Notes on MIT OpenCourseWare General Relativity

AC Williams

May 1, 2024

Video 4: Volumes and Volume Elements; Conservation Laws

We begin with apriori knowledge of the existence of a 4-velocity, which will be defined for a general observer to be

$$\underline{\mathbf{u}} = (\gamma,\ \gamma\vec{v}) \tag{1}$$

and

$$\underline{\mathbf{u}} \mathrel{\mathring{=}} (1,\ 0) \tag{2}$$

as measured by the "stationary" observer where any vector $\underline{\mathbf{A}}$ will represent a 4-vector, $\vec{A}$ will be a standard 3-vector and natural units ($\hbar = c = 1$) will be used unless explicity written otherwise. Similarly, the 4-momentum is defined as

$$\underline{\mathbf{p}} = m\underline{\mathbf{u}} = (\gamma m,\ \vec{p}) \tag{3}$$

These 4-vectors have invariant quantities

$$\begin{aligned} \underline{\mathbf{u}}\underline{\mathbf{u}} &= \underline{\mathbf{u}}\cdot\underline{\mathbf{u}} = u^\alpha u_\alpha = -1 \\ \underline{\mathbf{p}}\underline{\mathbf{p}} &= \underline{\mathbf{p}}\cdot\underline{\mathbf{p}} = p^\alpha p_\alpha = -m^2 \end{aligned} \tag{4}$$

For massless particles, it is convient to re-write the 4-momentum in terms of its 'angular' frequency ω and the unit vector in the direction it is traveling.

$$\underline{\mathbf{p}} = \omega(1,\ \hat{p}) \tag{5}$$

Building off of this, we can think about a collection of non-interacting particles. Lets say there is a "dust" cloud in the vacuum of space, where the collection of dust particles are not interacting with one another. The dust will have an energy density, but no collisions or forces between them. Imagine a small Gaussian cube that will represent the physical space the dust can move in and out of. In the rest frame of this Gaussian cube, we can ask how many particles are found within this cubic volumne of space, a.k.a. the number density. We can define n_0 to be the number density of the dust within some cubic volume of space in the rest frame of this space.

$$\begin{aligned} n_0 &\equiv \textit{number density in rest frame} \\ &\mathrel{\mathring{=}} \frac{N}{V_0} \end{aligned} \tag{6}$$

Generally speaking, we will want to define the dust in any moving frame. When we move out of the rest frame, we will boost into a moving frame with speed $|\vec{v}|$ relative to the dust cloud. Boosting into another frame will not change the number of particles that make up the dust cloud, but there will be a Lorentz contraction of the Gaussian cube, which means the volume will change under Lorentz contraction. The Lorentz contraction along the direction of motion ($x = x_0/\gamma$) will result in an additional factor of γ to the number density as measured in the rest frame.

$$\begin{aligned} n &\equiv \textit{number density in moving frame} \\ &= \frac{\gamma N}{V_0} = \gamma n_0 \end{aligned} \tag{7}$$

Also, there is now a "flow" of the dust moving through this volume of space when observing in the boosted frame. The cloud will be moving through space with velocity $\vec{v}$ in the boosted frame. If we define the flux of the number density to be $\vec{n}$, then we can intuitively relate the flux to the velocity.

$$\begin{aligned}\vec{n} &\equiv n\vec{v} \\ &= \gamma n_0 \vec{v}\end{aligned} \tag{8}$$

With both the number density and the flux of the number density we can attempt to construct a 4-vector from these two components, knowing that the components transform by a factor of γ in boosted frames. Let's call this 4-vector $\underline{\mathbf{N}}$.

$$\begin{aligned}\underline{\mathbf{N}} &= (n,\, \vec{n}) \\ &= (\gamma n_0,\, \gamma n_0 \vec{v}) \\ &= n_0\,(\gamma,\, \gamma\vec{v}) \\ &= n_0 \underline{\mathbf{u}}\end{aligned} \tag{9}$$

Being composed of the 4-velocity, $\underline{\mathbf{N}}$ can be called a 4-vector, which describes an invariant quantity.

$$\begin{aligned}\underline{\mathbf{N}}\underline{\mathbf{N}} &= -n_0^2 \\ n_0 &= \sqrt{-\underline{\mathbf{N}}\underline{\mathbf{N}}}\end{aligned} \tag{10}$$

BTW, this shows that the unit vector of the 4-number density, $\widehat{\underline{\mathbf{N}}}$, is equal to the 4-velocity, which is already a unit 4-vector ($\hat{\underline{\mathbf{u}}} = \underline{\mathbf{u}}$). This shows that our logic is consistent so far. Back to the topic at hand, suppose that we wanted to describe the flux of $\underline{\mathbf{N}}$ in the direction dx^α (or through the region $dx^\mu \bigwedge dx^\nu$, $\mu \neq \nu \neq \alpha$). We would simply need to contract the 4-number density with the "dual-vector" of dx^α. This dual-vector should just be the delta function if good coordinates are chosen ($dx^\mu \bigwedge dx^\nu \equiv (\tilde{dx}^\alpha)_\beta = \delta^\alpha_\beta$, $\mu \neq \nu \neq \alpha$).

$$\delta^\alpha_\beta N^\beta = N^\alpha \tag{11}$$

This is just a mathematically tidious way to say "pick the component of $\underline{\mathbf{N}}$ that is in the direction of interest". More generally, we are taking the projection of $\underline{\mathbf{N}}$ in the unit direction of the direction of interest.

$$\hat{x}_\mu N^\mu \equiv \textit{Flux of density of particles in direction } \hat{\underline{\mathbf{x}}} \tag{12}$$

The flux of $\underline{\mathbf{N}}$ through time is just its time component

$$d\hat{t} \cdot \underline{\mathbf{N}} = N^0 = n \tag{13}$$

which is the number density measured in the boosted frame.

An intuitive conservation law is to compare the rate of change of the dust in the cubic volumne of space to the flux of the dust through that space.

$$\frac{\partial n}{\partial t} = -\vec{\nabla} \cdot \vec{n} \tag{14}$$

I can also write this in integral form using the divergence theorem.

$$\begin{gathered}\int \vec{\nabla} \cdot \vec{n}\, d^3x = \oint \vec{n} \cdot d^2\vec{x} \\ \downarrow \\ \int \frac{\partial n}{\partial t}\, d^3x = -\oint \vec{n} \cdot d^2\vec{x}\end{gathered} \tag{15}$$

This is true for any inertial observer who has setup there coordinate basis in a specific way (which will be demonstrated soon following a generalization of the divergence theorem in n dimensions). Given our definition of the 4-number density, this equation can be re-written to have an invariant quantity of zero.

$$\partial_\alpha N^\alpha \;=\; 0 \tag{16}$$

It would be useful to generalize the integral forms of the conservation law(s) to n dimensions rather than just three. To start, lets focus on how we can mathematically express a volume in three dimensions. If we have a set of three vectors, we can express the volume of a parallelopiped constructed by these vectors by wedging all three vectors together. Given vectors $\vec{A}$, $\vec{B}$ and $\vec{C}$, the volume V can be defined as

$$V \;=\; \Im\left\{\vec{A}\wedge\vec{B}\wedge\vec{C}\right\} \;=\; \varepsilon_{ijk}A^iB^jC^k \tag{17}$$

where the Levi-Civita symbol is being defined as the components of a $\begin{pmatrix}0\\3\end{pmatrix}$ tensor. BTW, $\Im\{z\}$ is the imaginary component of z.

$$\begin{aligned}\varepsilon_{ijk} \;&=\; \Im\left\{\hat{e}_i\hat{e}_j\hat{e}_k\right\} \;=\; \Im\left\{\hat{e}_i\wedge\hat{e}_j\wedge\hat{e}_k\right\}\\ \varepsilon_{ijk} \;&=\; \begin{cases}1, & i\neq j\neq k \textit{ and any even permutation of elements}\\ -1, & i\neq j\neq k \textit{ and any odd permutation of elements}\\ 0, & \textit{otherwise}\end{cases}\end{aligned} \tag{18}$$

We will call this tensor $\underline{\underline{\varepsilon}}$. The volume in three dimensional space, which will now be denoted as V_3, can be generally defined as a tensor function that takes in up to three vectors as inputs.

