1 Flat Spacetime

1.1 Special Relativity

Special relativity, like Newtonian mechanics before it and general relativity after, is a theory for the structure of spacetime. In particular, SR is a theory without gravity, a theory of flat spacetime. Thus it is useful to start here.

Definition. *Spacetime is a four dimensional set, each element is labeled using three space coordinates and one time. An event is an individual point in spacetime.*

In Cartesian coordinates, each event can be represented using (t, x, y, z). In standard Newtonian mechanics, we care about the trajectories of particles: how they move over time. With time included in the tuple, we slightly modify the definition.

Definition. *A worldline is the path of a particle through spacetime, a one dimensional (parameterized) set of events.*

It's clear that time must be treated a bit more carefully than space. While particles can travel however they want (with some limitation as we will soon see) in space, they can only move forward in time. There is some subtlety involved when discussing trajectories in SR versus Newtonian mechanics. In particular, there is an unambiguous notion of *simultaneous events* in Newtonian mechanics whereas there is no such well-defined notion in SR.

In Newtonian mechanics, we have well defined "slices" of time, thus two events are simultaneous in they lie in the same slice of time. As long as two particles only move forward in time, there is no other constraint on their trajectories and, in particular, their velocities.

The only constraint in SR, is the concept of a lightcone. Consider a ray of light, the set of all possible paths it can take forms a cone where the "slope" is equal to the speed of light. The fact that particles cannot move faster than light requires that all worldlines remain inside a light cone.

Consider a standard 2D Cartesian plane. The distance between any two points can be easily compute using Pythagorean's theorem

$$(\Delta s)^2 = (\Delta x)^2 + (\Delta y)^2$$

If we were to rotate the plane so that the new axis are x' and y', we still have

$$(\Delta s)^2 = (\Delta x')^2 + (\Delta y')^2$$

Thus we say that distance is invariant under rotations. We would like to generalize this to include time. Without any justification, consider for now the spacetime interval

$$(\Delta s)^2 = -(c\Delta t)^2 + (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2$$

where c is just some conversion factor between time and space. It turns out that there exists a c (which we will see happens to be the speed of light) such that for any other frame (t', x', y', z'), we still have

$$(\Delta s)^2 = -(c\Delta t')^2 + (\Delta x')^2 + (\Delta y')^2 + (\Delta z')^2$$

This 4D spacetime is known as *Minkowski Space*.

We'll introduce some convenient notation. Spacetime coordinates will be denoted using a Greek subscript (μ, ν) which run from 0 to 3. Instead of a vector, we write an event as

$$x^\mu = (x^0, x^1, x^2, x^3) = (ct, x, y, z)$$

For the sake of simplicity, we will use "god-given units" in which $c = 1$. It is also common convention to use Latin indices for spatial coordinates, thus what we would normally write as a vector would be

$$x^i = (x^1, x^2, x^3) = (x, y, z)$$

To write the spacetime interval using this notation, we introduce a 4x4 matrix: the metric. Note that the previous tuples all are upper indices whereas the metric will carry two lower indices.[1]

$$\eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

The indices allow us to use Einstein's summation notation to write

$$(\Delta s)^2 = \eta_{\mu\nu} \Delta x^\mu \Delta x^\nu$$

A upper-lower index pair implies a sum over all possible values. For instance a dot product can be written

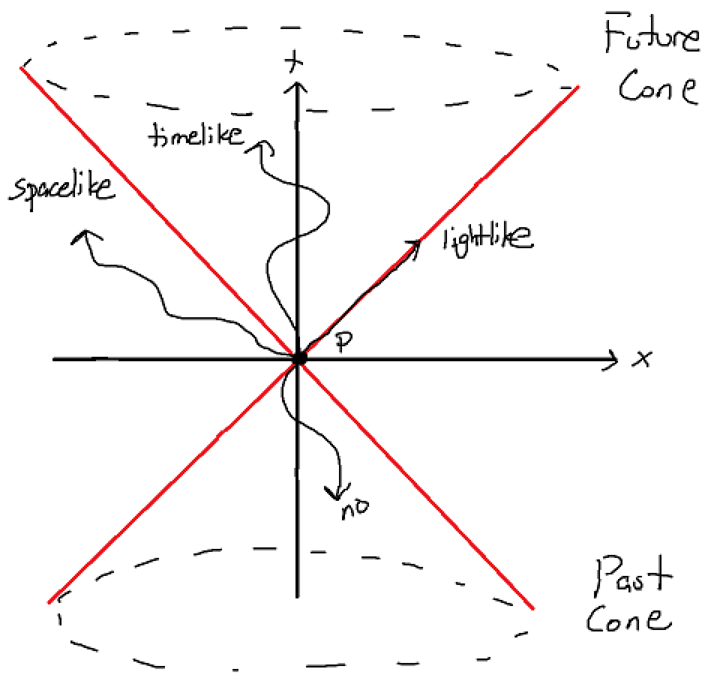
$$x \cdot y = x_i y^i = y_i x^i = \sum_{i=1}^{3} x_i y_i$$

While the choice of an upper or lower index may seem arbitrary, they will come to have meanings later. For now we can imagine an upper index as a column vector and a lower index as a row vector.

When discussing trajectories, it is useful to draw a spacetime diagram. Consider a point p at location $x = 0$ during time $t = 0$. We can draw a lightcone at that point and classify all the points around p:

- Points inside the past/present light cones are *timelike* separated from p
- Points outside the past/present light cones are *spacelike* separated from p
- Points on the past/present light cones are *lightlike* (or null) separated from p

[1] Some texts use $g_{\mu\nu}$ to denote the metric.



We can also classify these based on the spacetime interval

$$(\Delta s)^2 \begin{cases} < 0 & \text{timelike } (|x| < |t|) \\ = 0 & \text{lightlike } (|x| = |t|) \\ > 0 & \text{spacelike } (|x| > |t|) \end{cases}$$

Since (massive) particles travel slower than light, it is slightly unfortunate that the spacetime interval is negative for timelike points. Thus we define *proper time* τ as

$$(\Delta \tau)^2 = -(\Delta s)^2 = -\eta_{\mu\nu} \Delta x^\mu \Delta x^\nu$$

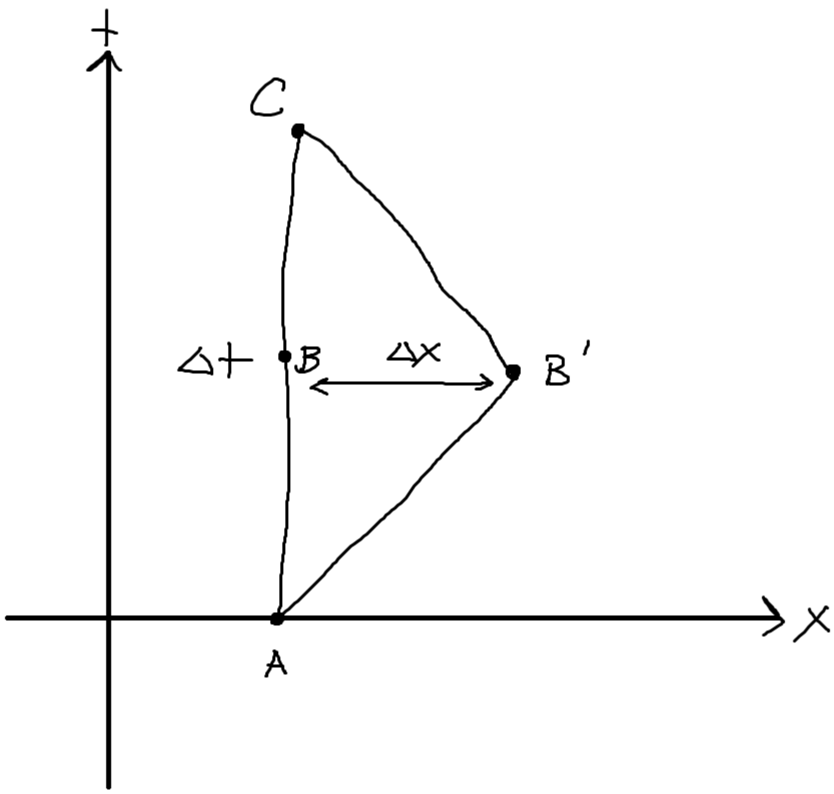
Note that if two events are separated only by time (for instance someone sitting still), then

$$(\Delta \tau)^2 = -\eta_{\mu\nu} \Delta x^\mu \Delta x^\nu = (\Delta t)^2 \implies \Delta \tau = \Delta t$$

Thus proper time is that time elapsed between two events as measured by an observer moving on a straight path between the events. However this is not always equal to coordinate time, which is what we typically interpret as "time elapsed." A common example of this is the twin paradox.

Example. Consider two trajectories undertaken by two twins:

1. $A \to B \to C$ is stationary, where B is half way between A and C
2. $A \to B' \to C$ where the observer travels at some velocity v to the point B' and then back at velocity $-v$.



The midpoints B and B' lie on at the same t on a spacetime diagram.

$$B = \left(\frac{1}{2}\Delta t, 0\right) \qquad B' = \left(\frac{1}{2}\Delta t, \frac{1}{2}v\Delta t\right)$$

We can compute the proper time for both trajectories

$$\Delta\tau_{AB} = \frac{1}{2}\Delta t = \Delta\tau_{BC} \qquad \Delta\tau_{AB'} = \frac{1}{2}\sqrt{1-v^2}\Delta t = \Delta\tau_{B'C}$$

Thus the total time elapsed for the traveler will be

$$\Delta\tau_{AB'C} = \sqrt{1-v^2}\Delta t < \Delta t = \Delta\tau_{ABC}$$

In particular, the twin which stays home will have aged more than the twin that left. This paradox is resolved by noting that the twins are in different inertial frames, or noting that there must be acceleration.[2]

It is important to note that proper time is measured by an observer along the trajectory and thus differs between trajectories. In Newtonian mechanics, there is some universal flow of time throughout all spacetime represented by t. This is not the case in relativity, t is just a coordinate like any other. Note that a straight line may be the shortest distance in space but it results in the longest proper time in spacetime.

