version of the equivalence principle: *In a sufficiently small region of spacetime, we can find a representation such that the laws of physics reduce to those of special relativity.* This formulation is knows as the *Einstein Equivalence Principle*, or EEP. It serves to guide us to understand how to formulate the laws of physics to situations in which gravity is important[1].

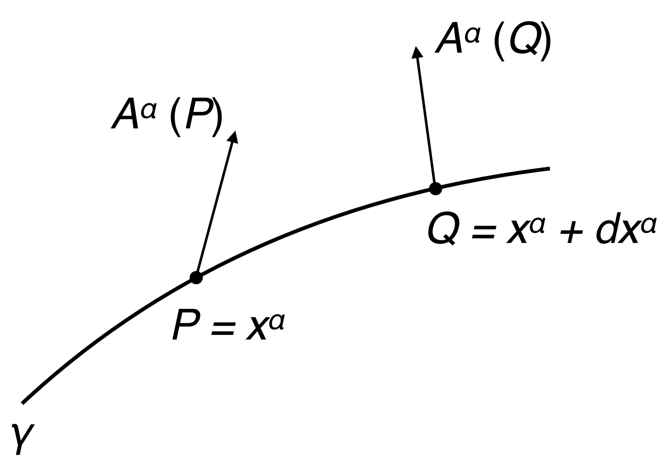
[1]It's worth noting that one other equivalence principle is often discussed: the Strong Equivalence Principle, or

7.2 Physics in a curved manifold: Parallel transport

A manifold with curvature is one on which initially parallel trajectories do not remain parallel. An example is the surface of a sphere: two north-bound trajectories at the equator will cross at the pole — the sphere's curvature caused these initial parallel paths to cross (think of lines of longitude on a globe, which all intersection the equator at 90°, but cross each other at the poles). *Not* an example is the surface of a cylinder: initial parallel paths on a surface remain parallel forever. A cylinder's surface is in fact "intrinsically" flat, though it appears to have *extrinsic* curvature due to the way we have embedded that 2-dimensional surface in 3-dimensional space.

As we think about doing physics in curved manifolds, we need to develop tools for putting things like vectors and tensors into the manifold. A useful concept for us will be the notion of a *tangent space*: the space that, at a particular point in a manifold, contains all the trajectories which pass tangent to the manifold at this point. Useful intuition is to think of the 2-dimensional surface of a sphere; the tangent space at each point on that surface is the plane which just touches the sphere at that point. The basis vectors lie in the tangent space to the manifold, and herein lies a big complication associated with doing physics in a curved manifold: *different points in the manifold have different tangent spaces.* This makes it difficult to compare vectors at different points — because the tangent spaces at event $\mathcal{P}$ and event $\mathcal{Q}$ are not the same, we must think very carefully about to relate these two vectors to each other.

In particular, how do we take derivatives of vector fields? Suppose that some curve γ lies in our manifold; the tangent vector to this curve is $\vec{u}$. We want to differentiate a vector field $\vec{A}$ that is defined in this manifold:



Our first guess for the derivative of $\vec{A}$ might be to just do the usual calculation thing:

$$\frac{\partial A^\alpha}{\partial x^\beta} = \frac{A^\alpha(\mathcal{Q}) - A^\alpha(\mathcal{P})}{dx^\beta} \,. \tag{7.1}$$

The problem with this is that the events $\mathcal{P}$ and $\mathcal{Q}$ don't have the same tangent space, so this derivative misses how the basis vectors vary from $\mathcal{P}$ to $\mathcal{Q}$. The resulting derivative is non-tensorial: though we hope to find

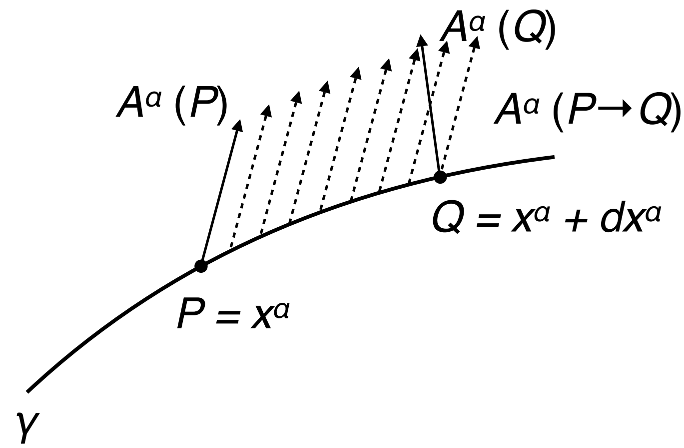
$$\partial_{\bar\nu} A^{\bar\mu} \stackrel{?}{=} \left(\frac{\partial x^\alpha}{\partial x^{\bar\nu}}\right)\left(\frac{\partial x^{\bar\mu}}{\partial x^\beta}\right)\partial_\alpha A^\beta \,, \tag{7.2}$$

we instead find

$$\partial_{\bar\nu} A^{\bar\mu} = \left(\frac{\partial x^\alpha}{\partial x^{\bar\nu}}\right)\left(\frac{\partial x^{\bar\mu}}{\partial x^\beta}\partial_\alpha A^\beta + \frac{\partial^2 x^{\bar\mu}}{\partial x^\beta \partial x^{\bar\nu}} A^\beta\right) \,. \tag{7.3}$$

SEP. We need to develop a lot more before we can use this one, which relates to the manner in which gravity itself gravitates.