$$V_3 \;=\; \underline{\underline{\varepsilon}}\left(\vec{A},\,\vec{B},\,\vec{C}\right) \;=\; \varepsilon_{ijk}A^iB^jC^k \tag{19}$$

Notice that if one of the vectors is left out of the entry, the Levi-Civita symbol remains, which means we construct new things depending on the number of vectors that are being put into this tensor function. For example, if $\vec{A}$ were not included, we would describe an object that is similar to a vector, but with "downstairs" indices. This is the 1-form of $\vec{A}$, whose indices will be labeled as Σ_i

$$\begin{aligned}\vec{\Sigma} \;&=\; \underline{\underline{\varepsilon}}\left(-,\,\vec{B},\,\vec{C}\right)\\ \Sigma_i \;&=\; \varepsilon_{ijk}B^jC^k\end{aligned} \tag{20}$$

where this sigma is geometrically representing a "surface" that is normal to $\vec{A}$. This can be understood better by remembering that this tensor function $\underline{\underline{\varepsilon}}$ is wedging the vectors (and ignoring the imaginary components that are a consequence of the Clifford Algebra involved with wedging orthonormal vectors). Returning to the divergence theorem,

$$\int \vec{\nabla}\cdot\vec{v}\,dV_3 \;=\; \oint \vec{v}\cdot d\vec{\Sigma} \tag{21}$$

where $\vec{v}$ is a general vector field, we can define the differential volume and surface elements.

$$dV_3 \;=\; d^3\vec{x} \;=\; \varepsilon_{ijk}dx^idx^jdx^k \tag{22}$$

$$d\vec{\Sigma} \;=\; d^2\vec{x} \;=\; \varepsilon_{ijk}dx^jdx^k \tag{23}$$

We now have a general form of the differentials that allows us to add or subtract as many dimensions as we please. Now, we can proceed to talking about 4 dimensions.

Imagine a 4 dimensional parallelepiped made from four vectors $\vec{A}$, $\vec{B}$, $\vec{C}$ and $\vec{D}$. The 4 dimensional volume corresponding to the parallelepiped will be $V_4 \;=\; \varepsilon_{\alpha\beta\gamma\delta}A^\alpha B^\beta C^\gamma D^\delta$. A "face" of this object will be a three dimensional volume, $\Sigma_\alpha \;=\; \varepsilon_{\alpha\beta\gamma\delta}B^\beta C^\gamma D^\delta$. Gauss' Theorem now reads

$$\int \partial_\alpha v^\alpha\,dV \;=\; \oint v^\alpha\,d\Sigma_\alpha \tag{24}$$

Replacing the general 4-vector $\underline{\mathbf{v}}$ with our 4-number density, the divergence theorem gives a new way to represent the conservation law we intuitively arrived at.

$$\oint N^{\alpha}\, d\Sigma_{\alpha} \;=\; 0 \tag{25}$$

This says the flux of the 4-number density through all Gaussian surfaces in space-time equals zero. "Whatever non-interacting material that flows into a spacetime volume will flow out of that spacetime volume". This expression can be expanded for spacetime coordinates $\underline{\mathbf{x}} \;=\; (x^0,\, x^1,\, x^2,\, x^3) \;=\; (t,\, \vec{x})$.

$$\begin{aligned}
&\left(\int N^0\left(x_2^0\right)\, \varepsilon_{0\alpha\beta\gamma}\, dx^{\alpha} dx^{\beta} dx^{\gamma} + \int N^0\left(x_1^0\right)\, \varepsilon_{0\alpha\beta\gamma}\, dx^{\alpha} dx^{\beta} dx^{\gamma}\right)_{time\ surfaces\ x_i^0 \,=\, f\left(x^1,\, x^2,\, x^3\right)} \\
&+ \left(\int N^1\left(x_2^1\right)\, \varepsilon_{\alpha 1\beta\gamma}\, dx^{\alpha} dx^{\beta} dx^{\gamma} + \int N^1\left(x_1^1\right)\, \varepsilon_{\alpha 1\beta\gamma}\, dx^{\alpha} dx^{\beta} dx^{\gamma}\right)_{x^1\ surfaces\ x_i^1 \,=\, f\left(x^0,\, x^2,\, x^3\right)} \\
&+ \ldots \;=\; 0
\end{aligned} \tag{26}$$

In Minkowski space, where we set up an orthogonal basis, where dx^{α} points in a unit direction, many terms drop out of the formula, allowing us to write the equation without the Levi-Civita nor the hidden Einstein summation of terms. Re-define the spacetime coordinates as $\underline{\mathbf{x}} \;=\; (t,\, x,\, y,\, z)$ to simplify.

$$\begin{aligned}
&\left(\int N^0\left(t_2\right)\, dxdydz + \int N^0\left(t_1\right)\, dxdydz\right) \\
&+ \left(\int N^1\left(x_2\right)\, dtdydz + \int N^1\left(x_1\right)\, dtdydz\right) \\
&+ \left(\int N^2\left(y_2\right)\, dtdxdz + \int N^2\left(y_1\right)\, dtdxdz\right) \\
&+ \left(\int N^3\left(z_2\right)\, dtdxdy + \int N^3\left(z_1\right)\, dtdxdy\right) \\
&= 0
\end{aligned} \tag{27}$$

We know that the direction of $\underline{\mathbf{N}}$ is going from an earlier time t_1 to a later time t_2, so there is a sign change, as the 4-number density "races towards" the surface of constant time t_1. Let's also imagine that this 4-D box is shrinking along the time axis, so that $t_2 \;=\; t_1 \,+\, dt$. Taking this into account and rearranging the "time" terms on the LHS and the "space" terms on the RHS gives the following:

$$\begin{aligned}
&\left(\int N^0\, dxdydz\right)_{t_1\,+\,dt} - \left(\int N^0\, dxdydz\right)_{t_1} = \\
&-\left[\left(\int N^1\, dydz\right)_{x_2} - \left(\int N^1\, dydz\right)_{x_1} + \left(\int N^2\, dxdz\right)_{y_2} - \left(\int N^2\, dxdz\right)_{y_1} + \left(\int N^3\, dxdy\right)_{z_2} - \left(\int N^3\, dxdy\right)_{z_1}\right] dt
\end{aligned} \tag{28}$$

Divide both sides by dt and take the limit $dt \rightarrow 0$ and we can see that this is a relationship between the time derivative of one term (LHS) and a dot product of spatial terms (RHS).

$$\frac{\partial}{\partial t}\int n\, dV \;=\; -\oint \vec{n}\cdot d\vec{a} \tag{29}$$

Notice that this is true only after one chooses the Minkowski coordinate system to observe the dust.

Another important example of matter is an electric current. Analogous to the dust's 4-number density, the 4-current density can be written in terms of the charge density (time piece) and the charge density moving with some speed through a volume of 3-D space according to some inertial observer (space piece). The charge density will be ρ and the current density will be $\vec{J} \;=\; \rho\vec{v}$.

$$\underline{\mathbf{J}} \;=\; \left(\rho,\, \vec{J}\right) \;=\; \rho\underline{\mathbf{u}} \tag{30}$$

Likewise, it has a conservation rule.

$$\frac{\partial\rho}{\partial t} = -\vec{\nabla}\cdot\vec{J} \tag{31}$$

or

$$\partial_\alpha J^\alpha = 0 \tag{32}$$

When thinking about describing electromagnetism in a space-time way, similar to the conservation law, there is a small problem with constructing more simple 4-vectors to use. The problem is that if one wanted to make a useful object out of the standard 3-vector way of thinking about E and B fields, it would be tough to construct two separate 4-vectors. Maybe a tensor? If the tensor only had 6 input components, then that could work. A typical tensor would have 16 components. If the tensor was symmetric, we would be down to only 10. Thats a start, but still not good enough. If we had an antisymmetric tensor, the diagonal elements would be zero, leaving only 6 elements in total to worry about. This gives us a tip that electromagnetism can be described with some antisymmetric tensor. This is already well established, but it is useful to have some intuition as to why it works. The Maxwell tensor is defined as $\underline{\underline{\mathbf{F}}}$.