[2]See wikipedia page for details.

Of course, not all trajectories will be composed of straight lines. It is often the case that we must integrate, the line element is given by

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu$$

Taking the square root of this is a bit unclear, so instead we parameterize the path as $x^\mu(\lambda)$. It's conventional in Newtonian mechanics to let λ be time, but this is not always the case in relativity. The path integral along a spacelike curve can thus be written

$$\Delta s = \int \sqrt{\eta_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}} d\lambda$$

For timelike paths, we calculate the proper time

$$\Delta \tau = \int \sqrt{-\eta_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}} d\lambda > 0$$

For lightlike paths, the integral will be zero in both forms ($\Delta s = 0$).

Now we formally discuss coordinate transformations of spacetime, in particular those that leave the interval invariant. A simple one is the translation

$$x^\mu \to x^{\mu'} = \delta^{\mu'}_{\mu} (x^\mu + a^\mu)$$

where a^μ are some constants. We use the Kronecker delta so that we can put the prime on the index, rather than on the x. This is to emphasize that the underlying geometrical object is the same, but it is just in a different coordinate system.

The other relevant transformations are spatial rotations and Lorentz boosts, which offset the coordinates by a constant velocity. These are written

$$x^{\mu'} = \Lambda^{\mu'}{}_{\nu} x^\nu$$

In matrix form, this is just $x' = \Lambda x$. Keeping with the matrix forms for everything, the requirement that the interval remain invariant is

$$(\Delta s)^2 = (\Delta x)^T \eta (\Delta x) = (\Delta x)^T \Lambda^T \eta \Lambda (\Delta x) \iff \eta = \Lambda^T \eta \Lambda$$

The set of transformations which satisfy this property form a (non-Abelian) group under matrix multiplication: the Lorentz group. elements of this group are known as Lorentz transformations. The group is typically denoted $O(3, 1)$ to note the relationship to the orthogonal group $O(N)$, the set of NxN matrices which satisfy $R^T R = 1$.

The orthogonal group $O(3)$, the group of spatial rotations satisfies the same condition that Λ does, just with $\eta = 1$ the (3x3) identity. This is just the metric for normal flat space, such a metric in which the eigenvalues are positive is *Euclidean*. A metric with a single minus sign, like η, is *Lorentzian*.

A standard rotation (e.g. in the x-y plane) takes the form

$$\Lambda^{\mu'}{}_{\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & \sin\theta & 0 \\ 0 & -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

where θ is some angle $\in [0, 2\pi)$. A boost can be thought of as a rotation between space and time, for instance a boost in x takes the form

$$\Lambda^{\mu'}{}_{\nu} = \begin{pmatrix} \cosh\phi & -\sinh\phi & 0 & 0 \\ -\sinh\phi & \cosh\phi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

The boost parameter ϕ can be any real number. The three rotations and three boosts generate the Lorentz group, so we see that it has six parameters. If we include translations, then we get the ten parameter Poincare group.

To explicitly see how boosts work, we can explicity write out the transform

$$\begin{aligned} t' &= t\cosh\phi - x\sinh\phi \\ x' &= -t\sinh\phi + x\cosh\phi \\ y' &= y \\ z' &= z \end{aligned}$$

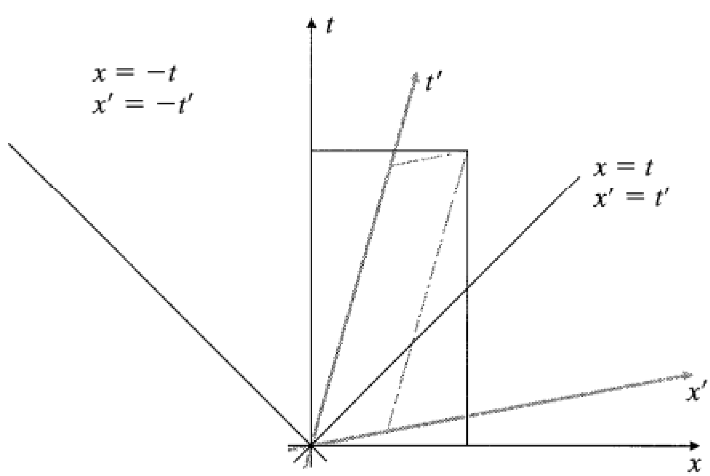
We see that the point at $x' = 0$ is moving with velocity

$$v = \frac{x}{t} = \tanh\phi$$

In particular if we rewrite the transform using v, we see that

$$\begin{aligned} t' &= \gamma(t - vx) \\ x' &= \gamma(x - vt) \end{aligned}$$

where $\gamma = 1/\sqrt{1 - v^2}$. This is the familiar form for a Lorentz transformation. Note that lightcones are unchanged under Lorentz transforms.



1.2 Vectors

To further investigate Minkowski spacetime, we must reintroduce vectors and tensors. Since spacetime is four dimensional, vectors are now *four-vectors*. These vectors are slightly different than normal spatial vectors, for instance we can no longer cross them.

One thing we should emphasize is that each vector is attached to a specific point in spacetime. In the past we have thought of vectors as just direction and magnitude, thus we can slide them along space however we see fit. This is no longer the case once we introduce curvature, so we must re-examine our intuitions.

Definition. *Given a point p of spacetime, the set of all vectors originating from that point is the tangent space at p, denoted T_p.*

Each tangent space will be four dimensional (since p is in Minkowski space). We will denote the basis elements using indices as well $\hat{e}_{(\mu)}$. An arbitrary vector can thus be written

$$A = A^\mu \hat{e}_{(\mu)}$$

and we call A^μ the components of the vector A. In fact we will sometimes refer to A^μ as the vector for convenience, dropping the basis entirely.[3]

As an example, consider a parameterized path $x^\mu(\lambda)$. The tangent vector is

$$V = V^\mu \hat{e}_{(\mu)} = \frac{dx^\mu}{d\lambda}\hat{e}_{(\mu)}$$

Under a Lorentz transform, we see that vector components transforms as

$$V^\mu \to V^{\mu'} = \Lambda^{\mu'}{}_{\nu} V^\nu$$

While the coordinates transform, the actual vector itself should be invariant. This will imply

$$V = V^\mu e_{(\mu)} = V^{\nu'}\hat{e}_{(\nu')} = \Lambda^{\nu'}{}_{\mu} V^\mu \hat{e}_{(\nu')} \Longrightarrow \hat{e}_{(\mu)} = \Lambda^{\nu'}{}_{\mu}\hat{e}_{(\nu')}$$

where $\hat{e}_{(\nu')}$ are the basis vectors in a transformed coordinate system. So to get the new basis elements from the starting point, we must get the inverse of Λ. But the inverse of a Lorentz transform is just another Lorentz transform, so we will denote this by simply swapping the primed and unprimed indices.

$$\Lambda^{\mu}{}_{\nu'}\Lambda^{\nu'}{}_{\rho} = \delta^\mu_\rho \qquad \Lambda^{\sigma'}{}_{\lambda}\Lambda^{\lambda}{}_{\tau'} = \delta^{\sigma'}_{\tau'}$$

Thus the transformation rule for basis vectors is

$$e_{(\nu)'} = \Lambda^{\mu}{}_{\nu'}\hat{e}_{(\mu)}$$

In other words, the basis vectors transform inversely to the components.

[3] The basis vectors have an extra parentheses around the index to indicate that this is a collection of vectors, rather than just components.

Given a vector space, there is the notion of a dual.

Definition. *The dual space to the tangent space T_p, called the cotangent space and denoted T_p^*, is the space of all linear maps $T_p \to \mathbb{R}$. In other words, if $\omega \in T_p^*$, then*

$$\omega(aV + bW) = a\omega(V) + b\omega(W) \in \mathbb{R}$$

The dual space can also have basis vectors, if we specify

$$\hat{\theta}^{(\nu)}(\hat{e}_{(\mu)}) = \delta^\nu_\mu$$

then we will have $\omega = \omega_\mu \hat{\theta}^{(\mu)}$. Just like vectors, we will often use ω_μ to refer to the dual vector. The elements of T_p, which have upper indices, are called contravariant vectors whereas the elements of T_p^*, which have lower indices, are called covariant vectors. Dual vectors are also called one-forms.

Using index notation, we can write the actual of a dual vector on a normal vector as

$$\omega(V) = \omega_\mu \hat{\theta}^{(\mu)}(V^\nu \hat{e}_{(\nu)}) = \omega_\mu V^\nu \hat{\theta}^{(\mu)}(\hat{e}_{(\nu)}) = \omega_\mu V^\nu \delta^\mu_\nu = \omega_\mu V^\mu$$

This is why we often just refer to vectors as just their components. This way of writing the action also seems to imply that vectors can act on dual vectors as

$$V(\omega) = \omega(V) = \omega_\mu V^\mu$$

which tells us that the dual to a dual vector space is just the original vector space.

Just as with before, we can derive transformation rules for components and basis dual vectors

$$\omega_{\mu'} = \Lambda^\nu{}_{\mu'} \omega_\nu \qquad \hat{\theta}^{(\rho')} = \Lambda^{\rho'}{}_\sigma \hat{\theta}^{(\sigma)}$$

A classic example of a dual vector is the gradient, which we can write as

$$\nabla\phi = \frac{\partial\phi}{\partial x^\mu}\hat{\theta}^{(\mu)}$$

It may seem strange to only see upper indices in this formula, but when it is in the denominator of a derivative, the "lower-ness" is implied. This will transform as

$$\frac{\partial\phi}{\partial x^{\mu'}} = \frac{\partial x^\mu}{\partial x^{\mu'}}\frac{\partial\phi}{\partial x^\mu} = \Lambda^\mu{}_{\mu'}\frac{\partial\phi}{\partial x^\mu}$$

Since the gradient is a dual vector, we will usually adopt the shorthand for partial derivatives

$$\frac{\partial\phi}{\partial x^\mu} = \partial_\mu\phi$$

This way there is no confusion about denominators and lower indices.

