Relativity II: Dynamics

Chapter 12 of Morin covers relativistic dynamics, as does chapter 13 of Kleppner, or chapter 12 of Wang and Ricardo, volume 2. For four-vectors in relativistic dynamics, finish chapter 13 of Morin, or chapter 14 of Kleppner. For a deeper explanation of four-vectors, see chapter 2 of Schutz. There is a total of **81** points.

1 Energy and Momentum

Idea 1

The relativistic generalizations of energy and momentum are

$$E = \gamma mc^2, \quad \mathbf{p} = \gamma m\mathbf{v}.$$

These quantities are conserved, and m is defined as the rest mass. Note that m is *not* conserved in inelastic processes, while E *is* conserved; this is precisely the opposite of what happens nonrelativistically. The relativistic energy E automatically counts all contributions to the energy, including internal energy and rest energy mc^2.

[5] **Problem 1.** A few useful facts about energy and momentum, for future reference.

(a) Recalling the definition of the four-velocity from **R1**, show that

$$(E/c, \mathbf{p}) = mu^\mu$$

where u^μ is the four-velocity. Setting $c = 1$, this establishes $p^\mu = (E, \mathbf{p})$ is a four-vector.

(b) Suppose a particle is at rest in frame S'. Confirm explicitly that the components of the four-momentum p^μ transform as expected when going to frame S.

(c) Setting $c = 1$ for all future parts, show that the norm of the four-momentum is

$$p^\mu p_\mu = E^2 - |\mathbf{p}|^2 = m^2.$$

This is a very useful result that can simplify the solutions to many problems below, especially ones that simply ask for a final mass m. In this case one can often compute a single four-momentum and find its norm to get the answer.

(d) The expressions in idea 1 for E and $\mathbf{p}$ don't work for photons, since γ is infinite and m is zero. Instead, show that for a photon we have $p^\mu = \hbar k^\mu$.

(e) For a system of particles with total energy and momentum E and $\mathbf{p}$, find the velocity of the center of mass, i.e. the velocity of the frame where the total momentum is zero.

(f) In Newtonian mechanics, the kinetic energy K of an object with fixed mass m satisfies $dK = \mathbf{v} \cdot d\mathbf{p}$. Show that this also holds in relativity, assuming the rest mass m is fixed.

(g) As we'll discuss in more detail below, the force three-vector is defined as $\mathbf{F} = d\mathbf{p}/dt$ in relativistic mechanics. Show that $dK = \mathbf{F} \cdot d\mathbf{x}$, continuing to assume that m is fixed.

Solution. (a) We saw in **R1** that $u^\mu = (\gamma c, \gamma \mathbf{v})$. Multiplying by γ,

$$mu^\mu = (\gamma mc, \gamma m\mathbf{v}) = (E/c, \mathbf{p})$$

as desired.

(b) Suppose we boost by velocity u. Then, the new speed is $\frac{u+v}{1+uv}$, so the new value of γ is

$$\gamma' = \left(1 - \frac{(u+v)^2}{(1+uv)^2}\right)^{-1/2} = (1+uv)\gamma_u\gamma.$$

Thus, the boosted values of E and p are

$$E' = \gamma' m = \gamma_u(E + up), \quad p' = \gamma' mv = \gamma_u(p + uE).$$

These are exactly the expected Lorentz transformation properties.

(c) The norm is

$$E^2 - p^2 = \gamma^2 m^2 - \gamma^2 m^2 v^2 = \gamma^2 m^2 (1-v)^2 = m^2$$

as desired.

(d) This follows directly from the de Broglie relations $E = \hbar\omega$ and $\mathbf{p} = \hbar\mathbf{k}$.

(e) In this frame, $p' = 0$. Then using the result of part (b), we have $p - vE = 0$ where v is the velocity of the center of mass in the original frame. Therefore, $\mathbf{v} = \mathbf{p}/E$.