$$F^{\mu\nu} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & B_z & -B_y \\ -E_y & -B_z & 0 & B_x \\ -E_z & B_y & -B_x & 0 \end{pmatrix} \tag{33}$$

where Maxwell's equations can be written in terms of derivatives of the Tensor components.

$$\partial_\mu F^{\mu\nu} = -\mu_0 J^\nu \tag{34}$$
$$\partial_\lambda F_{\mu\nu} + \partial_\nu F_{\lambda\mu} + \partial_\mu F_{\nu\lambda} = 0 \tag{35}$$

The conservation of current is a built-in feature of writing Maxwell's equations in this way. Looking at the divergence of the current density from the first of Maxwell's equations in this differential form, we get

$$\begin{aligned} \mu_0\partial_\mu J^\mu &= \partial_\mu\partial_\nu F^{\mu\nu} \\ &= \partial_\nu\partial_\mu F^{\nu\mu} \\ &= -\partial_\nu\partial_\mu F^{\mu\nu} \\ &= -\partial_\mu\partial_\nu F^{\mu\nu} \\ \therefore \partial_\mu J^\mu &= 0 \end{aligned} \tag{36}$$

This computation also reveals something in general about a contraction between antisymmetric ($\underline{\underline{\mathbf{A}}}$) and symetric tensors ($\underline{\underline{\mathbf{S}}}$).

$$A^{\mu\nu} S_{\mu\nu} \equiv 0 \tag{37}$$

Video 5: The Stress Energy Tensor and the Christoffel Symbol

Continue to imagine a cloud of dust in space where each particle of the dust has a rest mass m. In the rest frame of the dust element, its rest energy density is the density of the dust in that particular volume of space.

$$\rho_0 \mathrel{\hat{=}} mn_0 \tag{38}$$

Now, lets boost into a new frame moving with vecloity $\vec{v}$ with respect to the dust. The energy density in this boosted frame is changed. First, the energy of the individual moving particles gets a γ factor from their kinetic energy in this new frame. Second, the space (volume) gets contracted, which gives the number density a factor of γ as well.

$$\begin{aligned} \rho &= \gamma m \gamma n_0 \\ &= \gamma^2 \rho_0 \end{aligned} \tag{39}$$

If ρ was a 4-vector component, it would not transform with higher order powers of γ, which means it can't be part of a Lorentz vector (nor scalar). One other thing to recognize is that the energy density is built with a combination of two time-like components, namely the energy component of momentum $\underline{\mathbf{p}}$ and the number density component of the number vector $\underline{\mathbf{N}}$. This means that ρ can be written as

$$\rho = p^0 N^0 \equiv T^{00} \tag{40}$$

which implies that it is a part of some tensor. Let's write this tensor as the tensor product of the two 4-vectors.

$$\begin{aligned}\underline{\underline{\mathbf{T}}} &= \underline{\mathbf{p}} \bigotimes \underline{\mathbf{N}} \\ &= mn\, \underline{\mathbf{u}} \bigotimes \underline{\mathbf{u}} \\ &= \rho\, \underline{\mathbf{u}} \bigotimes \underline{\mathbf{u}}\end{aligned} \tag{41}$$

A component of this tensor is written with two indices.

$$\begin{aligned}T^{\mu\nu} &= p^\mu N^\nu \\ &= \rho u^\mu u^\nu\end{aligned} \tag{42}$$

This can physically be interpreted as the flux of the momentum component p^μ in the x^ν direction. With this interpretation, it is worth examaning the components of the tensor this way.

$$T^{00} \equiv \textit{flux of energy density in time "direction"} \tag{43}$$

$$T^{0i} \equiv \textit{flux of energy density in the } x^i \textit{ direction (through the } x^j \wedge x^k \textit{ plane)} \tag{44}$$

$$T^{i0} \equiv \textit{flux of momentum density in time "direction"} \tag{45}$$

$$T^{ij} \equiv \textit{flux of } p^i \textit{ in the } x^j \textit{ direction (units of pressure for } i = j \textit{ and shear for } i \neq j) \tag{46}$$

There is a symmetry to take note of, which is made apparent by writing it in terms of its components. I'll choose to write it in terms of the velocity components.

$$T^{00} = \gamma^2 \rho_0 \tag{47}$$

$$T^{0i} = \gamma^2 \rho_0 v^i \tag{48}$$

$$T^{i0} = \gamma^2 \rho_0 v^i \tag{49}$$

$$T^{ij} = \gamma^2 \rho_0 v^i v^j \tag{50}$$

This tensor $\underline{\underline{\mathbf{T}}}$ is called the Stress, Energy, Momentum tensor. The symmetry of $\underline{\underline{\mathbf{T}}}$ has been shown for dust. In general, this tensor is symmetric for all known physical structures in space-time (this is not proven here, but I suppose a rigorous proof can be found if one wants one). To deduce what $\underline{\underline{\mathbf{T}}}$ looks like, for some material being studied, we can apply the physical meaning of the tensor's individual components and construct it from there. Remember, we extracted the physical meaning of the components $T^{\mu\nu}$ from examining how dust behaves in inertial frames of reference.

For a new example, lets look at a perfect fluid. This is a fluid where there is no energy flow in some frame w.r.t. the fluid and no lateral stresses (basically no viscosity). As a neat "example" of a perfect fluid, one can think of "dry" water. The physics of such a fluid will be described by its energy density, ρ, and its pressure, P. Lets also assume that the pressure is isotropic to space (the pressure is the same in all three directions). In this frame of a perfect fluid, the momentum tensor only has four components.

$$T^{\mu\nu} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & P & 0 & 0 \\ 0 & 0 & P & 0 \\ 0 & 0 & 0 & P \end{pmatrix} \tag{51}$$

Picking out the rest frame of the fluid to be the frame of reference in question, where there is no energy flow, T^{00} is expressed in terms of the density and the four-velocity as the "dust" is described in the previous example.

$$T^{00} = \rho\, \underline{\mathbf{u}} \bigotimes \underline{\mathbf{u}} \tag{52}$$

The other three components are diagonal spatial terms which can be expressed with the Minkowski metric times the pressure term. The only caveat is that writting the pressure as this multiplication requires that we get rid of the extranuous "time" component by adding a pressure term to T^{00} we just defined. (We add it because we are working with a metric signature where the time element carries a minus sign in the metric)

$$\underline{\underline{\mathbf{T}}} = (\rho + P)\, \underline{\mathbf{u}} \bigotimes \underline{\mathbf{u}} + P\underline{\underline{\eta}} \tag{53}$$
$$T^{\mu\nu} = (\rho + P)\, u^{\mu}u^{\nu} + P\eta^{\mu\nu} \tag{54}$$

At some point we want to be able to express this tensor for a perfect fluid in a general space-time and not just a Minkowski space-time. For now, this is good enough for an intuition.

Let's consider the physics involving $\underline{\underline{\mathbf{T}}}$. Think of a cube of length l submerged in a fluid in space. We align the faces of the cube such that the normal vectors of the faces of the cube are parallel with the usual basis of Cartesian unit vectors $\hat{e}^i$. The pressure will be the force of the fluid acting on each face (surface area) of the cube. The force on a face of the cube is given classically as

$$\begin{aligned} \vec{F}_{on\ face\ whose\ normal\ vector\ is\ \hat{n}} &= \int \vec{P}_n da_n \\ &= \int \left(P_x\hat{x} + P_y\hat{y} + P_z\hat{z}\right)_n da_n \end{aligned}$$

which can be written in terms of components T^{ij}. The area, a, of each face is l^2. For each of the 6 faces of the cube, the force is found to be

$$\vec{F} = \mathbf{T} da_n \tag{55}$$
$$F^i = \pm T^{in} l^2;\ n = \{1,\ 2,\ 3\} \tag{56}$$

where $-$ represents a pressure acting on the outside of the cube, compressing it, and $+$ represents a pressure of the fluid flowing from the inside of the cube to the outside, if the cube allows for the fluid to flow through it. The letter n indicates the Cartesian direction parallel with the normal vector $\hat{n}$ of the face of the cube in question. In this fluid, the net force on the cube is zero. If the fluid can flow through the cube, the net flux is also zero. This is expected. An interesting way to give a physical argument for the spatial symmetry of T^{ij} is to consider torques and moments of inertia of this cube, still submersed in this fluid. Let's assume a axis of rotation to be passing through the center of the cube, aligned in the $\hat{e}^k$ direction (so the cube can rotate in the ij plane). Define $\vec{r}$ to point from the origin to the axis of rotation at the center of the cube and $\vec{r}'$ to point from the origin to a point on the surface of the cube that will be rotating about the axis. We can define a new variable, $\vec{r}_c = \vec{r} - \vec{r}'$, so that we can define the torque of a point on a face of the cube.