1.3 Tensors

A tensor is basically way to combine vectors and dual vectors, generalizing them both. Recall that a dual vector is a map from a vector to $\mathbb{R}$ and vice versa.

Definition. *A tensor T of type (or rank) (k, l) is a multilinear map*

$$T : \underbrace{T_p^* \times \cdots \times T_p^*}_{k \text{ times}} \times \underbrace{T_p \times \cdots \times T_p}_{l \text{ times}} \to \mathbb{R} \Longrightarrow T : T_p^{*k} \times T_p^l \to \mathbb{R}$$

A vector is a rank $(1, 0)$ tensor, a dual vector is a rank $(0, 1)$ tensor. Scalars are rank $(0, 0)$ and a matrix would be rank $(1, 1)$. For a fixed type (k, l), the set of all tensors form a vector space, so it is natural to look for a basis. This requires defining the notion of a tensor product.

Definition. *Let T be a (k, l) tensor and S a (m, n) tensor. The tensor product $T \otimes S$ is a $(k + m, l + n)$ type tensor which acts as*

$$\begin{aligned} T \otimes S(\omega^{(1)}, \ldots, \omega^{(k+m)}, V^{(1)}, \ldots, V^{(l+n)}) \\ = T(\omega^{(1)}, \ldots, \omega^{(k)}, V^{(1)}, \ldots, V^{(l)}) \\ \times S(\omega^{(k+1)}, \ldots, \omega^{(k+m)}, V^{(l+1)}, \ldots, V^{(l+n)}) \end{aligned}$$

In other words, a tensor product is just the product of T and S acting on the appropriate arguments. The basis elements will be of the form

$$\hat{e}_{(\mu_1)} \otimes \cdots \otimes \hat{e}_{(\mu_k)} \otimes \hat{\theta}^{(\nu_1)} \otimes \cdots \otimes \hat{\theta}^{(\nu_l)}$$

so we see that Minkowski space will have 4^{k+l} basis tensors for the space of rank (k, l) tensors. Note that in general $S \otimes T \neq T \otimes S$. Any arbitrary tensor can therefore be written as

$$T = T^{\mu_1 \cdots \mu_k}{}_{\nu_1 \cdots \nu_l} \, \hat{e}_{(\mu_1)} \otimes \cdots \otimes \hat{e}_{(\mu_k)} \otimes \hat{\theta}^{(\nu_1)} \otimes \cdots \otimes \hat{\theta}^{(\nu_l)}$$

Just like before, we will mostly refer to the components as *the* tensor. If we don't know the components, then we can recover them by acting on the basis vectors

$$T^{\mu_1 \cdots \mu_k}{}_{\nu_1 \cdots \nu_l} = T(\hat{\theta}^{(\mu_1)}, \ldots, \hat{\theta}^{(\mu_k)}, \hat{e}_{(\nu_1)}, \ldots, e_{(\nu_l)})$$

The order of the indices is important as the tensor may act on different arguments differently.

The transformation law for tensors is exactly what we expect it to be

$$T^{\mu_1' \cdots \mu_k'}{}_{\nu_1' \cdots \nu_l'} = \Lambda^{\mu_1'}{}_{\mu_1} \cdots \Lambda^{\mu_k'}{}_{\mu_k} \Lambda^{\nu_1}{}_{\nu_1'} \cdots \Lambda^{\nu_l}{}_{\nu_l'} T^{\mu_1 \cdots \mu_k}{}_{\nu_1 \cdots \nu_l}$$

While a tensor maps a collection of vectors and dual vectors to $\mathbb{R}$, there is nothing stopping us from only partially filling the arguments. For instance we can interpret a $(1, 1)$ tensor (which would normally map $T_p^* \times T_p \to \mathbb{R}$) as a map

$$T^\mu{}_\nu : V^\nu \mapsto T^\mu{}_\nu V^\nu$$

which is map from vectors to vectors $T_p \to T_p$. We can also act on tensors with other tensors,

$$U^\mu{}_\nu = T^{\mu\rho}{}_\sigma S^\sigma{}_{\rho\nu}$$

would be a $(1, 1)$ tensor.

Example. We've actually seen an example of a tensor already. The metric (which is actually shorthand for metric tensor) is a type $(0,2)$ tensor. The action given by the metric defines an inner product on spacetime

$$V \cdot W = \eta(V,W) = \eta_{\mu\nu} V^\mu W^\nu$$

The inner product of a vector with itself defines the norm. We classify vectors based on this number

$$\eta_{\mu\nu} V^\mu V^\nu \begin{cases} < 0 & \text{timelike} \\ = 0 & \text{lightlike} \\ > 0 & \text{spacelike} \end{cases}$$

This is terminology used previously, indeed the two definitions should be equal

$$\eta_{\mu\nu} V^\mu V^\nu = -t^2 + x^2 + y^2 + z^2 = s^2$$

Related is the Kronecker delta δ^μ_ν, which is just the identity map. Note that we don't space the indices because there is no difference between them

$$\delta^\mu{}_\nu = \delta_\nu{}^\mu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

The Kronecker delta allows us to define the inverse metric, a $(2,0)$ tensor $\eta^{\mu\nu}$ which satisfies

$$\eta^{\mu\nu}\eta_{\nu\rho} = \eta_{\rho\nu}\eta^{\nu\mu} = \delta^\mu_\rho$$

In fact the inverse metric has the same components of the normal metric, but this will only be the case in flat spacetime with Cartesian coordinates.

An interesting example is the Levi-Civita symbol (also called a tensor) which is useful in combinatorics and quantum mechanics.[4]

$$\tilde{\epsilon}_{\mu\nu\rho\sigma} = \begin{cases} +1 & \text{if } \mu\nu\rho\sigma \text{ is an even number of swaps from 0123} \\ -1 & \text{if } \mu\nu\rho\sigma \text{ is an odd number of swaps from 0123} \\ 0 & \text{otherwise} \end{cases}$$

A classic tensor is the electromagnetic field tensor. The electric field components E_i and magnetic field components B_i are components of a $(0,2)$ tensor

$$F_{\mu\nu} = \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & B_3 & -B_2 \\ E_2 & -B_3 & 0 & B_1 \\ E_3 & B_2 & -B_1 & 0 \end{pmatrix} = -F_{\nu\mu}$$

Note that swapping the lower indices transposes the matrix (swaps rows and columns).

[4] The tilde on top is sometimes omitted.

We can now discuss some operations that can be performed on tensors.

Definition. *Contraction is the operation of summing over one upper and one lower index to turn a $(k.l)$ tensor into a $(k-1, l-1)$ tensor. For instance*

$$S^{\mu\rho}{}_{\sigma} = T^{\mu\nu\rho}{}_{\sigma\nu}$$

Since the order of the indices matter, contracting in different ways will yield different tensors

$$T^{\mu\nu\rho}{}_{\sigma\nu} \neq T^{\mu\rho\nu}{}_{\sigma\nu}$$

The metric tensor and it's inverse can be used to raise/lower indices. One application is turn vectors into dual vectors and vice versa

$$V_\mu = \eta_{\mu\nu} V^\nu \qquad \omega^\mu = \eta^{\mu\nu}\omega_\nu$$

For a $(2,2)$ tensor, we can do something like

$$T_\mu{}^{\beta\nu}{}_\delta = \eta_{\mu\alpha}\eta^{\nu\gamma} T^{\alpha\beta}{}_{\gamma\delta}$$

We can also raise and lower indices being contracted by using both the metric and its inverse

$$A^\lambda B_\lambda = \eta^{\lambda\mu} A_\mu \eta_{\lambda\nu} B^\nu = \delta^\mu_\nu A_\mu B^\nu = A_\mu B^\mu$$

Definition. *A tensor is symmetric in its indices if it is invariant under exchange. For instance, a tensor is symmetric in its first two indices if*

$$S_{\mu\nu\rho} = S_{\nu\mu\rho}$$

while a totally symmetric tensor is symmetric in all indices

$$S_{\mu\nu\rho} = S_{\mu\rho\nu} = S_{\nu\mu\rho} = S_{\nu\rho\mu} = S_{\rho\mu\nu} = S_{\rho\nu\mu}$$

Similarly, a tensor is antisymmetric (or totally antisymmetric) if swapping (any) two indices introduces a negative sign. For instance, a tensor is antisymmetric in its first two indices if

$$A_{\mu\nu\rho} = -A_{\nu\mu\rho}$$

The metric and its inverse are (totally) symmetric, while the Levi-Civita symbol and the electromagnetic field strength tensor is totally antisymmetric.

Definition. *Any tensor can be symmetrized over some indices using the formula*

$$T_{[\mu_1 \cdots \mu_n]\rho}{}^{\sigma} = \frac{1}{n!} \sum \textit{permutations of } \mu_1 \cdots \mu_k$$

Similarly they can be antisymmetrized using an alternating sum[5]

$$T_{[\mu_1 \cdots \mu_n]\rho}{}^{\sigma} = \frac{1}{n!} \sum \textit{alternating sum of permutations of } \mu_1 \cdots \mu_k$$

[5] Sometimes the $1/n!$ factor is omitted in both formulas

The alternating sum adds a negative sign to all odd permutations of the given indices, e.g.

$$T_{[\mu\nu\rho]\sigma} = \frac{1}{6}(T_{\mu\nu\rho\sigma} - T_{\nu\mu\rho\sigma} + T_{\nu\rho\mu\sigma} - T_{\mu\rho\nu\sigma} + T_{\rho\mu\nu\sigma} - T_{\rho\nu\mu\sigma})$$

We can also leave out indices by surrounding them with vertical bars

$$T_{(\mu|\nu|\rho)} = \frac{1}{2}(T_{\mu\nu\rho} + T_{\rho\nu\mu})$$

(Anti)symmetrization has the unique property that contractions only interact with the (anti)symmetric parts, for instance in the case of a $(0,2)$ (or $(2,0)$) tensor

$$X^{(\mu\nu)}Y_{\mu\nu} = \frac{1}{2}(X^{\mu\nu} + X^{\mu\nu})Y_{\mu\nu} = \frac{1}{4}(X^{\mu\nu}Y_{\mu\nu} + X^{\mu\nu}Y_{\nu\mu} + X^{\mu\nu}Y_{\mu\nu} + X^{\mu\nu}Y_{\nu\mu}) = X^{(\mu\nu)}Y_{(\mu\nu)}$$

Another neat trick is that we an decompose a tensor in a sum of a symmetric and antisymmetric tensor given any two indices.

$$T_{\mu\nu\rho\sigma} = T_{(\mu\nu)\rho\sigma} + T_{[\mu\nu]\rho\sigma}$$

However this construction break down for three or more indices.