(f) Starting with $E^2 = p^2 + m^2$ and taking the differential of both sides,

$$2E\,dE = 2\mathbf{p} \cdot d\mathbf{p}.$$

Solving for dE, we have

$$dE = \frac{\mathbf{p}}{E} \cdot d\mathbf{p} = \mathbf{v} \cdot d\mathbf{p}$$

where we used the result of problem 1. Since K and E are the same up to a constant anyway, we conclude $dK = \mathbf{v} \cdot d\mathbf{p}$ as desired.

(g) We have $\mathbf{F} \cdot d\mathbf{x} = (\mathbf{F}\,dt) \cdot (d\mathbf{x}/dt) = \mathbf{v} \cdot d\mathbf{p} = dK$ using the result of part (f), as desired.

Idea 2

In relativistic dynamics problems, it is almost always better to work with energy and momentum than velocity; one typically shouldn't even mention velocities unless the problem asks for or gives them.

Example 1: Morin 12.2

Two photons each have energy E. They collide at an angle θ and create a particle of mass M. What is M?

Solution

The total four-momentum is

$$p^\mu = (2E, E(1+\cos\theta), E\sin\theta).$$

The mass is just the norm of the four-momentum, so

$$M = \sqrt{4E^2 - E^2(1+\cos\theta)^2 - E^2\sin^2\theta} = E\sqrt{2-2\cos\theta} = 2E\sin(\theta/2)/c^2$$

where we restored the factors of c at the end.

[1] **Problem 2** (KK 13.5)**.** A particle of rest mass m and speed v collides and sticks to a stationary particle of mass M. Find the final speed of the composite particle.

Solution. The four momentum of the first one is $(\gamma m, \gamma m v)$, and the second is $(M, 0)$. Thus, the final four momentum is $(\gamma m + M, \gamma m v)$. Thus the speed is

$$v = \frac{p}{E} = \frac{\gamma m v}{\gamma m + M}.$$

[2] **Problem 3.** USAPhO 2012, problem A1.

[3] **Problem 4.** USAPhO 2002, problem A2.

Example 2: Woodhouse 7.5

A particle of rest mass m moves with velocity $\mathbf{u}$ and collides elastically with a second particle, also of rest mass m, which is initially at rest. After the collision, the particles have velocities $\mathbf{v}$ and $\mathbf{w}$. Show that if θ is the angle between $\mathbf{v}$ and $\mathbf{w}$, then

$$\cos\theta = \frac{(1-\sqrt{1-v^2})(1-\sqrt{1-w^2})}{vw}.$$

Solution

First, a remark: in Newtonian mechanics, you learn that in an inelastic collision, the kinetic energy is dissipated into microscopic thermal motion. This often leads students to ask: if we keep track of the motion of all particles in detail, then are all collisions actually perfectly elastic? According to particle physics, the answer is no. You really can lose kinetic energy by converting it to mass-energy, in collisions which change the identity of the particles or produce new particles. Therefore, at particle colliders, we say a collision is elastic if the particles that come out are precisely the same as the ones that came in. For this example, that means the final particles still have rest mass m.

Conservation of energy and momentum imply

$$1 + \gamma_u = \gamma_v + \gamma_w, \quad \gamma_u \mathbf{u} = \gamma_v \mathbf{v} + \gamma_w \mathbf{w}.$$

To get an expression with $\cos\theta$, we take the norm squared of the momentum equation,

$$\gamma_u^2 u^2 = \gamma_v^2 v^2 + \gamma_w^2 w^2 + 2\gamma_v\gamma_w vw\cos\theta.$$