$$\vec{\tau} = \vec{r}_c \times \vec{F} \tag{57}$$

When writing the components of the torque, it's worth noting that $\vec{r}_c$ will always be pointing in the general direction of the unit normal vector of the face in question ($\vec{r}_c \cdot \hat{n} > 0$). This means that n will be restricted to which face of the cube you choose to calculate the torque with. To phrase this in a more direct way, the number n will equal the index chosen for the component of $\vec{r}_c$ in the calculation. For example, if the calculation requires the i^{th} component of $\vec{r}_c$, then $n \equiv i$. With this in mind, the torque of a point on a face of the cube about and axis parallel to one of the three Cartesian coordinate axes is written component wise as follows:

$$\begin{aligned} \tau_k &= \pm\varepsilon_{ijk} r_c^i F^j \\ &= \pm\varepsilon_{ijk} r_c^i T^{ji} l^2 \end{aligned} \tag{58}$$

where l^2 is the length of the cube squared (as to not confuse this exponent for one of the indices). So, suppose I wanted the z component of this torque. I can write my indices in terms of x, y and z now for the sake of simplicity.

$$\tau_z = \pm\left(r_c^x T^{yx} - r_c^y T^{xy}\right) l^2 \tag{59}$$

To simplify this further, lets also choose $\vec{r}_c$ such that $r_c^x = r_c^y = l/\sqrt{2}$.

$$\tau_z = \pm\frac{1}{\sqrt{2}}\left(T^{yx} - T^{xy}\right) l^3 \tag{60}$$

This shows that if the tensor is symmetric, then the torque completely vanishes. Even if it wasn't, the torque would still rapidly vanish as the size of the cube shrinks to zero. This isn't surprising.

Finally, lets consider the moment of inertia of this cube about the same axis we defined the torque. The inertia is

$$I = \int \vec{r}_c \cdot \vec{r}_c \, dm \tag{61}$$

The exact solution to this integral is not necessary. What's important to get out of this is the form of the inertia.

$$I \propto l^2 \rho l^3 = \rho l^5 \tag{62}$$

This is the cool part of the physics that's been being built towards. Remember that the torque is related to the moment of inertia in a similar way that force is related to mass. Also, notice that as $l \to 0$ the torque dies off as l^3, but the inertia dies off more rapidly as l^5. If we check the angular acceleration of the cube, we find that it is inversly related to the length of the cube.

$$\ddot{\varphi} \propto \frac{T^{yx} - T^{xy}}{l^2} \tag{63}$$

Assuming that the spatial tensor components were not symmetric, this would describe some angular acceleration term that would exist on small scales $l << 0$. Imagine if the molecules in a cup of water were to spontanously begin to spin and rotate faster and faster for no obvious reason other than the stress energy not having symmetric terms. Because we don't see this behavior in nature, we can use this as good motivation for saying that the stress energy tensor is symmetric. This example gives us a physical intuition about the symmetry of the spatial components, but again, another derivation of the stress energy tensor reveals the symmetry of the tensor in a more rigorous way, in both spatial and temporal components.

Its important to bring up the conservation of energy and momentum, as a great deal of physics understanding hinges on these two conservation laws. In terms of the stress-energy-momentum tensor, this single tensor is a conserved quantity. When defining the tensor from a field theory perspective, one can prove the conservation law directly.

$$\partial_\mu T^{\mu\nu} = 0 \tag{64}$$

This is similar to Maxwell's Equations in covariant form.

$$\partial_\mu F^{\mu\nu} = -\mu_0 J^\mu \tag{65}$$

The main difference is that there is no vector "source" term for the variation of the stress-energy-momentum terms. I suppose if one finds this strange acceleration in nature, this would prove that our theories need to be re-structured to allow for the non-conservation of energy and momentum!

Let's consider yet another common example of matter; lets consider a point mass, m_0 moving along a world line in space-time. The world line will be some four-vector of position parametrized by the proper time of the mass moving along that world line. The world line will be represented as $\underline{\mathbf{x}}(\tau)$. To represent the stress-energy-tensor of this point mass it's beneficial to make use of the delta function and write it in terms of the mass/energy density. Indicate the location of the point mass with an apostrophe ('). To start, we know the first element of the tensor is the same as that of the dust.

$$T^{00} = \rho u^0 u^0; \;\; [\rho] = \frac{mass}{length^3} \tag{66}$$

For the point mass in Euclidean space the mass density is described as

$$\rho = m_0 \delta^3 \left(\vec{x} - \vec{x}'\right); \;\; [\rho] = \frac{mass}{length^3} \tag{67}$$

but we want an expression for space-time that will adhere to the correct dimensions for ρ. The space-time delta function will carry a temporal dimension for any observer in any reference frame. Again, these space-time coordinates for the point mass are chosen to be expressed as a function of proper time. Therefore, this mass density with the space-time delta function must be integrated over all of "proper" time. This also maintains the dimensionality of ρ when given the 4D space-time delta function.

$$\rho = m_0 \int \delta^4 \left(\underline{\mathbf{x}} - \underline{\mathbf{x}}'\right) \, d\tau; \;\; [\rho] = \frac{mass}{length^3 \times time} \times time \tag{68}$$

We can express $\underline{\underline{\mathbf{T}}}$ of the point mass the same way we expressed the tensor for dust. The only difference is that the Dirac Delta function is serving a role here in place of the number density, since this point mass is obviously the only mass in question.

$$T^{\mu\nu} = m_0 \int u^\mu u^\nu \delta^4 \left(\underline{\mathbf{x}} - \underline{\mathbf{x}}'\right) \, d\tau \tag{69}$$

As a calculus reminder of integrals with delta functions,

$$\int f(x)\delta(g(x)) \, dx = \sum_{i=1}^{N} \frac{f(x_i)}{g'(x)|_{x=x_i}} \, |_{x_i=x_0} \tag{70}$$

where x_0 are the N roots of $g(x)$ and $g'(x)$ is the first derivative with respect to x. With this in mind, lets solve the integral over the proper time, choosing one of the four delta functions to work with.

$$\int \left[u^\mu u^\nu \delta^3 \left(\vec{x}(\tau) - \vec{x}(\tau)'\right)\right] \delta(t - \tau) \, d\tau = \frac{u^\mu u^\nu \delta^3 \left(\vec{x}(t) - \vec{x}(t)'\right)}{1} \tag{71}$$

Here, this 1 in the denominator (again, in natural units) is the zeroth component of the mass' 4-velocity in it's frame of reference. The final form of the stress-energy-momentum tensor for the point mass can now be written without that integral.

$$T^{\mu\nu} = \frac{m_0}{u^0} u^\mu u^\nu \delta^3 \left(\vec{x}(t) - \vec{x}(t)'\right) \tag{72}$$

To breifly summarize the introduction of the stress-energy-momentum tensor, we derived three examples listed below.

$$T^{\mu\nu}_{non-interacting\ dust} = \rho u^\mu u^\nu \tag{73}$$

$$T^{\mu\nu}_{perfect\ fluid} = (\rho + P) u^\mu u^\nu + P \eta^{\mu\nu} \tag{74}$$

$$T^{\mu\nu}_{point\ mass} = \frac{m_0}{u^0} u^\mu u^\nu \delta^3 \left(\vec{x}(t) - \vec{x}(t)'\right) \tag{75}$$

Let's consider the classical field equation for Newtonian gravity.

$$\nabla^2 \Phi = 4\pi G \rho \tag{76}$$

Recall that ρ is just one component of a tensor. For a better theory of gravity, we should not choose to pick out this one term out of the other 15 terms of the tensor. This means we should seek an equation that looks like this classical differential equation, but with $T^{\mu\nu}$ on the RHS. This idea will be brought up again later. We bring it up because we have at least began to get a feel for the stress-energy-momentum tensor, so we should keep gravity in mind in how it will be connected with this tensor.