To address this, we need to introduce a notion of *transporting* the field $\vec{A}$ from $\mathcal{P}$ to $\mathcal{Q}$. We thus compare the vector components at the same point, so that the same tangent space applies to them. The first notion of transport we will use is called "parallel transport." It amounts to just sliding the vector along the curve γ in sufficiently small steps that the change in the basis objects is negligible:



(Note that we have implicitly used this notion of transport throughout our physics careers to date. In a non-curved manifold, all points share the same tangent space, and parallel transport is trivial.) The transport will be "parallel" when the change associated with this notion of transport is zero.

To make this precise and quantitative, we need to define a mechanism to transport the vector between the two points. Let us assume a mathematical structure exists which transports as follows:

$$A^\alpha_{\rm PT}(\mathcal{P} \to \mathcal{Q}) = A^\alpha(\mathcal{P}) - \Pi^\alpha{}_{\beta\mu} dx^\beta A^\mu \ . \tag{7.4}$$

We then define our derivative operator D_β by comparing the vector field at $\mathcal{Q}$ to the vector field transported from $\mathcal{P}$ to $\mathcal{Q}$:

$$\begin{aligned} D_\beta A^\alpha &= \frac{A^\alpha(\mathcal{Q}) - A^\alpha_{\rm PT}(\mathcal{P} \to \mathcal{Q})}{dx^\beta} \\ &= \partial_\beta A^\alpha + \Pi^\alpha{}_{\beta\mu} A^\mu \ . \end{aligned} \tag{7.5}$$

The object we have called $\Pi^\alpha{}_{\beta\mu}$ is known as the "connection" — it connects separated points in the manifold, and allows us to compare them. We can further specify the properties of the connection by demanding it satisfy some additional properties. We will require that

$$\Pi^{\bar\alpha}{}_{\bar\beta\bar\mu} A^{\bar\mu} = \left(\frac{\partial x^{\bar\alpha}}{\partial x^\alpha}\right)\left(\frac{\partial x^\beta}{\partial x^{\bar\beta}}\right)\Pi^\alpha{}_{\beta\mu} A^\mu - \left(\frac{\partial^2 x^{\bar\alpha}}{\partial x^\beta \partial x^\mu}\right)\left(\frac{\partial x^\beta}{\partial x^{\bar\beta}}\right) A^\mu \ ; \tag{7.6}$$

this guarantees that $D_\beta A^\alpha$ transforms between representations as tensor components should, with the term involving the second derivative of the coordinate transformation matrix cancelling the analogous term that prevents the partial derivative from being tensorial.

We further require that $D_\gamma g_{\alpha\beta} = 0$ due to the equivalence principle: the metric has no derivative in a local Lorentz frame, so it cannot have a derivative in any frame. These properties tell us that the connection coefficient is in fact just the Christoffel symbol, and the derivative which follows from this transport operation is the covariant derivative:

$$\Pi^\alpha{}_{\beta\mu} = \Gamma^\alpha{}_{\beta\mu} \ , \qquad D_\beta = \nabla_\beta \ . \tag{7.7}$$

Suppose we transport a vector field along a curve parameterized by a quantity λ which increases as one moves along it[2]. The tangent vector along this curve has components $u^\alpha = dx^\alpha/d\lambda$. The vector field is parallel transported if

$$u^\beta \nabla_\beta A^\alpha \equiv \frac{DA^\alpha}{d\lambda} = 0 \,. \tag{7.8}$$

7.3 Parallel transport in the local Lorentz frame

The notion derivative operation $DA^\alpha/d\lambda$ describes how the vector field components A^α change as one moves along the curve γ. Even in the case of parallel transport,

$$\left.\frac{DA^\alpha}{d\lambda}\right|_{\rm PT} = 0 \,, \tag{7.9}$$

it should be emphasized that the vector components *do* change, since

$$u^\beta \nabla_\beta A^\alpha = u^\beta \partial_\beta A^\alpha + \Gamma^\alpha{}_{\beta\mu} u^\beta A^\mu \,. \tag{7.10}$$

Parallel transport can thus be regarded as defining an evolution for the components A^α that obeys the differential equation

$$\left.\frac{dA^\alpha}{d\lambda}\right|_{\rm PT} = -\Gamma^\alpha{}_{\beta\mu} u^\beta A^\mu \,. \tag{7.11}$$

Because within a local Lorentz frame first derivatives of the metric are zero, the Christoffel symbols are zero in that frame. This means that

$$\left.\frac{dA^\alpha}{d\lambda}\right|_{\rm PT} = 0 \qquad \text{within the LLF} \,. \tag{7.12}$$

Parallel transport thus corresponds to the components of a vector not changing within the LLF as we slide along the curve γ — hopefully, a very intuitively sensible correspondence.