This can be substantially simplified by noting that $\gamma_u^2 u^2 = \gamma_u^2 - 1$, giving

$$2vw\gamma_v\gamma_w\cos\theta = \gamma_u^2 - \gamma_v^2 - \gamma_w^2 + 1.$$

The appearance of so many squares motivates us to square both sides of the energy equation,

$$1 + 2\gamma_u + \gamma_u^2 = \gamma_v^2 + \gamma_w^2 + 2\gamma_v\gamma_w.$$

Using this to simplify the right-hand side of the previous equation,

$$2vw\gamma_v\gamma_w\cos\theta = 2\gamma_v\gamma_w - 2\gamma_u = 2(\gamma_v\gamma_w - \gamma_v - \gamma_w + 1) = 2(\gamma_v - 1)(\gamma_w - 1)$$

where in the second step we used conservation of energy. After solving for $\cos\theta$, we get the desired result. This was a bit of a slog, but it's representative of the hardest calculations you'll ever have to do for special relativity problems.

As a check on that result, note that in the nonrelativistic limit we get $\cos\theta = 0$, indicating a 90° angle, which you saw in **M3**. At relativistic speeds, the opening angle gets smaller, which is a manifestation of the "beaming" effect you saw in **R1**. This is a familiar effect, commonly observed in particle physics experiments.

[3] **Problem 5** (Morin 12.6)**.** A ball of mass M and energy E collides head-on elastically with a stationary ball of mass m. Show that the final energy of mass M is

$$E' = \frac{2mM^2 + E(m^2 + M^2)}{2Em + m^2 + M^2}.$$

This problem is a little messy, but you can save yourself some trouble by noting that $E' = E$ must be a root of the equation you get for E'.

Solution. Let the answer be x. The final momentum is $(E + m, p)$, split between $P_M = (x, p_M)$ and P_m. Now, $P_m = (E + m, p) - (x, p_M)$, so taking the norm squared, we see that

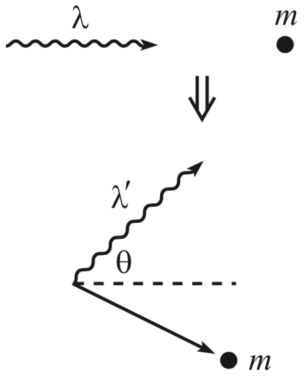
$$\begin{aligned}
&m^2 = (E + m - x)^2 - (\sqrt{E^2 - M^2} - \sqrt{x^2 - M^2})^2 \\
&\implies m^2 = (E^2 + m^2 + x^2 + 2Em - 2Ex - 2mx) - E^2 + M^2 - x^2 + M^2 + 2\sqrt{(E^2 - M^2)(x^2 - M^2)} \\
&\implies 0 = 2Em - 2Ex - 2mx + 2M^2 + 2\sqrt{(E^2 - M^2)(x^2 - M^2)} \\
&\implies (E^2 - M^2)(x^2 - M^2) = (mx + Ex - Em - M^2)^2.
\end{aligned}$$

This is manifestly a quadratic in x, and we know that one root is $x = E$, so applying Vieta's formulas and some tedious algebra reveals that

$$x = \frac{2mM^2 + E(m^2 + M^2)}{2Em + m^2 + M^2}$$

as desired.

[3] **Problem 6** (Morin 12.7)**.** In Compton scattering, a photon collides with a stationary electron.



(a) If the photon scatters at an angle θ, show that the resulting wavelength λ' is given in terms of the original wavelength λ by

$$\lambda' = \lambda + \frac{h}{mc}(1 - \cos\theta)$$

where m is the mass of the electron.

(b) While Compton scattering can occur for photons of any frequency, it is usually used in reference to X-rays, which have very high frequencies. Why?

Solution. (a) The original momentum of the system is $(E+m, E, 0)$ where E is the original energy of the photon. Letting x be the new energy of the photon. Then $P_\gamma = (x, x\cos\theta, x\sin\theta)$, and $P_m = (E+m, E, 0) - x(1, \cos\theta, \sin\theta)$. Taking the norm squared, we see that

$$\begin{aligned} m^2 &= (E+m-x)^2 - (E - x\cos\theta)^2 - x^2\sin^2\theta \\ \implies 0 &= 2Em - 2Ex - 2mx + 2Ex\cos\theta \\ \implies x &= \frac{Em}{m + E(1-\cos\theta)} = c^2\left(c^2/E + \frac{1}{m}(1-\cos\theta)\right)^{-1}. \end{aligned}$$

Now, $\lambda = hc/x = \lambda + \lambda_C(1-\cos\theta)$ where $\lambda_C = h/mc$.