As an introduction into curved space-time, let's first start with "simple" Minkowski space-time defined in curvilinear coordinates. In traditional Cartesian coordinates, our 4-vector coordinates are given as $\underline{\mathbf{x}} = (t, x, y, z)$. In this basis, the differential of the 4-vector can be written simply as contractions with the basis vectors with straightforward understanding.

$$\underline{\mathbf{dx}} = dx^\mu \hat{e}_\mu \tag{77}$$

The easiest way to switch to curvilinear coordinates is to change to cylinderical coordinates with the usual mapping. $\underline{\mathbf{x}} = (t, r, \varphi, z)$. In this basis, the differential of the 4-vector is written the same way, but unit analysis of the angular piece, $dx^2 \hat{e}_2 \equiv d\varphi \, \widehat{\varphi}$, shows that the unit vector $\widehat{\varphi}$ must carry the dimensionality of length to make the displacement vector make sense. This shows that these coordinates are not a "straightforward" basis. In a more math oriented way of putting it, the basis vectors are not always necessarily normalized. For the cylindrical coordinates, $\widehat{\varphi} \cdot \widehat{\varphi} \neq 1$. With two different coordinate systems established, its worth understanding how to transform from one set of coordinates to another in a general sense. Lets define the transformation from one coordinate system, $\underline{\mathbf{x}}$ (say Cartesian coordinates), to another coordinate system, $\underline{\tilde{\mathbf{x}}}$ (say cylinderical coordinates).

$$L^\alpha_\mu = \frac{\partial x^\alpha}{\partial \tilde{x}^\mu} \tag{78}$$

For these two specific coordinate systems, the form of $\underline{\underline{\mathbf{L}}}$ can be written as a matrix.

$$\underline{\underline{\mathbf{L}}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\varphi & \sin\varphi & 0 \\ 0 & -r\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{79}$$

$$\underline{\underline{\mathbf{L}}}^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi/r & 0 \\ 0 & \sin\varphi & \cos\varphi/r & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{80}$$

Video 6: The Principle of Equivalence

Continuing where the last video left off, we define the basis vectors for the cylindrical coordinates in terms of the Cartesian basis.

$$\begin{aligned} \widehat{r} &= \cos\varphi\,\hat{x} + \sin\varphi\,\hat{y} = L_1^{\alpha}\,\hat{e}_{\alpha} \\ \widehat{\varphi} &= -\sin\varphi\,\hat{x} + \cos\varphi\,\hat{y} = L_2^{\alpha}\,\hat{e}_{\alpha} \end{aligned} \tag{81}$$

Now, remember that one of the first tensors to work with is the metric for Minkowski space.

$$\eta_{\mu\nu} = \hat{e}_{\mu} \cdot \hat{e}_{\nu} \tag{82}$$

To simply define a general metric, we do the same thing with any coordinate basis. The general metric is labeled as $\underline{\underline{\mathbf{g}}}$.

$$g_{\mu\nu} = \hat{e}_{\mu} \cdot \hat{e}_{\nu} \tag{83}$$

At some point, we want to calculate derivatives. With curvilinear coordinate systems in general, the basis vectors vary with the coordinates.

$$\frac{\partial\hat{r}}{\partial r} = 0 \tag{84}$$

$$\frac{\partial\hat{r}}{\partial\varphi} = \frac{\widehat{\varphi}}{r} \tag{85}$$

$$\frac{\partial\widehat{\varphi}}{\partial r} = \frac{\widehat{\varphi}}{r} \tag{86}$$

$$\frac{\partial\widehat{\varphi}}{\partial\varphi} = -r\,\hat{r} \tag{87}$$

Lets say there is a vector $\underline{\mathbf{V}}$ that is defined in cylindrical coordinates as $V^{\alpha}\hat{e}_{\alpha}$. Supoose we wanted to write the "gradient" of this vector. In an abstract, ugly notation, this would be

$$\underline{\nabla\mathbf{V}} = \partial_{\beta}(V^{\alpha}\hat{e}_{\alpha})\,\widetilde{\omega}^{\beta} \tag{88}$$

where $\widetilde{\omega}^{\beta}$ are the basis one-forms ("inverses" of the basis vectors $\hat{e}_{\beta}$). Ignoring the $\widetilde{\omega}^{\beta}$ for a moment, the rest of this "gradient" should be some "tensorial" object. Lets expand that derivative term.

$$\frac{\partial\underline{\mathbf{V}}}{\partial x^{\beta}} = \frac{\partial V^{\alpha}}{\partial x^{\beta}}\hat{e}_{\alpha} + V^{\alpha}\frac{\partial\hat{e}_{\alpha}}{\partial x^{\beta}} \tag{89}$$

Note that the sum of these two components behaves as a tensor, but the individual addens do not. With that in the back of our head, lets look at the second term. We have shown (at least for cylindrical coordinates) that the derivatives of basis vectors can be written in terms of other basis vectors. So, we can choose to write the second term as a linear combination of basis vectors. To write it in the Einstein index summation notation, we introduce a new symbol to help define the coefficients for the linear combination.

$$\partial_{\beta}\hat{e}_{\alpha} = \Gamma^{\mu}_{\alpha\beta}\hat{e}_{\mu} \tag{90}$$

The Christoffel symbol Γ is not a tensor, but it can be used to get tensorial objects. The Christoffel symbol is understood with three indices. The top index shows which unit vector we are using as a basis for the linear combination, the first lower index shows which basis vector is being differentiated and the second lower index shows which coordinate we are differentiating with. For the cylindrical coordinates, the derivatives of the basis vectors we worked out earlier shows us the "relavent" Christoffel symbols.

$$\Gamma^r_{rr} = \Gamma^\varphi_{rr} = 0 \tag{91}$$
$$\Gamma^\varphi_{r\varphi} = \frac{1}{r} \tag{92}$$
$$\Gamma^\varphi_{\varphi r} = \frac{1}{r} \tag{93}$$
$$\Gamma^r_{\varphi\varphi} = -r \tag{94}$$

Now, the derivative of the 4-vector can be written in terms of the linear combination of basis vectors, or in terms of the Christoffel symbols.

$$\partial_\beta \underline{\mathbf{V}} = \partial_\beta V^\alpha \hat{e}_\alpha + V^\alpha \Gamma^\mu_{\alpha\beta} \hat{e}_\mu \tag{95}$$

Taking notice of the dummy indices required for the contractions, we can factor out the basis vectors $\hat{e}_\alpha$.

$$\partial_\beta \underline{\mathbf{V}} = \left(\partial_\beta V^\alpha + V^\mu \Gamma^\alpha_{\mu\beta}\right) \hat{e}_\alpha \tag{96}$$

The term in the parenthesis comes up A LOT, so lets give it some new notation and a fancy name to boot! It's called the covariant derivative. So, the term in the parenthesis is now the covariant derivative of component V^α with respect to x^β and is labeled as $\nabla_\beta V^\alpha$.

$$\begin{aligned} \partial_\beta \underline{\mathbf{V}} &= \left(\nabla_\beta V^\alpha\right) \hat{e}_\alpha \\ \nabla_\beta V^\alpha &= \partial_\beta V^\alpha + \Gamma^\alpha_{\mu\beta} V^\mu \end{aligned} \tag{97}$$

An immediate application for the covarient derivative can be for the divergence of a 4-vector.

$$\underline{\nabla} \cdot \underline{\mathbf{V}} = \nabla_\alpha V^\alpha = \partial_\alpha V^\alpha + \Gamma^\alpha_{\beta\alpha} V^\beta \tag{98}$$

For the cylindrical coordinates this has five terms.

$$\underline{\nabla} \cdot \underline{\mathbf{V}} = \partial_t V^t + \partial_r V^r + \partial_\varphi V^\varphi + \partial_z V^z + \frac{1}{r} V^r \tag{99}$$

What about taking derivatives of things other than 4-vectors? For scalars, they have no unit basis vectors by definition, so they're straightforward. How about a 1-form? For a 1-form, let's start with our understanding of covariant derivatives for scalars and 4-vectors, combined with the knowledge that the contraction of a 4-vector with a 1-form results in a Lorentz scalar.

