ASTR 498: Special Relativity Practice Problems

Part 1: Basics

A four-vector is written with a greek superscript: x^μ or x^α are examples. This has four components, one for each of the four spacetime coordinates you have chosen. For example, in Cartesian coordinates, the differential separation vector is $dx^\mu = (dt, dx, dy, dz)$.

The invariant interval ds (an interval between events that is the same as measured by any observer) is related to the differential separation vector dx^μ through the metric tensor $g_{\mu\nu}$:

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

where we use the Einstein summation convention that for a repeated up and down index, we sum over all spacetime components. For example, in Cartesian coordinates, $\mu = t, x, y, z$ and the same with ν, so

$$ds^2 = g_{tt} dt^2 + g_{tx} dt dx + g_{ty} dt dy + g_{tz} dt dz + \ldots \tag{2}$$

In this problem set we will opt for simplicity and always use Cartesian coordinates. Note, though, that in general we could use other coordinate systems (e.g., spherical polar coordinates). We are also focusing on flat spacetime. With these restrictions, we have

$$g_{tt} = -c^2, \qquad g_{xx} = g_{yy} = g_{zz} = 1 \tag{3}$$

and all other components vanish. In general, one writes the metric tensor in a symmetric way, meaning that $g_{\mu\nu} = g_{\nu\mu}$ for any ν and μ. It is not, however, always the case that the metric is diagonal.

It is common in special and general relativity to use a system of units in which the speed of light c is set to 1; if it makes you more comfortable, you can think of this as using a coordinate $\hat{t}$ that is defined as a length-like quantity, $\hat{t} \equiv ct$. In that case, the metric would be

$$ds^2 = -d\hat{t}^2 + dx^2 + dy^2 + dz^2 \ . \tag{4}$$

This is sometimes written with the special Minkowski metric tensor $\eta_{\mu\nu} = (-1, 1, 1, 1)$ along the diagonal and zero elsewhere. That is, $ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu$.

To raise and lower indices, we use the metric tensor. In general, for example, $u_\nu = g_{\mu\nu} u^\mu$. Since in this expression we have μ as a lower and upper index, we sum over them:

$$u_\nu = g_{t\nu} u^t + g_{x\nu} u^x + g_{y\nu} u^y + g_{z\nu} u^z \ . \tag{5}$$

In our case the metric is diagonal, meaning that for any given ν, all but one of these terms vanishes. For example, $u_x = g_{xx} u^x$. However, for the more general metrics that we will encounter in general relativity, you'll need to sum over the terms.

You can also go the other way around: $u^\mu = g^{\mu\nu} u_\nu$. Here $g^{\mu\nu}$ is the matrix inverse of $g_{\mu\nu}$. For a diagonal metric such as the one we consider, the matrix inverse is another diagonal metric, where each entry is the reciprocal of the corresponding entry in the original metric. Therefore, for the Minkowski metric, the inverse is equal to the original. This is of course not true in general.

Finally, note that if you have summed over all free indices, then you have a scalar, i.e., a pure number. Pure numbers are the same in all frames, hence they are invariants. These are very useful indeed in figuring out relativistic problems. As an example, the squared length of any four-vector is measured to be the same by any observer: $A^\mu A_\mu$ is a scalar. Now, the actual value could be different than the squared length of a different vector: $A^\mu A_\mu \neq B^\mu B_\mu$ in general. But everyone will agree on $A^\mu A_\mu$ even if the individual components of A^μ appear different in different frames. You can extend this further as well. For example, the metric tensor $g_{\mu\nu}$ has two free indices, so it is a rank two tensor. But $g_{\mu\nu} g^{\mu\nu}$ has no free indices, so it is a scalar and thus an invariant. Similarly, it turns out that there is an electromagnetic tensor, usually represented $F^{\alpha\beta}$ (note that I'm changing μ and ν to α and β, not because it is important but to show that these are dummy variables; we can call them anything). The individual components of this tensor, which turn out to be components of the electric and magnetic fields, are measured differently by different observers. However, $F^{\alpha\beta} F_{\alpha\beta}$ is a scalar. Let me highlight the importance of scalars:

- Because a scalar is the same in all reference frames, you can go into an especially convenient reference frame to calculate the value of the scalar. You are then guaranteed it is the same in all frames. For example, if you want to know something about a particle, a good frame is the one in which the particle is at rest, i.e., in which the spatial components of the four-velocity all vanish.