(b) The wavelength shift is independent of frequency, and since $c = f\lambda$ the frequency shift (which is what we measure directly) is larger if the frequency begins large. The energy loss for visible photons is hardly noticeable, while it is very large for X-rays.

Indeed, for such photons we usually talk about Thomson scattering (as in **E7**) which does not change the frequency of the photon at all. At the level of relativistic dynamics, Thomson scattering is nothing more than the low-frequency limit of Compton scattering. Incidentally, at even higher frequencies, the result has more subtle corrections due to quantum field theory effects, and the cross section is given by the Klein–Nishina formula.

[3] **Problem 7.** USAPhO 2017, problem A4. However, to make it a little harder, solve part (a) without assuming E_b is small.

2 Relativistic Systems

Idea 3

The truly nonintuitive part of the result $E = mc^2$ is that changes in internal energy cause changes in mass. As a simple example, if you take a box of gas and heat it up, it'll have more mass than before, in every sense: the system will have more inertia, it'll have more momentum and kinetic energy when moving, it'll be heavier, and it'll exert more gravitational force on other objects. Some of the questions below illustrate how this can occur.

[3] **Problem 8.** The facts that $E = \gamma mc^2$ and $\mathbf{p} = \gamma m\mathbf{v}$ are conserved are fundamentally new results of relativity, so the logically cleanest way to set up the theory is to simply make these postulates, without any further justification. But this certainly isn't the most *convincing* way, if you don't already believe that relativity is true.

The most striking new result is the huge rest energy $E = mc^2$. Throughout his life, Einstein came up with many derivations of this result, starting from more familiar postulates. In this problem, we'll cover Einstein's 1946 derivation of $E = mc^2$. Specifically, we will prove that when the energy content of a body at rest decreases by ΔE, its mass decreases by $\Delta E/c^2$. The result then follows if one assumes that a zero-mass object has no rest energy.

Consider an object of mass M at rest, and suppose it emits photons with equal and opposite momenta p_γ upward and downward simultaneously. Let m be the final mass of the object.

(a) Now consider the same process in a frame moving with speed $v \ll c$ to the left. By using conservation of momentum in the x direction, show that

$$M = m + \frac{2p_\gamma}{c}.$$

Don't use the relativistic momentum formula here, because we're trying to work from first principles. Just use the fact that at $v \ll c$ the Galilean formula works.

(b) Using energy conservation, conclude the desired result.

(c) The derivation also works if one considers a frame moving upward with speed $v \ll c$. Carry out this analysis.

(d) The physicist Hans Ohanian has claimed that all of Einstein's derivations of $E = mc^2$, including this one, were circular. What do you think?

Solution. (a) The initial momentum is Mv. After emitting the photons, the body still has the same speed, so its final momentum is mv. Using Galilean velocity addition, the photons are emitted at a slight angle in this frame, contributing momentum $2p_\gamma/c$.

(b) Since the speeds are low, the $mv^2/2$ and $Mv^2/2$ contributions to the energy are second order and hence negligible. Energy $2p_\gamma c$ goes into photons, so an equal amount must have come out of rest energy. But the change in mass is $2p_\gamma/c$, so $\Delta E = \Delta M\, c^2$.

Finally, assuming that the rest energy of a particle goes to zero as its mass does to zero, which seems reasonable, gives $E = Mc^2$.