$$\begin{aligned} \nabla_\rho \left(W_\alpha V^\alpha\right) &= \partial_\rho \left(W_\alpha V^\alpha\right) \\ \nabla_\rho W_\alpha V^\alpha + W_\alpha \nabla_\rho V^\alpha &= \partial_\rho W_\alpha V^\alpha + W_\alpha \partial_\rho V^\alpha \\ \nabla_\rho W_\alpha V^\alpha + W_\alpha \nabla_\rho V^\alpha &= \partial_\rho W_\alpha V^\alpha + W_\alpha \nabla_\rho V^\alpha - W_\alpha \Gamma^\alpha_{\mu\rho} V^\mu \\ \nabla_\rho W_\alpha V^\alpha + \cancel{W_\alpha \nabla_\rho V^\alpha} &= \partial_\rho W_\alpha V^\alpha + \cancel{W_\alpha \nabla_\rho V^\alpha} - W_\alpha \Gamma^\alpha_{\mu\rho} V^\mu \\ \nabla_\rho W_\alpha V^\alpha &= \partial_\rho W_\alpha V^\alpha - W_\alpha \Gamma^\alpha_{\mu\rho} V^\mu \\ \left(\nabla_\rho W_\alpha\right) \cancel{V^\alpha} &= \left(\partial_\rho W_\alpha - W_\mu \Gamma^\mu_{\alpha\rho}\right) \cancel{V^\alpha} \\ \nabla_\rho W_\alpha &= \partial_\rho W_\alpha - \Gamma^\mu_{\alpha\rho} W_\mu \end{aligned} \tag{100}$$

For tensors we can deduce how to write it based on index balancing and whether or not the index is "upstairs" or "downstairs". This is not "rigorous", but it gives enough intuition to determine how many terms we need (an additional Christoffel symbol per index) and which sign it needs (+ for upstairs and − for downstairs).

$$\nabla_\rho \Phi = \partial_\rho \Phi \tag{101}$$

$$\nabla_\rho V^\alpha = \partial_\rho V^\alpha + \Gamma^\alpha_{\sigma\rho} V^\sigma \tag{102}$$

$$\nabla_\rho W_\beta = \partial_\rho W_\beta - \Gamma^\sigma_{\beta\rho} W_\sigma \tag{103}$$

$$\nabla_\rho T^{\alpha\beta\ldots\mu}_{\gamma\delta\ldots\nu} = \partial_\rho T^{\alpha\beta\ldots\mu}_{\gamma\delta\ldots\nu} + \left(\Gamma^\alpha_{\sigma\rho} T^{\sigma\beta\ldots\mu}_{\gamma\delta\ldots\nu} + \Gamma^\beta_{\sigma\rho} T^{\alpha\sigma\ldots\mu}_{\gamma\delta\ldots\nu} + \ldots + \Gamma^\mu_{\sigma\rho} T^{\alpha\beta\ldots\sigma}_{\gamma\delta\ldots\nu}\right) - \left(\Gamma^\sigma_{\gamma\rho} T^{\alpha\beta\ldots\mu}_{\sigma\delta\ldots\nu} + \Gamma^\sigma_{\delta\rho} T^{\alpha\beta\ldots\mu}_{\gamma\sigma\ldots\nu} + \ldots + \Gamma^\sigma_{\nu\rho} T^{\alpha\beta\ldots\mu}_{\gamma\delta\ldots\sigma}\right) \tag{104}$$

This information will help to define a better way of finding Christoffel symbols than building the symbols piece by piece, solving the derivatives of the basis vectors one by one. There are still some features of tensors and derivatives that need to be seen before that "better way" can be shown. To ease our way into it, lets think of a strange operator acting on a scalar. Namely, what would the "double gradient" of a scalar look like in a Cartesian basis?

$$\underline{\nabla\nabla}\Phi = \partial_\alpha \partial_\beta \Phi \, \tilde{\omega}^\alpha \tilde{\omega}^\beta \tag{105}$$

Its clear that these partial derivatives are symmetric and can be swapped. This means that the "tensorial" object this creates is symmertrical in indices α and β. What about a general representation of this "double gradient"? The only difference is that it will be dependent on covarient derivatives.

$$\underline{\nabla\nabla}\Phi = \nabla_\alpha \nabla_\beta \Phi \, \tilde{\omega}^\alpha \tilde{\omega}^\beta \tag{106}$$

One thing to realize about this is that if this double gradient is a tensor then, regardless of which coordinate representation we choose, it should be symmetric, because it is already shown to be obviously symmetric under the Cartesian coordinate basis. Given that this tensor should be symmetric means that the order of operations of two covarient derivatives can be swapped, just like partial derivatives.

$$\nabla_\alpha \nabla_\beta \Phi = \nabla_\beta \nabla_\alpha \Phi \tag{107}$$

Lets expand these covariant derivatives out.

$$\begin{aligned} \nabla_\alpha \left(\partial_\beta \Phi\right) &= \nabla_\beta \left(\partial_\alpha \Phi\right) \\ \cancel{\partial_\alpha \left(\partial_\beta \Phi\right)} - \Gamma^\sigma_{\beta\alpha} \left(\partial_\sigma \Phi\right) &= \cancel{\partial_\beta \left(\partial_\alpha \Phi\right)} - \Gamma^\sigma_{\alpha\beta} \left(\partial_\sigma \Phi\right) \\ \left(\Gamma^\sigma_{\alpha\beta} - \Gamma^\sigma_{\beta\alpha}\right) \partial_\sigma \Phi &= 0 \end{aligned} \tag{108}$$

This reveals something about the Christoffel symbols: If we require that this tensorial operation be symmetric in all representations as a true tensor, the Christoffel symbols can also be shown to be symmetric given that the single gradient of the scalar is non-zero.

In light of seeing symmetric vs anti-symmetric tensors it's useful now to introduce some more notation. We can define "simplified" notation to express a "symmetric" operation and an "anti-symmetric" operation of a tensor.

$$T_{\{\alpha\beta\}} \equiv \frac{1}{2}\left(T_{\alpha\beta} + T_{\beta\alpha}\right) \tag{109}$$

$$T_{[\alpha\beta]} \equiv \frac{1}{2}\left(T_{\alpha\beta} - T_{\beta\alpha}\right) \tag{110}$$

Given what we've just shown for the Christoffel symbols, we can write some expressions.

$$\Gamma^\sigma_{\{\alpha\beta\}} = \Gamma^\sigma_{\alpha\beta} \tag{111}$$

$$\Gamma^\sigma_{[\alpha\beta]} = 0 \tag{112}$$

We can also remember what happens when a symmetric matrix gets contracted with an anti-symmetric matrix, $\underline{\underline{\mathbf{A}}}$.

$$\Gamma^\sigma_{\alpha\beta} A^{\alpha\beta} = \Gamma^\sigma_{\{\alpha\beta\}} A^{[\alpha\beta]} = 0 \tag{113}$$

With all of this in mind, lets finally derive the "better way" to get Christoffel symbols. We begin the derivation with the gradient of the metric tensor.

$$\underline{\nabla}\,\underline{\underline{\mathbf{g}}} = \nabla_\gamma g_{\alpha\beta} \, \tilde{\omega}^\beta \bigotimes \tilde{\omega}^\alpha \bigotimes \tilde{\omega}^\gamma \tag{114}$$

Again, we assert that we want tensorial operations like this to be the same in any given coordinate representation. So, this gradient of a general metric should be the same as the gradient of the Minkowski metric.

$$\underline{\nabla}\,\underline{\underline{\eta}} \;=\; \nabla_\gamma \eta_{\alpha\beta}\, \widetilde{\omega}^\beta \bigotimes \widetilde{\omega}^\alpha \bigotimes \widetilde{\omega}^\gamma \;\equiv\; 0 \tag{115}$$

This leads us to require that the covariant derivative of any general metric must be equal to zero.