Part 2: Practice Problems

1. Write out explicitly the squared magnitude of some four-vector A^μ in Cartesian coordinates and flat spacetime.

Answer:

The squared magnitude is $A^\mu A_\mu$. Expanding this, we have $A^t A_t + A^x A_x + A^y A_y + A^z A_z$. We also know that the components with lowered indices are related to the upper via, e.g., $A_t = g_{\nu t} A^\nu = g_{t\nu} A^\nu$ (since the metric tensor is symmetric, we can write it either way). Since the metric tensor we are using is diagonal, this reduces to $A_t = g_{tt} A^t = -A^t$ because $g_{tt} = -1$. Doing the same thing with the other components, we get

$$A^\mu A_\mu = -(A^t)^2 + (A^x)^2 + (A^y)^2 + (A_z)^2 \ . \tag{6}$$

This has no free indices, therefore this is a invariant (i.e., a scalar). As a sanity check, we note that for slow relative motion between frames, the component A^t is the same in all frames (i.e., what we expect in normal Euclidean geometry). Our statement is therefore that $(A^x)^2 + (A^y)^2 + (A^z)^2$ is the same in all frames. This is the familiar point that the length of a three-vector is independent of the coordinate system, so it makes sense in the slow-motion limit.

2. Consider two four-vectors **A** and **B**. Construct an invariant dot product between them.

Answer:

We can write this in components. For example, we have A^μ and B^ν. However, $A^\mu B^\nu$ has two free indices, so it is not an invariant scalar. We might try something like $A^\mu B^\mu$, but that has two up indices instead of one up and one down, so that doesn't have to be invariant either. Forms with one being a vector and the other being a one-form seem more promising, but it might appear that there are two possibilities: $A^\mu B_\mu$ and $A_\nu B^\nu$. Are these necessarily equal?

To determine this, let's consider the construct

$$\mathbf{A} \cdot \mathbf{B} = g_{\mu\nu} A^\mu B^\nu \ . \tag{7}$$

Note that this has μ as an up and a down index, and also ν, so there are no free indices and this should indeed be a scalar. Remembering how we lower indices, we can rewrite this as

$$g_{\mu\nu} A^\mu B^\nu = (g_{\mu\nu} A^\mu) B^\nu = A_\nu B^\nu \ . \tag{8}$$

But we can also write

$$g_{\mu\nu} A^\mu B^\nu = g_{\mu\nu} B^\nu A^\mu = (g_{\mu\nu} B^\nu) A^\mu = B_\mu A^\mu = A^\mu B_\mu \ . \tag{9}$$

The forms are indeed the same.

In our particular metric, which is diagonal, the dot product is

$$\mathbf{A} \cdot \mathbf{B} = -A^t B^t + A^x B^x + A^y B^y + A^z B^z \ . \tag{10}$$

Let's check the slow motion limit again. In this limit, A^t and B^t are the same in all frames, so we ignore that term. Constancy of our dot product then implies that $A^x B^x + A^y B^y + A^z B^z$ is invariant, i.e., even if we rotate our axes this answer will be the same. That is indeed true in Euclidean three-dimensional geometry, so this checks out.

3. The four-velocity u^μ of a particle with nonzero rest mass is defined as $u^\mu = dx^\mu / d\tau$, where τ is the "proper time". The proper time is the time measured by an observer riding along

with the particle. Such an observer therefore sees no motion in the spatial coordinates: $ds^2 = -c^2dt^2 + dx^2 + dy^2 + dz^2 = -c^2dt^2$. Remember that the time interval dt depends on the observer; for this special observer (who has $dx = dy = dz = 0$), we call the time coordinate τ. The proper time interval $d\tau$ is therefore related to the invariant interval ds by $ds^2 = -c^2d\tau^2$. With this in mind, what is the squared magnitude of the four-velocity?

Answer:

The squared magnitude of any vector A^μ is $A^\mu A_\mu$. With lowering of indices, this can be rewritten as $g_{\mu\nu}A^\mu A^\nu$. Therefore,

$$u^\mu u_\mu = g_{\mu\nu}u^\mu u^\nu = g_{\mu\nu}(dx^\mu/d\tau)(dx^\nu/d\tau) = (g_{\mu\nu}dx^\mu dx^\nu)/d\tau^2 = ds^2/d\tau^2 = -c^2 \ . \qquad (11)$$

Let's try to look at this from another point of view. Consider an observer riding along with the particle, who therefore sees $dx = dy = dz = 0$. The only coordinate that is changing is t. Therefore, after some time dt, the spacetime "distance" traveled (i.e., the interval) is given by $ds^2 = -c^2dt^2$. Recall that the minus sign, which I admit seems counterintuitive, is the result of our choice of metric signature (we could have chosen a positive sign for the time component and a negative sign for the rest; it's just a convention). This implies that the squared four-velocity is $-c^2$. However, since the squared four-velocity, like the squared magnitude of any vector, is a scalar, this must be the answer in *any* reference frame.

4. The four-momentum of a particle with nonzero rest mass m is $p^\mu = mu^\mu$. Note that "rest mass" means the mass measured by an observer riding along with the particle. What is the squared magnitude of the four-momentum?