Definition. *The trace of a* $(1,1)$ *tensor* $X^{\mu}{}_{\nu}$ *is the contraction*

$$X = X^{\mu}{}_{\mu}$$

Since a $(1,1)$ tensor is really just a matrix, the trace is just the normal matrix trace. We can also take the trace of $(2,0)$ and $(0,2)$ tensors, but we must raise/lower an index first. For instance

$$Y = Y^{\mu}{}_{\mu} = \eta^{\mu\nu}Y_{\mu\nu}$$

Example. The trace of the metric is

$$\eta = \eta^{\mu\nu}\eta_{\mu\nu} = \delta^{\mu}_{\mu} = 4$$

In n dimensions, $\delta^{\mu}_{\mu} = n$. Antisymmetric $(0,2)$ tensors are always traceless.

The partial derivative can be defined on a (k,l) tensor which yields a $(k, l+1)$ tensor

$$T_{\alpha}{}^{\mu}{}_{\nu} = \partial_{\alpha}R^{\mu}{}_{\nu}$$

But this only transforms properly in flat spacetime. In general spacetime, we will have to replace this partial derivative is a more general covariant derivative.

1.4 Classical Field Theory

We'll now apply tensors to some actual physics. Recall Maxwell's equations

$$\begin{gathered}\nabla \times \mathbf{B} - \partial_t \mathbf{E} = \mathbf{J} \\ \nabla \cdot \mathbf{E} = \rho \\ \nabla \times \mathbf{E} + \partial_t \mathbf{B} = 0 \\ \nabla \cdot \mathbf{B} = 0\end{gathered}$$

where $\mathbf{E}$ and $\mathbf{B}$ are the electric and magnetic field vectors respectively, $\mathbf{J}$ is electric current, and ρ is the charge density. To see that these equations are actually Lorentz invariant, we will write then using tensors. First note that we can write the cross product as

$$\mathbf{B} = \nabla \times \mathbf{A} \iff B^i = \tilde{\epsilon}^{ijk} \partial_j A_k$$

where we use the 3-vector version of the Levi-Civita symbol. Maxwell's equation can now be written

$$\begin{gathered}\tilde{\epsilon}^{ijk} \partial_j B_k - \partial_0 E^i = J^i \\ \partial_i E^i = J^0 \\ \tilde{\epsilon}^{ijk} \partial_j E_k + \partial_0 B^i = 0 \\ \partial_i B^i = 0\end{gathered}$$

We combine the current and charge density into a single 4-vector by setting $J^0 = \rho$. We introduced the electromagnetic field strength tensor earlier, we can rewrite it with upper indices using

$$F^{0i} = E^i \qquad F^{ij} = \tilde{\epsilon}_{ijk} B_k$$

So we can rewrite all the equations in terms of the EM field strength tensor

$$\begin{gathered}\partial_j F^{ij} - \partial_0 F^{0i} = J^i \\ \partial_i F^{0i} = J^0\end{gathered}$$

Since $F^{\mu\nu}$ is antisymmetric and $F^{00} = 0$, we can combine the first two equations

$$\begin{gathered}\partial_j F^{ij} + \partial_0 F^{i0} = J^i \qquad \partial_j F^{0j} + \partial_0 F^{00} = J^0 \\ \therefore \partial_\nu F^{\mu\nu} = J^\mu\end{gathered}$$

The remaining two equations can be expressed as[6]

$$\partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} + \partial_\lambda F_{\mu\nu} = \partial_{[\mu} F_{\nu\lambda]} = 0$$

Thus we see that the four original Maxwell's equations can be re-expressed into two Lorentz invariant equations using tensors.

[6] The process involves calculating the dual tensor $G_{\mu\nu}$ using the Levi-Civita symbol, then repeating the previous construction. Returning back to $F_{\mu\nu}$ gives an alternating sum and we exploit anti-symmetry.

Next, we will re-examine how physics works in Minkowski spacetime. A worldline can be interpreted as a map $\mathbb{R} \to M$ where M is the manifold representing spacetime. We can imagine this as a parameterized curve $x^\mu(\lambda)$, for massive particles the tangent vector at any point will always be timelike.

Definition. *For a massive particle, we can compute the proper time $\tau(\lambda)$ and invert (if λ is a good parameter) to get $\lambda(\tau)$. Then the 4-velocity is the tangent vector*

$$U^\mu = \frac{dx^\mu}{d\tau}$$

Note that the 4-velocity will be normalized by definition

$$d\tau^2 = -\eta_{\mu\nu} dx^\mu dx^\nu \implies \eta_{\mu\nu} U^\mu U^\nu = -1$$

While spatial 3-vectors can take on different magnitudes (speeds), we see that the 4-vector will only every has a speed of -1, which is the speed at which we travel through spacetime.

Definition. *The 4-momentum is*

$$p^\mu = mU^\mu$$

where m is the (rest) mass of the particle. The energy is the first component

$$E = p^0$$

In the rest frame we will have

$$E = p^0 = m \iff E = mc^2$$

For other frames, we can Lorentz boost the rest momentum. Consider a particle traveling with velocity v along the x axis. The 4-momentum is

$$p^\mu = (\gamma m, v\gamma m, 0, 0) \qquad \gamma = \frac{1}{\sqrt{1 - v^2}}$$

Recalling the the 4-velocities are normalized,

$$p^\mu p_\nu = -m^2 = E^2 + \mathbf{p}^2 \iff E^2 = m^2 + \mathbf{p}^2$$

Now that we have a way to describe the energy and momentum of individual particles, we need a way to describe systems of many particles. The best way to do so is to describe the system as a fluid, characterized by macroscopic quantities such as density and viscosity. A single 4-momentum field is not enough to describe a fluid, instead we must use a tensor.

Definition. *The energy-momentum tensor is a $(0, 2)$ tensor $T^{\mu\nu}$ which describes the flux of a 4-momentum p^μ across a surface x.*

The definition is not particularly illuminating, but we can see a few things from it. For instance,

$$T^{00} = \rho \qquad T^{ii} = p_i$$

where ρ is the (rest-frame) energy density and p_i are the spatial pressures. The spatial components T^{ij} in general give the stress and the off diagonal components of those represent shear.

Example. To make this more concrete, consider a collection of particles at rest which we will call dust. Clearly the 4-velocity field $U^\mu(x)$ is going to be constant. Define the *number-flux 4-vector*

$$N^\mu = nU^\mu$$

where n is the number density[7] of the particles measured in the rest frame. N^0 is just the number density and N^i is the flux of particles in the x^i direction. We can express the energy density as

$$\rho = mn$$

This completely specifies the dust in the rest frame. But note that m and n are actually 0-components of two 4-vectors

$$N^\mu = (n, 0, 0, 0) \qquad p^\mu = (m, 0, 0, 0)$$

So ρ is actually the 00 component of the tensor product $p \otimes N$. If we wanted to describe the dust in other frames, we are inclined to define the energy-momentum tensor as

$$T^{\mu\nu} = p^\mu N^\nu = mnU^\mu U^\nu = \rho U^\mu U^\nu$$

Dust is an idealized approximation of most fluids and we can do better. We can generalize to the concept of *perfect fluids*, which are completely specified by two quantities: a rest-frame energy density ρ and and an isotropic (same in all directions) rest-frame pressure p.

An consequence of isotropy is that the energy-momentum tensor will be diagonal in its rest frame. Furthermore all the spatial components are equal, so we may guess that

$$T^{\mu\nu} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{pmatrix}$$

By comparing this to the dust energy-momentum tensor, we see that a covariant general form would be

$$T^{\mu\nu} = (\rho + p)U^\mu U^\nu + p\eta^{\mu\nu}$$

A neat property of the energy-momentum tensor is that it is symmetric, furthermore it is conserved in the sense that

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

If we want to make the jump to general relativity, we will need a feel for field theories. We'll start by considering a classical field theory on flat spacetime. Consider a particle in one dimension with coordinate $q(t)$. The equations of motion can be derived using the principle of least action.

[7] Number of particles per unit area

Definition. *The action of a trajectory $q(t)$ is the integral*

$$S = \int dt\, L(q, \dot{q})$$

where $L(q, \cdot q)$ is the Lagrangian, typically

$$L = K - V$$

for a point-particle where K is kinetic energy and V the potential energy.