(c) Initially, the mass M has momentum downwards of Mv, and after the photons are emitted, the mass m has momentum mv which is made up for by the photons of different momenta

due to Doppler shifting. Since energy and momenta are proportional to frequency, which is proportional to $1 \pm v/c$, the difference in the momenta of the photons is $p_\gamma(2v/c)$ so we get $M = m + 2p_\gamma/c$. For energy, we have $\frac{1}{2}Mv^2 + \Delta E = \frac{1}{2}mv^2 + p_\gamma c(1 + v/c + 1 - v/c)$, and with second order v terms we have $\Delta E = 2p_\gamma c = \Delta M c^2$. The rest will be the same as above.

(d) This is a very subjective question, so opinions will vary. Here's my personal opinion.

Special relativity contains nonrelativistic mechanics as a special case. Therefore, there is no need to derive the results of special relativity from nonrelativistic physics – it stands on its own. Instead one can derive the results of nonrelativistic physics by taking limits of the results of special relativity. (It's just like quantum mechanics: you don't derive Schrodinger's equation from $F = ma$, you derive $F = ma$ as a limiting behavior of Schrodinger's equation.) Because of this, there is absolutely nothing illogical about simply *defining* $E = \gamma mc^2$. We then believe it because it reduces to results we already know about ($E = mv^2/2$ in the nonrelativistic limit) and also produces new verified predictions (nuclear power works).

(It's also worth noting that in nonrelativistic physics, the definition of energy simply follows from it being the conserved quantity associated with time translations. If we continue to define energy that way in special relativity, we automatically get $E = \gamma mc^2$. So it's not like $E = \gamma mc^2$ is some ad hoc, independent assumption on top of what we assumed in **R1**.)

Given the above, what is the point of trying to derive the rest energy expression at all? It's just to make people more comfortable with the new ideas of relativity. In physics you can often derive the same result in multiple ways. The rest energy follows automatically from the full framework of relativity, but it also follows by using *part* of the framework of relativity and part of the framework of nonrelativistic physics. This is useful if you're trying to explain why rest energy makes sense, to people who don't already believe in it: you get to the result using fewer unfamiliar assumptions, and possibly only ones that have already been tested experimentally. That's why arguments like these were important historically, when scientists were first grappling with relativity, and pedagogically, when students first encounter relativity.

A derivation using this kind of "hybrid" framework is necessarily weaker. For example, we had to make the somewhat random assumption above that a zero-mass object has no rest energy. You could argue that the only way to deduce *that* is to start with $E = mc^2$, making the argument "circular". But that doesn't really matter. The point of such a derivation is just to provide motivation, by explaining something new and unfamiliar in terms of things that are more believable. If you find the result that a zero-mass object has no rest energy believable, then the derivation works for you.

[3] **Problem 9.** Consider a completely black cube of density ρ and side length L sitting in free space. In some particular frame, plane electromagnetic waves of intensity I (in units of $\mathrm{W/m^2}$) approach the cube from the left and right, striking two faces of it head on. Neglect any radiation from the cube. If the cube has an initial velocity $v \ll c$ in this frame, find its displacement after a long time. (Hint: solving the problem exactly will be very messy; it's better to approximate early, since we only want an answer correct in the limit $v/c \to 0$.)

Solution. We'll set $c = 1$ for convenience, and expand everything to lowest order in v. The cube sees the frequency of the light beams redshifted and blueshifted, by a factor of $1 - v$ and $1 + v$ respectively. This affects the intensities by the same factor, so the net force on the box is

$$\frac{dp}{dt} = -2Pv, \quad P = IL^2, \quad m = \rho L^3.$$

In the nonrelativistic limit $p = mv$, so it's tempting to conclude that

$$\frac{dp}{dt} = m\frac{dv}{dt} = -2Pv.$$

However, this is incorrect, even in the limit of small v! The reason is that as v gets smaller, the drag force gets smaller, but the cube still is absorbing energy at the same rate, and this energy can be sizable. So even in the nonrelativistic limit, we have to account for the change of the rest mass of the cube. Specifically, we have