7.4 Lie transport

Parallel transport is just one way to define transport of a vector or tensor along the curve γ. A second notion of transport comes by defining the event $\mathcal{P}$ as at coordinate x^α, the event $\mathcal{Q}$ as at coordinate $x^\alpha + dx^\alpha \equiv (x')^\alpha$. Since the curve γ has tangent vector $u^\alpha = d^\alpha/d\lambda$, we can further write $(x')^\alpha = x^\alpha + u^\alpha d\lambda$.

We now regard the shift from $\mathcal{P}$ to $\mathcal{Q}$ as a kind of coordinate transformation; let's apply this to a vector field $\vec{A}$:

$$\begin{aligned} A^\alpha_{\rm LT}(\mathcal{P}\to\mathcal{Q}) &= \text{components } A^\alpha \text{ ``transformed'' from coordinate } x^\alpha \text{ to } (x')^\alpha \\ &= \frac{\partial (x')^\alpha}{\partial x^\beta} A^\beta(\mathcal{P}) \\ &= \left(\delta^\alpha{}_\beta + (\partial_\beta u^\alpha) d\lambda\right) A^\beta(\mathcal{P}) \,, \end{aligned} \tag{7.13}$$

yielding

$$A^\alpha_{\rm LT}(\mathcal{P}\to\mathcal{Q}) = A^\alpha(\mathcal{P}) + (\partial_\beta u^\alpha) A^\beta(\mathcal{P}) d\lambda \,. \tag{7.14}$$

If the displacement $d\lambda$ along γ is small, we can also expand the vector field at event $\mathcal{Q}$ in a Taylor series:

$$\begin{aligned} A^\alpha(\mathcal{Q}) &= A^\alpha(x^\beta + dx^\beta) \\ &= A^\alpha(x^\beta) + dx^\beta(\partial_\beta A^\alpha) \\ &= A^\alpha(\mathcal{P}) + (u^\beta\, d\lambda)(\partial_\beta A^\alpha) \,. \end{aligned} \tag{7.15}$$

[2] We will discuss and more precisely define a particular λ parameter for certain very important curves in spacetime in an upcoming lecture.

We now define a derivative along the curve γ by comparing $A^\alpha(\mathcal{Q})$ with $A^\alpha_{\rm LT}(\mathcal{P} \to \mathcal{Q})$:

$$\begin{aligned} \mathcal{L}_{\vec u} A^\alpha &\equiv \frac{A^\alpha(\mathcal{Q}) - A^\alpha_{\rm LT}(\mathcal{P} \to \mathcal{Q})}{d\lambda} \\ &= u^\beta(\partial_\beta A^\alpha) - A^\beta(\partial_\beta u^\alpha) \; . \end{aligned} \tag{7.16}$$

This construction is known as the *Lie derivative*, and tells us how the field $\vec{A}$ varies as one "flows" along the vector field $\vec{u}$. Lie derivatives are often encountered in fluid dynamics, since they express a very natural notion of how a quantity evolves as one moves along the flow lines associated with the motion of a fluid.

It's not hard to show that

$$\mathcal{L}_{\vec u} A^\alpha = u^\beta \nabla_\beta A^\alpha - A^\beta \nabla_\beta u^\alpha \; . \tag{7.17}$$

The connection coefficients which enter into the covariant derivative cancel out, reducing the expression in terms of covariant derivatives to one that only involves partial derivatives. However, the fact that it can be written using covariant derivatives tells us that the Lie derivative is itself tensorial: the components of the Lie derivative of some tensorial object transform between frames as a tensor.

If we repeat this exercise but apply the Lie derivative to a scalar, we find

$$\mathcal{L}_{\vec u} \Phi = u^\alpha \partial_\alpha \Phi = u^\alpha \nabla_\alpha \Phi \; . \tag{7.18}$$

For a 1-form, we find

$$\begin{aligned} \mathcal{L}_{\vec u} p_\alpha &= u^\beta \partial_\beta p_\alpha + p_\beta \partial_\alpha u^\beta \\ &= u^\beta \nabla_\beta p_\alpha + p_\beta \nabla_\alpha u^\beta \; . \end{aligned} \tag{7.19}$$

The generalization to general rank tensors is probably not surprising at this point:

$$\begin{aligned} \mathcal{L}_{\vec u} T^\alpha{}_\beta &= u^\mu \partial_\mu T^\alpha{}_\beta - T^\mu{}_\beta \partial_\mu u^\alpha + T^\alpha{}_\mu \partial_\beta u^\mu \\ &= u^\mu \nabla_\mu T^\alpha{}_\beta - T^\mu{}_\beta \nabla_\mu u^\alpha + T^\alpha{}_\mu \nabla_\beta u^\mu \; . \end{aligned} \tag{7.20}$$

In 8.962, we use the Lie derivative for one particularly important task: characterizing important symmetry and conservation principles associated with motion along certain trajectories in spacetime.