Using calculus-of-variations, it can be shown that the trajectory which minimizes action will solve the *Euler-Lagrange equations*

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

When we move to field theory, the coordinate $q(t)$ becomes a spacetime dependent set of fields $\Phi^i(x^\mu)$. The Lagrangian is replaced with a *Lagrange density* which will satisfy

$$L = \int d^3x\, \mathcal{L}(\Phi^i, \partial_\mu \Phi^i)$$

and the action is now

$$S = \int dt\, L = \int d^4x\, \mathcal{L}(\Phi^i, \partial_\mu \Phi^i)$$

In field theory, it is common to use "natural units" in which $\hbar = c = k = 1$. We can repeat the variation procedure to get equations of motion for the field

$$\frac{\delta S}{\delta \Phi^i} = \frac{\partial \mathcal{L}}{\partial \Phi^i} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \Phi^i)}\right) = 0$$

Example. The simplest example is that of a single real scalar field

$$\phi(x^\mu) : M \to \mathbb{R}$$

There are three contributions to the energy density: the kinetic energy $\frac{1}{2}\dot{\phi}^2$, the potential energy $V(\phi)$, and a third contribution called the gradient energy $\frac{1}{2}(\boldsymbol{\nabla}\phi)^2$. We combine the kinetic and gradient energies into a Lorentz invariant form

$$\frac{1}{2}\dot{\phi}^2 - \frac{1}{2}(\boldsymbol{\nabla}\phi)^2 = -\frac{1}{2}\eta^{\mu\nu}(\partial_\mu\phi)(\partial_\nu\phi) = (\partial\phi)^2$$

So a reasonable Lagrangian (density) would be

$$\mathcal{L} = (\partial\phi)^2 - V(\phi)$$

To calculate the derivatives required for the equations of motion, we will use the general rule for indices on derivatives

$$\frac{\partial V_\alpha}{\partial V_\beta} = \delta_\alpha^\beta$$

Using this, we relabel the indices and calculate

$$\frac{\partial}{\partial(\partial_\mu \phi)}[\eta^{\rho\sigma}(\partial_\rho \phi)(\partial_\sigma \phi)] = \eta^{\rho\sigma}[\delta^\mu_\rho(\partial_\sigma\phi) + (\partial_\rho\phi)\delta^\mu_\sigma]$$
$$= \eta^{\mu\sigma}(\partial_\sigma\phi) + \eta^{\rho\mu}(\partial_\rho\phi) = 2\eta^{\mu\nu}\phi$$

Thus the derivatives are

$$\frac{\partial\mathcal{L}}{\partial\phi} = -\frac{dV}{d\phi} \qquad \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)} = -\eta^{\mu\nu}\partial_\nu\phi$$

Putting these together gives us the equation of motion for a real scalar field

$$\eta^{\mu\nu}\partial_\mu\partial_\nu\phi - \frac{dV}{d\phi} = \Box\phi - \frac{dV}{d\phi} = 0$$

where we define $\Box = \eta^{\mu\nu}\partial_\mu\partial_\nu = \partial_\mu\partial^\mu$ as the d'Alembertian. There will be a direct procedure introduce to calculate the energy-momentum tensor (Chapter 4), but we will quote the result here

$$T^{\mu\nu}_{\text{scalar}} = \eta^{\mu\lambda}\eta^{\nu\sigma}(\partial_\lambda\phi)(\partial_\sigma\phi) - \eta^{\mu\nu}\left(\frac{1}{2}\eta^{\lambda\sigma}(\partial_\lambda\phi)(\partial_\sigma\phi) + V(\phi)\right)$$

Example. A slightly more difficult example would be electromagnetism. The field components are actually the vector potentials

$$\Phi^i = \{A_0, A_1, A_2, A_3\} = \{\phi, A_x, A_y, A_z\}$$

The field strength tensor is related to the vector potential (field components) by

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$$

From this we see that the tensor is gauge invariant, that is if we perform a gauge transformation

$$A_\mu \to A_\mu + \partial_\mu\lambda(x)$$

then there will be no change in the field strength tensor

$$F_{\mu\nu} \to F_{\mu\nu} + \partial_\mu\partial_\nu\lambda - \partial_\nu\partial_\mu\lambda = F_{\mu\nu}$$

All observable electromagnetic quantities are gauge invariant, so while the actual field of uses A_μ, physical quantities are typically expressed using $F_{\mu\nu}$. Drawing some inspiration from the covariant Maxwell's equations, we will atke the Lagrangian to be

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + A_\mu J^\mu$$

Just as with before, we must relabel indices and do some extra work to obtain our derivatives. Once that's all done, we are left with

$$\frac{\partial\mathcal{L}}{\partial A_\nu} = J^\nu \qquad \frac{\partial\mathcal{L}}{\partial(\partial_\mu A_\nu)} = -F^{\mu\nu}$$

and our equation of motion forms the first two Maxwell's equations

$$\partial_\mu F^{\nu\mu} = J^\nu$$

The energy-momentum tensor for electromagnetism is

$$T^{\mu\nu}_{\text{EM}} = F^{\mu\lambda}F^\nu{}_\lambda - \frac{1}{4}\eta^{\mu\nu}F^{\lambda\sigma}F_{\lambda\sigma}$$

1.5 Exercises

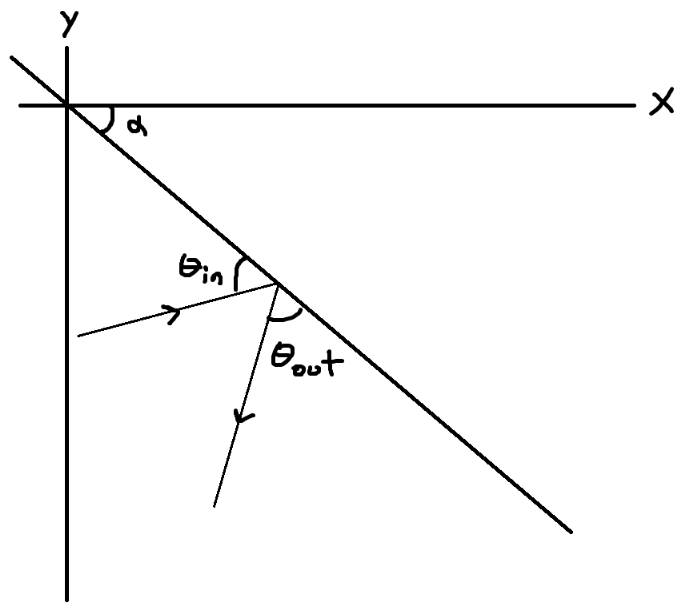
Exercise 1. The transform for a boost along y is

$$\begin{cases} t' &= \gamma(t - vy) \\ x' &= x \\ y' &= \gamma(y - vt) \\ z' &= z \end{cases}$$

where $\gamma = 1/\sqrt{1-v^2}$ is the Lorentz factor. Then the wall can be expressed in the S frame as

$$x' = -y' \iff x = -\gamma(y - vt)$$

So the angle between the wall and the horizontal is $\alpha = \tan^{-1}(1/\gamma) = \tan^{-1}(\sqrt{1-v^2})$. Suppose the ball has velocity w and hits the wall at an angle θ as in the picture below.



We can solve for the angle the wall's velocity makes with the horizontal. The components are of the ball's velocity is

$$w_x = w\cos\left(\frac{\pi}{2} - \alpha + \theta_{in}\right)$$
$$w_y = w\sin\left(\frac{\pi}{2} - \alpha + \theta_{in}\right)$$

In the S' frame, these velocity components are

$$\begin{aligned} w'_x = \frac{dx'}{dt'} &= \frac{dx}{\gamma(dt - v\,dy)} \\ &= \frac{dx/dt}{\gamma(1 - v\,dy/dt)} = \frac{w_x}{\gamma(1 - vw_y)} \end{aligned}$$

$$\begin{aligned} w'_y = \frac{dy'}{dt'} &= \frac{\gamma(dy - v\,dt)}{\gamma(dt - v\,dy)} \\ &= \frac{dy/dt - v}{1 - v\,dy/dt} = \frac{w_y - v}{1 - vw_y} \end{aligned}$$

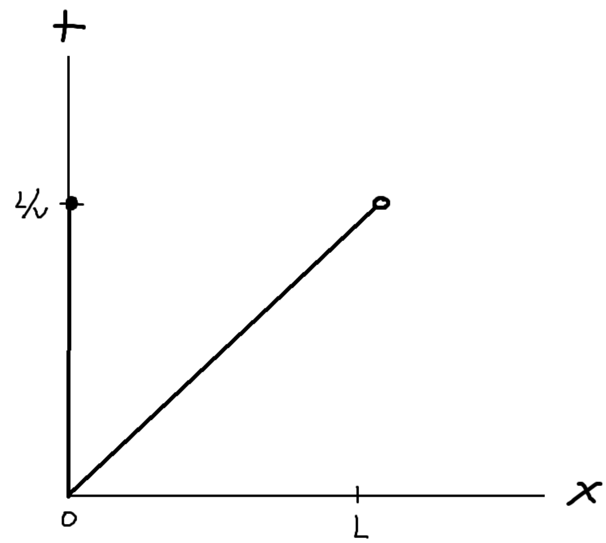
When we reflect over $x' = -y'$, we get the new components as $u'_x = -w'_y$ and $u'_y = -w'_x$ then convert back into the S frame by boosting backwards

$$\begin{aligned} u_x &= \frac{u'_x}{\gamma(1 + vu'_y)} = \frac{v - w_y}{\gamma(1 - vw_y) - vw_x} \\ u_y &= \frac{u'_y + v}{1 + vu'_y} = \frac{\gamma v(1 - vw_y) - w_x}{\gamma(1 - vw_y) - vw_x} \end{aligned}$$

Now we can solve for the new magnitude and reflected magnitude

$$\begin{aligned} \theta_{out} &= \tan^{-1}\left(\frac{u_y}{u_x}\right) - \alpha \\ |u| &= \sqrt{u_x^2 + u_y^2} \end{aligned}$$