Answer:

$$p^\mu p_\mu = m^2u^\mu u_\mu = -m^2c^2 \ . \qquad (12)$$

We can motivate this further if we realize that in a given frame, $p^\mu = (E/c, p^x, p^y, p^z)$, where E is the total energy including rest mass and p^x etc. are the normal linear momentum components. Since $p^\mu p_\mu$ is a scalar, we can pick an especially convenient reference frame, in which the particle is at rest. Then $p^x = p^y = p^z = 0$, and the energy is just the rest mass energy $E = mc^2$. Therefore, in this frame, $p^\mu = (mc, 0, 0, 0)$. Writing things out in components,

$$p^\mu p_\mu = g_{\mu\nu}p^\mu p^\nu = g_{tt}(p^t)^2+g_{xx}(p^x)^2+g_{yy}(p^y)^2+g_{zz}(p^z)^2 = g_{tt}(p^t)^2 = -(p^t)^2 = -m^2c^2 \ . \qquad (13)$$

5. Consider a particle with nonzero rest mass m moving in the x direction with a speed that

gives a Lorentz factor γ. The energy is then $E = \gamma mc^2$. What is the momentum in the x direction?

Answer:

We have $p^\mu = (E/c, p^x, p^y, p^z) = (\gamma mc, p^x, 0, 0)$. Then from the previous problem,

$$\begin{aligned} p^\mu p_\mu = -m^2c^2 &= g_{\mu\nu}p^\mu p^\nu \\ &= g_{tt}p^t p^t + g_{xx}p^x p^x \\ -m^2c^2 &= -\gamma^2 m^2 c^2 + (p^x)^2 \\ (p^x)^2 &= (\gamma^2 - 1)m^2c^2 \\ &= [1/(1 - v^2/c^2) - 1]m^2c^2 \\ &= [1/(1 - v^2/c^2) - (1 - v^2/c^2)/(1 - v^2/c^2)]m^2c^2 \\ &= [(v^2/c^2)/(1 - v^2/c^2)]m^2c^2 \\ &= [1/(1 - v^2/c^2)]m^2v^2 \\ &= \gamma^2 m^2 v^2 \\ p^x &= \gamma mv \ . \end{aligned} \tag{14}$$

6. Consistent with our general $p^\mu p_\mu = -m^2c^2$ formula, it is the case that for photons (which have zero rest mass), $p^\mu p_\mu = 0$. Given this, can an electron absorb a single photon in free space?

Answer:

We are asked whether the process $e + \gamma \to e$ is possible with nothing else around, where here we use γ to represent a photon and not the Lorentz factor. The four-momentum must be conserved in this interaction. Let us represent the four-momentum of the initial electron (on the left hand side) by p, of the final electron (on the right hand side) by p', and of the photon by q. Four-momentum conservation then implies $p^\mu + q^\mu = p'^\mu$. Squaring gives us

$$\begin{aligned} (p^\mu + q^\mu)^2 &= (p'^\mu)^2 \\ (p^\mu + q^\mu)(p_\mu + q_\mu) &= (p'^\mu)(p'_\mu) \\ p^\mu p_\mu + 2p^\mu q_\mu + q^\mu q_\mu &= -m_e^2c^2 \\ -m_e^2c^2 + 2p^\mu q_\mu + 0 &= -m_e^2c^2 \\ 2p^\mu q_\mu &= 0 \ . \end{aligned} \tag{15}$$

Note that $p^\mu q_\mu = q^\mu p_\mu$, as we demonstrated before, which is why we were able to combine the two terms. We are left with a scalar equation, meaning that we can go into a convenient frame. The one we choose is the one in which the initial electron was at rest. Then $p^\mu = (m_e c, 0, 0, 0)$. Suppose that in this frame the energy of the photon was E; then $q^\nu = (E/c, q^x, q^y, q^z)$. The product is then

$$p^\mu q_\mu = g_{\mu\nu}p^\mu q^\nu = g_{tt}p^t q^t + g_{xx}p^x q^x + g_{yy}p^y q^y + g_{zz}p^z q^z = -(m_e c)(E/c) = -m_e E \ . \tag{16}$$

Therefore, our condition for four-momentum conservation in the $e + \gamma \to e$ process reduces to

$$-2m_e E = 0 \ . \tag{17}$$

But since $-2m_e \neq 0$, this means $E = 0$, i.e., there can't be a photon there. The process is not possible in free space. The process *can* occur if there is something else around, e.g., a nucleus or magnetic field.