$$\nabla_\gamma g_{\alpha\beta} \;=\; 0 \tag{116}$$

With this choice of three indices (γ, α and β) this metric relation can also be expressed by strategic swapping of these indices. For this derivation, we shuffle the indices in the same way we shuffle indices for the Levi-Civita symbol that keep them equal to each other. For example:

$$\varepsilon_{abc} \;=\; \varepsilon_{cab} \;=\; \varepsilon_{bca}$$

Also, for each of these three ways to express this metric relation, we expand the definition of the covariant derivative with the Christoffel symbols. This is what will lead into the "better" form of the Christoffel symbol.

$$\nabla_\gamma g_{\alpha\beta} \;=\; \partial_\gamma g_{\alpha\beta} \;-\; \Gamma^\mu_{\alpha\gamma} g_{\mu\beta} \;-\; \Gamma^\mu_{\beta\gamma} g_{\alpha\mu} \;=\; 0 \tag{117}$$

$$\nabla_\beta g_{\gamma\alpha} \;=\; \partial_\beta g_{\gamma\alpha} \;-\; \Gamma^\mu_{\gamma\beta} g_{\mu\alpha} \;-\; \Gamma^\mu_{\alpha\beta} g_{\gamma\mu} \;=\; 0 \tag{118}$$

$$\nabla_\alpha g_{\beta\gamma} \;=\; \partial_\alpha g_{\beta\gamma} \;-\; \Gamma^\mu_{\beta\alpha} g_{\mu\gamma} \;-\; \Gamma^\mu_{\gamma\alpha} g_{\beta\mu} \;=\; 0 \tag{119}$$

All of these covariant derivatives are equal to zero, hence all of these expressions are equal to each other; Expression (117) = Expression (118) = Expression (119). This means that Expression (117) - Expression (118) = Expression (119). This ALSO means that Expression (117) - Expression (118) - Expression (119) = 0. Without much thought, these seem to be ugly, arbitrary computations. But, by taking advantage of the symmetry of the Christoffel symbols, many terms cancel from this combination of the expressions and leaves behind one Christoffel symbol and a bunch of metric terms. Without writing all of the cancellations and simplifications explicity, this gives the "better way" of expressing the Christoffel symbols.

$$\Gamma^\mu_{\alpha\beta} \;=\; -\frac{1}{2} g^{\mu\gamma} \left(\partial_\gamma g_{\alpha\beta} \;-\; \partial_\beta g_{\gamma\alpha} \;-\; \partial_\alpha g_{\beta\gamma}\right) \tag{120}$$

Video 7: The Principle of Equivalence Continued; Parallel Transport

When we have gravity we cannot cover all of space-time with an inertial frame like we can in Special Relativity. One feature of gravity is redshifted light when light travels away from a gravitational source. This reshifting phenomena has been experimentally measured and proved to exist. A fun thought experiment in which gravitational redshift manifests is to think about dropping a rock with rest energy m from a height y here near the surface of the Earth on top of a hill. For the fun part of the thought experiment, imagine that there is a device, a photonulator, at the bottom of the hill that can convert the rock into a raw photon. Lets also assume that this device managed to convert all of the rock's energy the moment before hitting the ground. This means that the fallen rock will have all of its energy converted to that of a photon at the last moment before impact (Again, remember that we're using natural units.)

$$m \;\rightarrow\; m \,+\, mgy \;= \rightarrow \omega_{fall} \tag{121}$$

If this photon was completely and perfectly reflected back up the hill, assuming no loss of energy of the photon during the reflection, the photon will travel back with the kinetic energy (and rest energy) of the rock.

$$E_{photon} \;=\; \omega_{fall} \;=\; m \,+\, mgy \tag{122}$$

If the photon went back up the hill and entered into a "re-rockalizer", to convert the photon back into the rock that we dropped in the first place, there would be a difference in energy between when the rock was initially dropped and anytime the rock came back up as a photon. This is because the photon is now carrying the energy of the rest particle, but also the gravitational potential energy that the rock initially converted to kinetic energy. The rock could even be interpreted as having more mass when it comes back vs. what it started with before falling. This thought experiment dilemma is resolved by allowing the frequency of the returning light to be lower than the frequency of the "falling" light. This frequency change would be determined by the gravitational potential. The light moving away from the pull of gravity would be

$$\omega_{rise} \;=\; \omega_{fall}\,(1 \;-\; gy) \tag{123}$$

$$NOTE : This\ Argument\ has\ Issues!!!$$

Despite the wierd (and probably wrong) argument here, gravitational redshift is an established fact that also has a more solid theoretical framework. Now, suppose that there was a large region of space-time that could be covered by a single Lorentz frame. Thinking about the light that climbs out of a gravitational potential well, we know experimentally that the light changes in energy. But, taking small snapshots of the photon from one region of space-time to the next in this frame would show that they are identical. The photon in one snapshot cannot be the same as the photon in the other if the energy is changing. The energy is changing, therefore the frequency is changing, therefore the period is changing. This just shows that we cannot think about global coordinates on a large lorentz frame when thinking about gravity and how it effects other things. The Lorentz frames are designed to specify inertial observers, but gravity does not allow for inertial observers. This means that we cannot describe space-time globally with gravity, but we can describe it locally. From Einstein's prospective, the next best thing to describe, rather than an inertial frame, was a freely falling frame. Objects that are freely falling with respect to other freely falling objects, that are also assumed to not be effecting each other gravitationally in a noticable way, "appear" to be in an inertial frame. In other words, the net force acting on those objects in the freely falling frame amounts to zero. This makes sense because the frame itself is required to accelerate with the object in a free fall frame of reference. The one caveat is that tides breakdown the notion of inertial freely falling frames. This shows that, in order to think about space-time for freely falling frames, we need to think about local regions of space-time when dealing with gravity.

Let $\underline{\mathbf{x}}$ be our starting coordinates with the metric $g_{\alpha\beta}$. Let there be another set of coordinates $\underline{\mathbf{y}}$ in which space-time is locally flat in the vicinity of some event P. Assume that there is a mapping between these two sets of coordinates, $x^{\alpha} \;=\; x^{\alpha}\left(\underline{\mathbf{y}}\right)$, allowing a transformation between the two.

$$L^{\alpha}_{\mu} \;=\; \frac{\partial x^{\alpha}}{\partial y^{\mu}} \tag{124}$$

The goal is to find a coordinate system such that

$$g_{\mu\nu} \;=\; L^{\alpha}_{\mu} L^{\beta}_{\nu} g_{\alpha\beta} \;\equiv\; \eta_{\mu\nu}$$

over as large of a region as possible. First, recognize that these are all functions. We can expand these functions as a Taylor series about the event P. It is reasonable to do this because we expect the metric to be locally equal to the Minkowski metric. The coordinate transformation can be whatever we choose.

$$g_{\alpha\beta} \;\approx\; g_{\alpha\beta}|_{\mathbf{P}} \;+\; \left(x^{\gamma} \;-\; y^{\gamma}_{P}\right)\partial_{\gamma} g_{\alpha\beta}|_{\mathbf{P}} \;+\; \frac{1}{2}\left(x^{\gamma} \;-\; y^{\gamma}_{P}\right)\left(x^{\sigma} \;-\; y^{\sigma}_{P}\right)\partial_{\gamma}\partial_{\sigma} g_{\alpha\beta}|_{\mathbf{P}} \tag{125}$$

$$L^{\alpha}_{\mu} \;\approx\; L^{\alpha}_{\mu}|_{\mathbf{P}} \;+\; \left(x^{\gamma} \;-\; y^{\gamma}_{P}\right)\partial_{\gamma} L^{\alpha}_{\mu}|_{\mathbf{P}} \;+\; \frac{1}{2}\left(x^{\gamma} \;-\; y^{\gamma}_{P}\right)\left(x^{\sigma} \;-\; y^{\sigma}_{P}\right)\partial_{\gamma}\partial_{\sigma} L^{\alpha}_{\mu}|_{\mathbf{P}} \tag{126}$$

Before continuing with the computation, its important to take note of the constraints and the degrees of freedom. The metric $g_{\alpha\beta}$ is what's given, so this metric, along with its first and second derivatives are the constraints. Since we are free to choose the new set of coordinates to be whatever we like, the transformation matrix and its first and second derivatives are the degrees of freedom.

$$L^{\alpha}_{\mu} L^{\beta}_{\nu} g_{\alpha\beta} \;\approx\; \left(L^{\alpha}_{\mu}|_{\mathbf{P}}\right)\left(L^{\beta}_{\nu}|_{\mathbf{P}}\right)\left(g_{\alpha\beta}|_{\mathbf{P}}\right) \;+\; \left(x^{\gamma} \;-\; y^{\gamma}_{P}\right)(First\ Order\ Terms) \;+\; \frac{1}{2}\left(x^{\gamma} \;-\; y^{\gamma}_{P}\right)\left(x^{\sigma} \;-\; y^{\sigma}_{P}\right)(Second\ Order\ Terms) \tag{127}$$

At zeroth order, the first constraint is $g_{\alpha\beta}$. Being a rank 2, 4 dimensional symmetric tensor, this gives $Red(2,4) \;=\; 10$ constraints. The transformation matrix, on the other hand, is not symmetric. This means there are 16 degrees of freedom to ensure the 10 constraints are satisfied. This means that, at zeroth order, we can easily satisfy the constraints (find a local Lorentz frame sufficiently close to event P). The 6 additional degrees of freedom 'left over' are the three boosts and three rotations we expect.