$$\frac{dm}{dt} \approx \frac{dE}{dt} = 2P$$

where we used $E \approx m$, valid in the nonrelativistic limit. Then

$$\frac{dp}{dt} = m\frac{dv}{dt} + \frac{dm}{dt}v = m\frac{dv}{dt} + 2Pv.$$

We thus arrive at the equation

$$m\frac{dv}{dt} = -4Pv.$$

Multiplying both sides by dt and integrating both sides, we find

$$\Delta x = \frac{mv_0}{4P} = \frac{\rho L v_0}{4I}.$$

Thus, even the naively negligible term changes the answer by a factor of 2.

[4] **Problem 10.** A rocket of initial mass M_0 starts from rest and propels itself forward along the x axis by emitting photons backward.

(a) Show that the final velocity of the rocket relative to the initial frame is

$$\frac{v}{c} = \frac{x^2 - 1}{x^2 + 1} = \tanh(\log x), \quad x = \frac{M_0}{M_f}$$

where M_f is the final rest mass of the rocket. (Hint: for this part, no integration is needed.)

(b) More generally, show that if the rocket fuel comes out at a speed u relative to the rocket,

$$\frac{v}{c} = \frac{x^{2u/c} - 1}{x^{2u/c} + 1} = \tanh((u/c)\log x)$$

where x is defined as above. (Hint: to avoid nasty differential equations, relate dm and dv.)

(c) Show that this reduces to the nonrelativistic rocket equation in the limit $u/c \to 0$.

(d) Show that in the limit $v/c \to 0$, the result of part (a) *also* reduces to the nonrelativistic rocket equation with exhaust speed c. Why does this work, given that photons are the most relativistic possible things?

Solution. (a) We see that the four momentum goes from $(M_0, 0)$ to $(\gamma M_f, \gamma M_f v)$. Since the difference is given by photons, we must have

$$-\gamma M_f v = \gamma M_f - M_0 \implies \gamma M_f(1 + v) = M_0 \implies \frac{1+v}{1-v} = x^2.$$

Solving for v and restoring c, we have

$$\frac{v}{c} = \frac{x^2 - 1}{x^2 + 1}$$

as desired.

(b) This part does require integration. The reason that part (a) didn't require integration is that all the emitted photons have the same speed in the original frame, because light always travels at c. But in this case, the emitted fuel will have varying speed in the original frame, depending on when it was emitted, so we have to actually do the calculation.

Since our variable x is in terms of mass, it's useful to relate the decrease in mass dm of the rocket with its increase in speed dv. Let's consider the very first instant the rocket is on. The decrease in the rocket's energy is dm (the kinetic energy is picks up is proportional to dv^2, which is negligible). All of this energy must be in the fuel, which is traveling with speed u, which means the mass of the fuel obeys

$$dm = \gamma_u \, dm_f.$$

The momentum carried by this bit of fuel is

$$dp = \gamma_u u \, dm_f = u \, dm.$$

This is equal to the momentum change of the rocket, $dp = m \, dv$. So combining everything,

$$-\frac{dm}{m} = \frac{dv}{u}.$$

In fact, this is exactly the same as the first half of the derivation of the ordinary rocket equation.

Now, in general this equation works as long as we're working in the momentarily comoving frame of the rocket. The place the relativity comes in is that the dv in this frame is not the same as the dv in the original frame. If the rocket has speed v in the original frame, then after accelerating by dv in its momentarily comoving frame, it ends up with speed

$$v' = \frac{v + dv}{1 + v \, dv} \approx v + dv - v^2 \, dv = v + (1 - v^2) \, dv$$

in the original frame. Therefore, we actually have in general

$$-\frac{dm}{m} = \frac{1}{u} \frac{dv}{1 - v^2}$$

and integrating both sides gives

$$\log x = \frac{1}{u} \int_0^v \frac{dv}{1 - v^2} = \frac{1}{2u} \int_0^v \frac{dv}{1 - v} + \frac{dv}{1 + v} = \frac{1}{2u} \log \frac{1 + v}{1 - v}.$$

Solving for v gives the result.