7. Can a single photon in free space split into an electron and positron?

Answer:

We are asked whether $\gamma \to e^- + e^+$ can occur with nothing else around. Call the four-momentum of the photon q, of the electron p^-, and of the positron p^+. Then four-momentum conservation means $q^\mu = (p^-)^\mu + (p^+)^\mu$. Squaring gives

$$\begin{aligned} q^\mu q_\mu &= ((p^-)^\mu + (p^+)^\mu)((p^-)_\mu + (p^+)_\mu) \\ 0 &= (p^-)^\mu (p^-)_\mu + (p^+)^\mu (p^+)_\mu + 2(p^-)^\mu (p^+)_\mu \\ 0 &= -2m_e^2 c^2 + 2(p^-)^\mu (p^+)_\mu \ . \end{aligned} \tag{18}$$

This is a frame-independent equation, so again we can pick a convenient frame. A good choice is the one in which the electron is at rest, meaning that $(p^-)^\mu = (m_e c, 0, 0, 0)$. Suppose that in this frame the positron has Lorentz factor γ, so that $(p^+)^\mu = (\gamma m_e c, (p^+)^x, (p^+)^y, (p^+)^z)$. Then

$$(p^-)^\mu (p^+)_\mu = g_{\mu\nu}(p^-)^\mu (p^+)^\nu = g_{tt}(p^-)^t (p^+)^t + g_{xx}(p^-)^x (p^+)^x + g_{yy}(p^-)^y (p^+)^y + g_{zz}(p^-)^z (p^+)^z \tag{19}$$

or

$$(p^-)^\mu (p^+)_\mu = -(p^-)^t (p^+)^t = -\gamma m_e^2 c^2 \ . \tag{20}$$

Our four-momentum conservation equation therefore becomes

$$0 = -2m_e^2 c^2 - 2\gamma m_e^2 c^2 \ . \tag{21}$$

This is impossible, because the right hand side is always negative. Single-photon pair production in free space can't happen. However, with something nearby (e.g., a nucleus or a magnetic field), it can happen.

8. Suppose two photons hit head-on. In a particular frame, the energies of the photons are E_1 and E_2. What is the condition on E_1 and E_2 such that it is possible to produce an electron-positron pair?

Answer:

The process is $\gamma + \gamma \to e^- + e^+$. We could go through four-momentum conservation as before and verify that this process is possible in principle. However, let's take a slightly more physical approach in this case.

If pair production happens, everyone must agree that it has happened. It is not possible that one observer sees two photons enter and an electron and a positron leave, whereas another observer never sees the production of the electron and positron. Therefore, the conditions for pair production must exist in all frames. One obvious condition is that the total energy of the photons must exceed the mass-energy of the electron plus the mass-energy of the positron (i.e., twice the mass-energy of an electron). That is, we know that $E_1 + E_2 \geq 2m_e c^2$ in any frame. We therefore need, in some sense, to find the frame in which this is most challenging. As always, we look to invariants for some guidance.

Suppose that we call the four-momentum of the first photon q_1 and of the second q_2. Then we know that $q_1^\mu q_{2\mu}$ is invariant. If we assume that the first one is moving in the $+x$ direction and the second is therefore moving in the $-x$ direction, then $q^\mu q_\mu = 0$ for photons means $q_1^\mu = (E_1/c, E_1/c, 0, 0)$ and $q_2^\mu = (E_2/c, -E_2/c, 0, 0)$. Therefore $q_{2\mu} = (-E_2/c, -E_2/c, 0, 0)$ (using the same procedures we have on previous problems) and

$$q_1^\mu q_{2\mu} = -E_1 E_2/c^2 - E_1 E_2/c^2 = -2E_1 E_2/c^2 \tag{22}$$

is invariant. Given that -2 and c^2 are also invariant, it means that $E_1 E_2$ is invariant for head-on collisions. We have to be careful here: of course this means that $q_1^\mu q_{2\mu}$ is equal to $-2E_1 E_2/c^2$ in any frame at all (even those in which the collision is *not* head-on), but when we say that the product of the energies is invariant, we are specifying to frames in which the collision is head-on. That is, $E_1 E_2$ is fixed for observers moving at arbitrary speeds in the x direction, but not for general observers.

We therefore fix $E_1 E_2$, and our physical condition was that $E_1 + E_2 \geq 2m_e c^2$. If this is true in the frame that minimizes $E_1 + E_2$, it will be true in all frames. We thus need to minimize $E_1 + E_2$ given that $E_1 E_2 = A$, where A is some constant. Rewriting, we need to minimize $E_1 + A/E_1$. Taking the derivative relative to E_1 and setting to zero gives

$$1 - A/E_1^2 = 0 \Rightarrow E_1 = \sqrt{A} \Rightarrow E_1 = E_2 = E \; . \tag{23}$$

Then $E_1 + E_2 \geq 2m_e c^2 \Rightarrow E \geq m_e c^2$. The product is then $E_1 E_2 \geq (m_e c^2)^2$, which is the condition to allow pair production.