Next, at first order, the constraints are from derivatives of the metric $\partial_{\gamma} g_{\alpha\beta}$. That's $4 \cdot Red(2,4) \;=\; 40$ constraints. The first derivative of the transformation matrix, $\partial_{\gamma} L^{\alpha}_{\mu}$ is symmetric with respect to the bottom indices (as they are both referencing derivatives). This will have degrees of freedom as the number of constraints that must be satisfied.

Finally, at second order, the constraints are from a second order derivative of the metric, where indices from the derivatives are symmetric as well. This gives $Red^{2}(2,4) \;=\; 100$ constraints. Whereas the degrees of freedom from the second order derivative of the transformation matrix (which looks like four symmetric rank 3, 4 dimensional tensors) are fewer than the constraints; $4 \cdot Red(3,4) = 80$ degrees of freedom.

What this shows is that any metric can be described as flat space only up to second order corrections.

$$g_{\mu\nu} = \eta_{\mu\nu} + \mathfrak{O}\left\{\partial^2 g\right\} \tag{128}$$

The second derivative of the metric is alluding to the definition of the curvature of the spacetime. It turns out that the remaining 20 terms that are missing from the previous counting scheme are found in the Riemann curvature tensor. The size of the spacetime region where this approximation method is effective is proportional to the inverse of a length scale.

$$\ell \sim \frac{1}{\sqrt{\partial^2 g}} \tag{129}$$

Let's define what we mean by curved manifolds. A curved manifold is one in which initially parallel trajectories do not remain parallel. An easy to visualize example is that of the surface of a 2-sphere. An interesting side note of a surface that LOOKS curved, but is not (by this definition) is the surface of a cylinder. Imagine an infinitely long cylinder (no end caps) with initially parallel vectors on the surface. The transport of those vectors will remain parallel throughout. Another way to think of it is that you can find a way to perfectly flatten the cylinder's surface, or completely project the surface onto a 2D plane without "ripping" the surface, like what would need to be done for flattening the surface of a 2-sphere.

We want a way to handle vectors and tensors in this curved spacetime. One thing to help us is to realize the space that we're dealing with when we think of drawing vectors on a piece of paper. The vectors are living in a tangent space; always tangent to the "surface" of the space they occupy. So, what does that tell us about vectors along the surface of a 2-sphere? Two vectors, at first glance, seem difficult to compare to one another depending on where these vectors are located on the sphere. So, consider a curve, γ, that exists in a curved manifold. Along γ there exists two points, P and Q, where $P = x^\alpha$ and $Q = x^\alpha + dx^\alpha$. Now, suppose there is a vector field that fills this manifold. The vector field $\underline{\mathbf{A}}$ has values $A^\alpha(P)$ and $A^\alpha(Q)$ at the points P and Q respectively. We want to know how to take the derivative of the vector field from points P to Q. Let's start with the definition of a derivative as a first attempt.

$$\partial_\beta A^\alpha = \frac{\partial A^\alpha}{\partial x^\beta} \equiv \lim_{dx^\beta \to 0}\left\{\frac{A^\alpha(Q) - A^\alpha(P)}{dx^\beta}\right\} \tag{130}$$

This seems to be ok, but a closer look (perhaps a picture on my part would make this more obvious) reminds us that P and Q do not share the same tangent space, as they exist along γ in a given curved manifold. This means that their basis vectors point in different directions, and thus the basis vectors will be changing along γ in general as we try to compute the derivative. One can still compute the derivative like this, but the result will not be a tensorial object, which is what we are after for the sake of physics. A tensorial object would have the flexibility of being written in another coordinate basis, say y^μ.

$$\begin{aligned}\partial_\nu A^\mu &= \frac{\partial A^\alpha}{\partial x^\beta}\frac{\partial x^\beta}{\partial y^\nu}\frac{\partial y^\mu}{\partial x^\alpha} \\ &= \frac{\partial x^\beta}{\partial y^\nu}\frac{\partial y^\mu}{\partial x^\alpha}\partial_\beta A^\alpha\end{aligned} \tag{131}$$

For this argument, let's assume that $\underline{\mathbf{A}}$ is already a tensorial object, meaning

$$A^\mu = \frac{\partial y^\mu}{\partial x^\alpha}A^\alpha \tag{132}$$

$$\partial_\nu = \frac{\partial x^\beta}{\partial y^\nu}\partial_\beta \tag{133}$$

Carrying on with the math by substituting these two equations into the left hand side of Equation (131) gives the following:

$$\begin{aligned}\frac{\partial x^\beta}{\partial y^\nu}\partial_\beta\left[\frac{\partial y^\mu}{\partial x^\alpha}A^\alpha\right] &\stackrel{?}{=} R.H.S. \\ \frac{\partial x^\beta}{\partial y^\nu}\frac{\partial y^\mu}{\partial x^\alpha}\partial_\beta A^\alpha + \frac{\partial x^\beta}{\partial y^\nu}\frac{\partial^2 y^\mu}{\partial x^\alpha \partial x^\beta}A^\alpha &\neq R.H.S.\end{aligned} \tag{134}$$

which shows that there is an extra term that makes this derivative not a tensor. A way of fixing this issue is to have the two vectors at the same point. So, to take the derivative of the vector field between these two points, we need to first transport $A^\alpha(P)$ to $A^\alpha(Q)$. The first notion of transportation is called "parallel transport".

Video 8: Lie Transport, Killing Vectors, Tensor Densities

What has been shown in the last video is that the partial derivative of a tensor does not produce another tensor, which is what we expect from several different arguments: visually, algebraically and remembering the necessity of the Cristoffel symbols for appropriate derivatives of tensors to account for general curvilinear coordinate bases. A proper way to differentiate the vector field between two close points is to parallel transport one vector to the other and then perform the differentiation as you'd expect. Using the same example as the previous video (section), let's transport the vector from point P to point Q. The parallel transport of the vector is done to preserve its tensorial nature. Lets assume that we can define an object $\mathfrak{T}^{\alpha}_{\beta\mu}$ that allows us to write the transported vector as a difference between the vector at the original point and the infinitesimal distance each component of the vector was moved to reach the next point. We will define this transported vector from P to Q as $A^{\alpha}_{T}(P \to Q)$

$$A^{\alpha}_{T}(P \to Q) = A^{\alpha}(P) - \mathfrak{T}^{\alpha}_{\beta\mu} dx^{\beta} A^{\mu} \tag{135}$$

With the vector "transported", the derivative, which we will label with a capital letter D, is now straightforward.

$$\begin{aligned} \mathfrak{D}_{\beta} A^{\alpha} &\equiv \lim_{dx^{\beta} \to 0} \left\{ \frac{A^{\alpha}(Q) - A^{\alpha}_{T}(P \to Q)}{dx^{\beta}} \right\} \\ &= \partial_{\beta} A^{\alpha} + \mathfrak{T}^{\alpha}_{\beta\mu} A^{\mu} \end{aligned} \tag{136}$$

In this notation $\mathfrak{T}^{\alpha}_{\beta\mu}$, in general, is known as "the connection" term.