(c) We can use the approximation

$$x^{2u/c} = e^{(2u/c) \log x} \approx 1 + \frac{2u}{c} \log x$$

to arrive at

$$\frac{v}{c} \approx \frac{(2u/c) \log x}{2} \approx \frac{u}{c} \log x.$$

In other words, $v = u \log x$ which is precisely the nonrelativistic rocket equation. (Here we have implicitly assumed that $(u/c) \log x$ is small, which is equivalent to assuming that the rocket doesn't get to relativistic speeds. If u/c is nonrelativistic, this should be true for any reasonable value of x.)

(d) At first glance, this shouldn't make any sense. When $u/c \to 1$, the rocket fuel is always moving extremely relativistically, so how can we take the nonrelativistic limit? But pressing on, let's consider the limit $v/c \to 0$. This corresponds to $x \to 0$, so

$$v \approx \frac{(x-1)(x+1)}{2}\, c \approx (x-1)\, c = \frac{M_0 - M_f}{M_f}\, c.$$

On the other hand, the nonrelativistic rocket equation gives

$$v = u \log \frac{M_0}{M_f} = c \log \frac{M_0}{M_f} = c \log \left(1 + \frac{M_0 - M_f}{M_f}\right) \approx \frac{M_0 - M_f}{M_f}\, c$$

which matches.

Why does this work? Notice that the first half of the derivation in part (b) gives precisely the same result as the ordinary rocket equation; the only thing that matters from the standpoint of propelling the rocket is how much momentum you get from the fuel per energy spent. In the nonrelativistic limit, this ratio is $p/E \approx p/mc^2 = u/c^2$. When we apply the nonrelativistic rocket equation to relativistic fuel, we're implicitly using the "dumb" extrapolation $p/E = u/c^2$ for all speeds u. But this is in fact exactly true in relativity, because the factors of γ cancel out! For example, for photons we indeed have $p/E = 1/c$.

Thus, the only step where we actually need relativity is the velocity addition performed in the second half of part (b), but this effect is negligible as long as v/c is small, no matter how big u/c is.

[3] **Problem 11** (Grad)**.** An empty box of total mass M and perfectly reflecting walls is at rest in the lab frame. Then N photons are introduced into the box, each with frequency ω_0 in a standing wave configuration; one can think of these photons as continually bouncing back and forth with velocity $\pm c\,\hat{\mathbf{x}}$, with zero total momentum.

(a) State what the rest mass M_{tot} of the system will be when the photons are present.

(b) Consider the momentum of the system in an inertial frame moving along the x axis with speed $v \ll c$. Using the first order Doppler shift and assuming that at any moment, half the photons are moving left and half the photons are moving right, show that $p = M_{\text{tot}} v$. This provides a dynamical explanation of exactly how photons contribute to the inertia of an object.

(c) ⋆ Unfortunately, it is *not* true that half the photons are moving right at any given time. Show that the fraction of photons moving to the right is modified by an amount of order v/c, and find the total momentum accounting for this effect.

The analysis of part (b) is nice and neat, and you can sometimes find it in textbooks. But part (c) shows that this simple analysis is wrong! The official solution will show how to resolve the paradox, but it requires using the stress-energy tensor, which is well beyond the scope of Olympiad physics.

Solution. (a) Since $E = mc^2$, the rest mass is

$$M_{\text{tot}} = M + \frac{N\hbar\omega_0}{c^2}.$$

(b) Since $v \ll c$, we will use the equation $p = M_{\text{tot}}v$. We clearly have momentum Mv from the box itself. Meanwhile, the photons are Doppler shifted, so their total momentum is

$$p_\gamma = \frac{N}{2}\frac{\hbar\omega_0}{c}(1+v/c) - \frac{N}{2}\frac{\hbar\omega_0}{c}(1-v/c) = \frac{Nv\hbar\omega_0}{c^2}.$$

Dividing the momentum by v, we find the same result as in part (a).