Exercise 2. The worldlines look something like

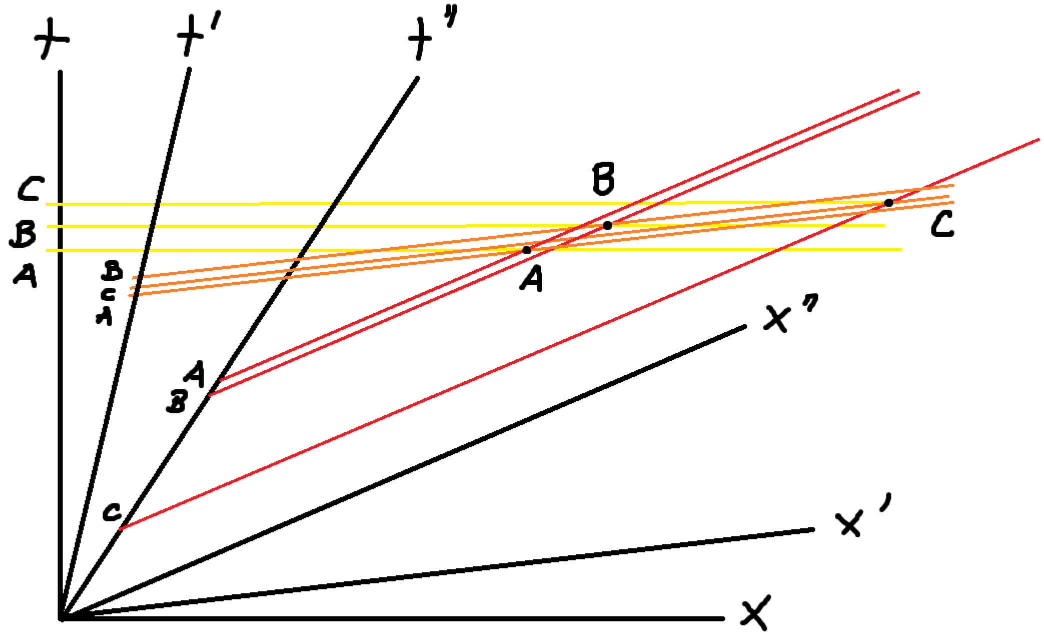


A better way to visualize would be to imagine the x axis as a circle so that the $x - t$ plane is actually a cylinder. The stationary observer has a vertical worldline while the traveling observer has a helix. We can clearly see that for the traveling observer

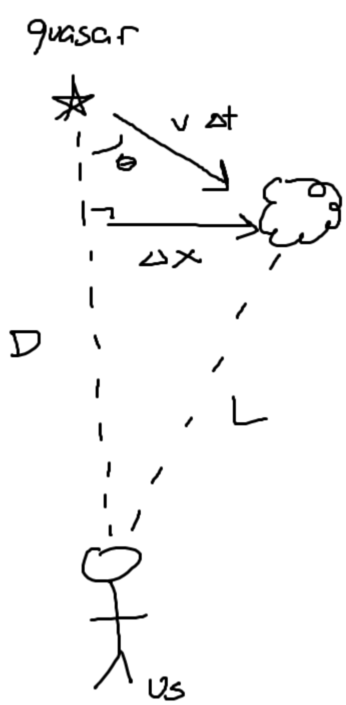
$$\Delta\tau = \sqrt{\frac{L^2}{v^2} - L^2} = \frac{L}{v}\sqrt{1 - v^2} < \frac{L}{v}$$

since $v < 1$ (recall that we defined $c = 1$). So the traveling observer will still experience less proper time than the stationary one, as expected. However, this results in a paradox where someone can go in a circle and end up younger. We can explain this as resulting from the fact that the universe is periodic.

Exercise 3. Suppose observer O see three events in the order ABC and another observer seeing them in the order CBA. It is indeed possible for a third observer to see ACB, which can be seen by carefully choosing a boost.



Exercise 4. Honestly I had no idea what I was doing until I found some hints and a helpful picture online. Here is an artistic reinterpretation.



Recalling that $c = 1$, the time at which we see the cloud first be omitted is

$$t_0 = D$$

After some time dt has passed, we observe the cloud at its new position at time

$$t_1 = L + dt$$

We can calculate the perpendicular (perceived) distance traveled to get

$$v_{app} = \frac{\Delta x}{\Delta t} = \frac{v \sin\theta \, dt}{L + dt - D}$$

Using some basic trig, we can calculate L

$$L^2 = (D - v\cos\theta \, dt)^2 + (vT\sin\theta)^2 = (v\,dt)^2 + D^2 - 2v\cos\theta \, dt$$

Since quasar's are quite far away, we can assume $D \gg v\,dt$ and expand to first order

$$L = D\sqrt{1 + \left(\frac{v\,dt}{D}\right)^2 - \frac{2v\,dt}{D}\cos\theta} \approx D - v\cos\theta \, dt$$

Thus the apparent velocity of the cloud to us is

$$v_{app} = \frac{v\sin\theta \, dt}{dt - v\cos\theta \, dt} = \frac{v\sin\theta}{1 - v\cos\theta}$$

This can very easily be greater than 1, for instance take $\theta = \pi/4$

$$v_{app} > 1 \implies v > \frac{1}{\sin\theta + \cos\theta} = 0.71$$

Exercise 5. I don't believe this is covered in the text, but relativistic kinematics is a bit different, we must modify our equations to account for relativistic effects

$$m = \gamma m_0 \qquad p = mv = \gamma m_0 v$$

where m_0 is the rest mass and $\gamma = 1/\sqrt{1 - v^2}$ is the Lorentz factor. We can also rewrite

$$p = \gamma m_0 v = \gamma m_0 \sqrt{\frac{\gamma^2 - 1}{\gamma^2}} \iff \gamma = \sqrt{\frac{p^2 + m_0^2}{m_0^2}} = \frac{E}{m_0}$$

The rest mass and energy are given to us

$$\gamma = \frac{1000 \text{ GeV}}{0.106 \text{ Gev}} = 9433.96$$

If we imagine the particle as traveling in a straight line rather than a circle, we see that the proper time is related to the experimentalist's time by

$$\Delta\tau = \sqrt{1 - v^2}\Delta t \iff \frac{\Delta t}{\Delta\tau} = \gamma$$

Thus from the experimentalist's point of view, we will measure a lifetime of

$$\Delta t = \gamma \Delta \tau = 9433.96 \cdot 2.19 \cdot 10^{-6} \approx 0.021 \text{ s}$$

or about 21 milliseconds. During this lifetime it will travel[8]

$$\Delta\theta = \frac{vc\Delta t}{r} = \frac{0.021c}{500}\sqrt{1 - \frac{1}{9433.96^2}} = 12591.3 \text{ rad}$$

which is roughly 2004 full revolutions.

Exercise 6. A) The gradients of note are

$$\begin{aligned}
\boldsymbol{\nabla} x(\lambda) &= (1, 2\lambda - 2, -1) \\
\boldsymbol{\nabla} x(\mu) &= (-\sin\mu, \cos\mu, 1) \\
\boldsymbol{\nabla} x(\sigma) &= (2\sigma, 3\sigma^2 + 2\sigma, 1)
\end{aligned}$$

p occurs at $\lambda = 1, \mu = 0, \sigma = -1$ so the tangent vectors are

$$\begin{aligned}
\boldsymbol{\nabla} x(1) &= (1, 0, -1) \\
\boldsymbol{\nabla} x(0) &= (0, 1, 1) \\
\boldsymbol{\nabla} x(-1) &= (-2, 1, 1)
\end{aligned}$$

B) f depends on the path taken

$$\begin{aligned}
\frac{df}{d\lambda} &= 4\lambda^3 - 9\lambda^2 + 10\lambda - 3 \\
\frac{df}{d\mu} &= (1-\mu)\cos\mu - \sin\mu \\
\frac{df}{d\sigma} &= 6\sigma^5 + 10\sigma^4 + 4\sigma^3 - 3\sigma^2
\end{aligned}$$

Exercise 7. Consider the following tensor and 4-vector

$$X^{\mu\nu} = \begin{pmatrix} 2 & 0 & 1 & -1 \\ -1 & 0 & 3 & 2 \\ -1 & 1 & 0 & 0 \\ -2 & 1 & 1 & -2 \end{pmatrix} \qquad V^\mu = (-1, 2, 0, -2)$$

AB) We use the metric to raise/lower indices[9][10]

$$X^\mu{}_\nu = X^{\mu\lambda}\eta_{\lambda\nu} = \begin{pmatrix} -2 & 0 & 1 & -1 \\ 1 & 0 & 3 & 2 \\ 1 & 1 & 0 & 0 \\ 2 & 1 & 1 & -2 \end{pmatrix} \qquad X^\mu{}_\nu = \eta_{\mu\lambda}X^{\lambda\nu} = \begin{pmatrix} -2 & 0 & -1 & 1 \\ -1 & 0 & 3 & 2 \\ -1 & 1 & 0 & 0 \\ -2 & 1 & 1 & -2 \end{pmatrix}$$

[8] Note that we defined $c = 1$, so we should account for that here

[9] An easy way to interpret contraction is as matrix multiplication, but be careful that the order is correct.

[10] Of course we can do all these problems by explicitly writing out the implied sums, but that is a bit tedious, so the matrix multiplication shortcut helps.

CD) We can symmetrize and antisymmetrize $X^{\mu\nu}$

$$X^{(\mu\nu)} = \frac{1}{2}(X^{\mu\nu} + X^{\nu\mu}) = \frac{1}{2}\begin{pmatrix} 4 & -1 & 0 & -3 \\ -1 & 0 & 4 & 3 \\ 0 & 4 & 0 & 1 \\ -3 & 3 & 1 & -4 \end{pmatrix}$$

$$X_{[\mu\nu]} = \frac{1}{2}\eta_{\mu\rho}X^{[\rho\sigma]}\eta_{\sigma\nu} = \frac{1}{2}\begin{pmatrix} 0 & 1 & 2 & 1 \\ -1 & 0 & 2 & 1 \\ -2 & -2 & 0 & -1 \\ -1 & -1 & 1 & 0 \end{pmatrix}$$

E) The trace is given by

$$X = X^{\lambda}{}_{\lambda} = -2 - 2 = -4$$

F) This is the norm of V

$$V^{\mu}V_{\mu} = V^{\mu}V^{\nu}\eta_{\nu\mu} = 1 + 4 + 4 = 9$$

G) This is essentially multiplying a row vector with a matrix

$$V_{\mu}X^{\mu\nu} = (4, -2, 5, 7)$$

Exercise 8. Since $T^{\mu\nu}$ is symmetric, this is equivalent to saying

$$\partial_{\nu}T^{\mu\nu} = \partial_{\nu}T^{\nu\mu} = Q^{\mu}$$

This implies that the metric tensor is not conserved, so the spatial vector Q^i would represent some form of external force on the system.

Exercise 9. To get rid of the Dirac delta, we should integrate other some volume. Consider some volume element V and let's just integrate the whole thing, we'll also raise the indices right away

$$T^{\mu\nu} = \frac{1}{V}\sum_a \int_V \frac{p^{\mu\,(a)}p^{\nu\,(a)}}{p^{0(a)}}\delta^3(\mathbf{x} - \mathbf{x}^{(a)})dV = \frac{1}{V}\sum_a \frac{p^{\mu\,(a)}p^{\nu\,(a)}}{p^{0(a)}}$$

The difference is that the sum is now only taken over particles which pass through the given volume element. Now we consider specific elements

$$T^{00} = \frac{1}{V}\sum_a p^{0\,(a)} = \rho$$

Since the particles are identical and velocities isotropically distributed, we can define the (rest-frame) energy density ρ and it will be consistent regardless of the specific volume chosen. For the remaining elements

$$T^{i0} = T^{0i} = \frac{1}{V}\sum_a p^{i\,(a)}$$

This (spatial) vector must be zero because otherwise it would indicate some sort of preferred direction for the momenta, a contradiction to isotropy. Finally, we have the inner 3x3

$$T^{ij} = \frac{1}{V}\sum_a \frac{p^{i\,(a)}p^{j\,(a)}}{p^{0(a)}}$$

This should be a symmetric matrix, but due to isotropy it is also diagonal. This is because any preferred direction would be reflected by its eigenvectors, isotropy would imply that all three eigenvectors are the same (3-fold degenerate). For a symmetric matrix, this only occurs if it is proportional the the identity matrix, in tensor notation

$$T^{ij} = p\delta^{ij}$$

This p is now the pressure. Putting this all together, we can write

$$T^{\mu\nu} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{pmatrix} = (\rho + p)U^\mu U^\nu + p\eta^{\mu\nu}$$

where U is a stationary 4-vector. This is the perfect energy-momentum tensor as required.

Exercise 10. A) The transformation matrix (tensor) for a rotation about y is

$$\Lambda^{\mu'}{}_{\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & 0 & \sin\theta \\ 0 & 0 & 1 & 0 \\ 0 & -\sin\theta & 0 & \cos\theta \end{pmatrix}$$

This transformation is for $F^{\mu\nu} \to F^{\mu'\nu'}$ so we'll reintroduce that.

$$F^{\mu\nu} = \begin{pmatrix} 0 & E_1 & E_2 & E_3 \\ -E_1 & 0 & B_3 & -B_2 \\ -E_2 & -B_3 & 0 & B_1 \\ -E_3 & B_2 & -B_1 & 0 \end{pmatrix}$$

The transformed tensor is[11]

$$F^{\mu'\nu'} = \Lambda^{\mu'}{}_{\mu} F^{\mu\nu} \Lambda^{\nu'}{}_{\nu}$$

$$= \begin{pmatrix} 0 & E_1\cos\theta + E_3\sin\theta & E_2 & E_3\cos\theta - E_1\sin\theta \\ -E_1\cos\theta - E_3\sin\theta & 0 & B_3\cos\theta - B_1\sin\theta & -B_2 \\ -E_2 & B_1\sin\theta - B_3\cos\theta & 0 & B_1\cos\theta + B_3\sin\theta \\ E_1\sin\theta - E_3\cos\theta & B_2 & -B_1\cos\theta - B_3\sin\theta & 0 \end{pmatrix}$$

We see that both the electric and magnetic field components are rotated by the same matrix, i.e. both are rotated about the y axis in the same way.

$$\mathbf{E}, \mathbf{B} = \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix} \mathbf{E}, \mathbf{B}$$

[11] Note that this is essentially $F' = \Lambda F \Lambda^T$.

B) The transformation matrix for a boost in z is

$$\Lambda^{\mu'}{}_{\nu} = \begin{pmatrix} \cosh\phi & 0 & 0 & -\sinh\phi \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\sinh\phi & 0 & 0 & \cosh\phi \end{pmatrix}$$

Proceeding the same way as before,

$$F^{\mu'\nu'} = \Lambda^{\mu'}{}_{\mu} F^{\mu\nu} \Lambda^{\nu'}{}_{\nu}$$

$$\begin{pmatrix} 0 & E_1\cosh\phi - B_2\sinh\phi & E_2\cosh\phi + B_1\sinh\phi & E_3 \\ -E_1\cosh\phi + B_2\sinh\phi & 0 & B_3 & E_1\sinh\phi - B_2\cosh\phi \\ -E_2\cosh\phi - B_1\sinh\phi & -B_3 & 0 & E_2\sinh\phi + B_1\cosh\phi \\ -E_3 & -E_1\sinh\phi + B_2\cosh\phi & -E_2\sinh\phi - B_1\cosh\phi & 0 \end{pmatrix}$$

So we see that there is some mixing between the x and y components of the electric/magnetic fields, while the z component of both fields are kept constant.

Exercise 11. Since $F_{\mu\nu} = -F_{\nu\mu}$ we can easily verify that

$$\begin{aligned} \partial_{[\mu}F_{\nu\lambda]} &= \frac{1}{6}[\partial_\mu(F_{\nu\lambda} - F_{\lambda\nu}) + \partial_\nu(F_{\lambda\mu} - F_{\mu\lambda}) + \partial_\lambda(F_{\mu\nu} - F_{\nu\mu})] \\ &= \frac{1}{3}(\partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} + \partial_\lambda F_{\mu\nu}) = 0 \\ \therefore 0 &= \partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} + \partial_\lambda F_{\mu\nu} \end{aligned}$$

To see that these actually give the last two Maxwell's equations, first note that if $\mu = \nu = \lambda$ then we just get $0 = 0$ since the there are no diagonal elements in $F^{\mu\nu}$. The anti-symmetry of $F^{\mu\nu}$ also implies that if any two indices are equal, we just get $0 = 0$, for instance

$$\partial_\mu F_{\nu\nu} + \partial_\nu F_{\nu\mu} + \partial_\nu F_{\mu\nu} = \partial_\mu F_{\nu\nu} = 0$$

So there are only 24 possible combinations of indices that yield a useful equation. Since the equation is invariant under cyclic permutation of indices and swaps

$$\partial_\mu F_{\lambda\nu} + \partial_\lambda F_{\nu\mu} + \partial_\nu F_{\mu\lambda} = -(\partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} + \partial_\lambda F_{\mu\nu}) = 0$$

Thus we can consider just the 4 combinations where the indices are strictly increasing, which is much more manageable than the original 64 combinations. The 4 equations given are

$$\begin{aligned} 0 &= \partial_t F_{12} + \partial_x F_{20} + \partial_y F_{01} \\ &= \partial_t B_z + \partial_x E_y - \partial_y E_x \\ 0 &= \partial_t F_{13} + \partial_x F_{30} + \partial_z F_{01} \\ &= -\partial_t B_y + \partial_x E_z - \partial_z E_x \\ 0 &= \partial_t F_{23} + \partial_y F_{30} + \partial_z F_{02} \\ &= \partial_t B_x + \partial_y E_z - \partial_z E_y \\ 0 &= \partial_x F_{23} + \partial_y F_{31} + \partial_z F_{12} \\ &= \partial_x B_x + \partial_y B_y + \partial_z B_z \end{aligned}$$

With some rearranging we see that the first three equations correspond to the statement

$$\nabla \times \mathbf{E} + \partial_t \mathbf{B} = \tilde{\epsilon}^{ijk}\partial_j E_k + \partial_0 B^i = 0$$

and the last statement corresponds to

$$\nabla \cdot \mathbf{B} = \partial_i B^i = 0$$

Thus these are indeed the covariant forms of Maxwell's equations.

Exercise 12. A) Recall that the two energy-momentum tensors are

$$T^{\mu\nu}_{\text{scalar}} = \eta^{\mu\lambda}\eta^{\nu\sigma}(\partial_\lambda\phi)(\partial_\sigma\phi) - \eta^{\mu\nu}\left(\frac{1}{2}\eta^{\lambda\sigma}(\partial_\lambda\phi)(\partial_\sigma\phi) + V(\phi)\right)$$
$$T^{\mu\nu}_{\text{EM}} = F^{\mu\lambda}F^{\nu}{}_{\lambda} - \frac{1}{4}\eta^{\mu\nu}F^{\lambda\sigma}F_{\lambda\sigma}$$

We consider the scalar energy-momentum tensor first

$$\begin{aligned}
T^{00} &= (\partial_0\phi)^2 + \frac{1}{2}\eta^{\lambda\sigma}(\partial_\lambda\phi)(\partial_\sigma\phi) + V(\phi)\\
&= \frac{1}{2}\left(\dot\phi^2 + (\nabla\phi)^2\right) + V(\phi)\\
T^{i0} = T^{0i} &= -\dot\phi(\partial_i\phi) = -\dot\phi\nabla\phi\\
T^{ij} &= (\partial_i\phi)(\partial_j\phi) - \delta^{ij}\left(\frac{1}{2}\eta^{\lambda\sigma}(\partial_\lambda\phi)(\partial_\sigma\phi) + V(\phi)\right)\\
&= \nabla\phi\otimes\nabla\phi - \delta^{ij}\left(\frac{1}{2}\left((\nabla\phi)^2 - \dot\phi^2\right) + V(\phi)\right)
\end{aligned}$$

where $\otimes$ is the tensor product, also known as the outer product, of two vectors.

For the electromagnetic energy-momentum tensor[12]

$$\begin{aligned}
T^{00} &= E_iE^i + \frac{1}{2}(-E_iE^i + B_iB^i)\\
&= \frac{1}{2}\left(\mathbf{E}^2 + \mathbf{B}^2\right)\\
T^{i0} = T^{0i} &= (E_2B_3 - E_3B_2, E_3B_1 - E_1B_3, E_1B_2 - E_2B_1)\\
&= \mathbf{E}\times\mathbf{B}\\
T^{ij} &= \begin{pmatrix} -E_1^2 + B_3^2 + B_2^2 & -E_1E_2 - B_1B_2 & -E_1E_3 - B_1B_3\\ -E_1E_2 - B_1B_2 & -E_2^2 + B_3^2 + B_1^2 & -E_2E_3 - B_2B_3\\ -E_1E_3 - B_1B_3 & -E_2^2 + B_3^2 + B_1^2 & -E_3^2 + B_2^3 + B_3^3 \end{pmatrix} - \frac{\delta^{ij}}{2}(\mathbf{E}^2 + \mathbf{B}^2)\\
&= \frac{\delta^{ij}}{2}\left(\mathbf{E}^2 + \mathbf{B}^2\right) - \mathbf{E}\otimes\mathbf{E} - \mathbf{B}\otimes\mathbf{B}
\end{aligned}$$

[12] For the last one, I multiplied out the matrices.

B) We want to show that both energy-momentum tensors satisfy $\partial_\mu T^{\mu\nu} = 0$, in the case of the scalar field

$$\begin{aligned}\partial_\mu T^{\mu\nu} &= \eta^{\mu\lambda}\eta^{\nu\sigma}\left((\partial_\mu\partial_\lambda\phi)(\partial_\sigma\phi) + (\partial_\lambda\phi)(\partial_\mu\partial_\sigma\phi)\right)\\ &\quad - \frac{1}{2}\eta^{\mu\nu}\eta^{\lambda\sigma}((\partial_\mu\partial_\lambda\phi)(\partial_\sigma\phi) + (\partial_\lambda\phi)(\partial_\mu\partial_\sigma\phi)) - \eta^{\mu\nu}\partial_\mu V(\phi)\end{aligned}$$

We can relabel some indices in the second line (e.g. swap λ and σ) to see that the third and fourth terms are the same and can be combined.[13]

$$\partial_\mu T^{\mu\nu} = \eta^{\mu\lambda}\eta^{\nu\sigma}\left((\partial_\mu\partial_\lambda\phi)(\partial_\sigma\phi) + (\partial_\lambda\phi)(\partial_\mu\partial_\sigma\phi)\right) - \eta^{\mu\nu}\eta^{\lambda\sigma}(\partial_\mu\partial_\lambda\phi)(\partial_\sigma\phi) - \eta^{\mu\nu}\partial_\mu V(\phi)$$

The third term can be rewritten using the symmetry of η, commutativity of ∂, and more relabelling[14]

$$\begin{aligned}\eta^{\mu\nu}\eta^{\lambda\sigma}(\partial_\mu\partial_\lambda\phi)(\partial_\sigma\phi) &= \eta^{\sigma\nu}\eta^{\lambda\mu}(\partial_\sigma\partial_\lambda\phi)(\partial_\mu\phi)\\ &= \eta^{\sigma\nu}\eta^{\mu\lambda}(\partial_\sigma\partial_\mu\phi)(\partial_\lambda\phi)\\ &= \eta^{\mu\lambda}\eta^{\nu\sigma}(\partial_\mu\partial_\sigma\phi)(\partial_\lambda\phi)\end{aligned}$$

Thus we are left with

$$\partial_\mu T^{\mu\nu} = \eta^{\mu\lambda}\eta^{\nu\sigma}(\partial_\mu\partial_\lambda\phi)(\partial_\sigma\phi) - \eta^{\mu\nu}\partial_\mu V(\phi)$$

Note that we can write (by factoring)

$$\eta^{\mu\nu}\partial_\mu V(\phi) = \eta^{\mu\nu}(\partial_\mu\phi)\frac{dV(\phi)}{d\phi}$$

Recall that the equation of motion was

$$\eta^{\mu\nu}\partial_\mu\partial_\nu\phi - \frac{dV}{d\phi} = 0$$

We can use this to show, after yet more relabelling,

$$\partial_\mu T^{\mu\nu} = \eta^{\nu\sigma}(\partial_\sigma\phi)\left(\eta^{\mu\lambda}\partial_\mu\partial_\lambda\phi - \frac{dV}{d\phi}\right) = 0$$

Now for the EM energy-momentum tensor, the two equations of motions are the covariant Maxwell's equations

$$\begin{gathered}\partial_\mu F^{\nu\mu} = J^\nu = 0\\ \partial_{[\mu}F_{\nu\lambda]} = \partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} + \partial_\lambda F_{\mu\nu} = 0\end{gathered}$$

[13] Relabelling is valid since we are just changing what we call the dummy indices, however we should be relabel ν since that is what gives the parameters of the final vector.

[14] I think it is ok to relabel μ as long as everything gets contracted the same way

It is more convenient in this case to work using just the vectors, so we'll have four equations of motion

$$\begin{aligned}\nabla \times \mathbf{B} - \partial_t \mathbf{E} = 0\\ \nabla \cdot \mathbf{E} = 0\\ \nabla \times \mathbf{E} + \partial_t \mathbf{B} = 0\\ \nabla \cdot \mathbf{B} = 0\end{aligned}$$

The components of $\partial_\mu T^{\mu\nu}$ are

$$\begin{aligned}\partial_\mu T^{\mu 0} &= \partial_0 T^{00} + \partial_i T^{i0}\\ &= \frac{1}{2}\frac{d}{dt}\left(\frac{\mathbf{E}^2 + \mathbf{B}^2}{2}\right) + \nabla \cdot (\mathbf{E} \times \mathbf{B})\\ &= \mathbf{E} \cdot \dot{\mathbf{E}} + \mathbf{B} \cdot \dot{\mathbf{B}} + \mathbf{B} \cdot (\nabla \times \mathbf{E}) - \mathbf{E} \cdot (\nabla \times \mathbf{B}) = 0\end{aligned}$$

The spatial components require some additional thought. There are four terms, the first is

$$\partial_0 T^{0i} = \partial_0 T^{i0} = \frac{d}{dt}(\mathbf{E} \times \mathbf{B}) = (\dot{\mathbf{E}} \times \mathbf{B}) + (\mathbf{E} \times \dot{\mathbf{B}}) = 0$$

since a vector cross itself is zero.[15] Consider just the x component $i = 1$ for now.

$$\begin{aligned}\partial_j T^{ij} = &-E_x \partial_x E_x + E_y \partial_x E_y + E_z \partial_x E_z - B_x \partial_x B_x + B_y \partial_x B_y + B_z \partial_x B_z\\ &- E_x \partial_y E_y - E_y \partial_y E_x - B_x \partial_y B_y - B_y \partial_y B_x\\ &- E_x \partial_z E_z - E_z \partial_z E_x - B_x \partial_z B_z - B_z \partial_z B_x\\ = &-E_x(\nabla \cdot E) + E_y(\partial_x E_y - \partial_y E_x) + E_z(\partial_x E_z - \partial_z E_x)\\ &- B_x(\nabla \cdot B) + B_y(\partial_x B_y - \partial_y B_x) + B_z(\partial_x B_z - \partial_z B_x)\end{aligned}$$

All expression in parentheses vanish due to Maxwell's equations. The same approach applies to $i = 2$ and $i = 3$, showing that the EM energy-momentum tensor is conserved.

Exercise 13. A) Using the symmetry properties of $F^{\mu\nu}$, we find that

$$\begin{aligned}\mathcal{L}' &= \tilde{\epsilon}_{\mu\nu\rho\sigma} F^{\mu\nu} F^{\rho\sigma}\\ &= 8(F^{01}F^{23} - F^{02}F^{13} + F^{03}F^{12})\\ &= 8(E_1 B_1 + E_2 B_2 + E_3 B_3)\\ &= 8E_i B^i = 8\mathbf{E} \cdot \mathbf{B}\end{aligned}$$

[15] However the cross product is not associative, so we actually have an intermediate step where we use the vector triple product.

B) Let's re-derive the equations of motion. The first derivative doesn't change, but we will have to re-examine the second one. To avoid running out of Greek letters, we will keep the indices upper. However this requires us to calculate

$$\begin{aligned}
\frac{\partial}{\partial(\partial_\alpha\partial_\beta)}F^{\mu\nu} &= \eta^{\mu\rho}\eta^{\nu\sigma}\frac{\partial}{\partial(\partial_\alpha\partial_\beta)}F_{\rho\sigma} \\
&= \eta^{\mu\rho}\eta^{\nu\sigma}\frac{\partial}{\partial(\partial_\alpha\partial_\beta)}(\partial_\rho A_\sigma - \partial_\sigma A_\rho) \\
&= \eta^{\mu\rho}\eta^{\nu\sigma}(\delta^\alpha_\rho\delta^\beta_\sigma + \delta^\alpha_\sigma\delta^\beta_\rho) \\
&= \eta^{\mu\alpha}\eta^{\nu\beta} + \eta^{\mu\beta}\eta^{\nu\alpha}
\end{aligned}$$

Thus the second derivative is

$$\begin{aligned}
\frac{\partial}{\partial(\partial_\alpha\partial_\beta)}\tilde{\epsilon}_{\mu\nu\rho\sigma}F^{\mu\nu}F^{\rho\sigma} &= \tilde{\epsilon}_{\mu\nu\rho\sigma}(\eta^{\mu\alpha}\eta^{\nu\beta} + \eta^{\mu\beta}\eta^{\nu\alpha})F^{\rho\sigma} + F^{\mu\nu}(\eta^{\rho\alpha}\eta^{\sigma\beta} + \eta^{\rho\beta}\eta^{\sigma\alpha}) \\
&= 2\tilde{\epsilon}_{\mu\nu\rho\sigma}(\eta^{\mu\alpha}\eta^{\nu\beta} + \eta^{\mu\beta}\eta^{\nu\alpha})F^{\rho\sigma}
\end{aligned}$$

But $\eta^{\mu\nu}$ is diagonal, so for this to be nonzero would require $\mu = \alpha = \beta = \nu$, but that would cause the Levi-Civita symbol to return zero. So this contribution to the derivative ends up being 0. In other words, this has no effect on the Maxwell's equations. I have no idea why this is the case.