(c) The fraction of photons moving to the right/left is $(1 \pm v/c)(N/2)$, which implies that

$$p_\gamma = \frac{N}{2}\frac{\hbar\omega_0}{c}(1+v/c)^2 - \frac{N}{2}\frac{\hbar\omega_0}{c}(1-v/c)^2 = \frac{2Nv\hbar\omega_0}{c^2}.$$

This appears to ruin the conclusion of part (b), and there is no other first-order effect to fix it.

Now we resolve the paradox. For simplicity, we'll analyze the system only at first order in v/c. There are numerous other effects at second order, such as the relativistic corrections to the Doppler shift and momentum, but these will complicate the analysis without adding much insight.

The resolution is very subtle, so to warm up, let's consider a simpler situation. In **R3**, you will learn that the charge density and current density can be combined into a four-vector $J^\mu = (\rho, \mathbf{J})$. If you integrate J^0 over all of space, you get the total electric charge Q. And it can be shown that whenever you integrate the zeroth component of a four-vector over all space, you get a Lorentz scalar. That is, the total charge is the same in all frames.

However, this *isn't* always true if you don't integrate over all of space. For example, suppose we had a segment of wire with a perfectly steady current flowing through it. In the wire's frame, it's neutral, and each new charge enters the left end as another charge exits the right end. But in a frame with a velocity along the wire, the loss of simultaneity effect implies that the wire has a net charge! That is, "the amount of charge on the wire" is *not* a Lorentz scalar. (This insight is essential to solving many of the problems in **R3**.) The amount of charge in a system is only necessarily a Lorentz scalar when there's no current flowing through it.

The same subtlety applies to energy and momentum. The total four-momentum of an isolated system (i.e. through which no external energy or momentum enters or leaves) is indeed a four-vector. That's why, for all the collision problems in this problem set, we could treat the four-momenta of particles long before or after the collision as four-vectors. But the photons in the box are not a closed system, because they are constantly interacting with the box, and as a result their four-momentum is not a four-vector. That's why the total momentum of the photons, in a frame where the box is moving, is not what we expect. However, the total momentum of the photons and box together is exactly what we expect, i.e. it is precisely $M_{\text{tot}}v$ in the nonrelativistic limit. The rest of the solution will show this explicitly.

To do this properly, we must introduce the stress-energy tensor $T^{\mu\nu}$, which is analogous to p^μ in the same way that J^μ is analogous to Q. Concretely, in a one-dimensional universe with only x and t directions, it is

$$T^{\mu\nu} = \begin{pmatrix} u & S \\ S & \sigma \end{pmatrix}$$

where the components have the following meanings.

- $T^{00} = u$ is the energy density.
- $T^{01} = S$ is the momentum density, i.e. what we must integrate over space to get momentum. We call this S because it coincides with the Poynting vector for a light wave.

- T^{10} is the current of energy in the x direction. For example, a particle of mass m and velocity v would have $T^{10} = mc^2 v$. It turns out that in general $T^{10} = T^{01}$.
- T^{11} is the current of x-momentum in the x direction, i.e. it has units of momentum per time. Physically, a flow of momentum is equivalent to a pressure.

Upon a Lorentz transformation, the stress energy tensor transforms differently from a four-vector. For a four-vector we would have

$$\begin{pmatrix} x' \\ t' \end{pmatrix} = \gamma \begin{pmatrix} 1 & -v \\ -v & 1 \end{pmatrix} \begin{pmatrix} x \\ t \end{pmatrix}$$

but for the stress-energy tensor we have

$$\begin{pmatrix} u' & S' \\ S' & \sigma' \end{pmatrix} = \gamma^2 \begin{pmatrix} 1 & -v \\ -v & 1 \end{pmatrix} \begin{pmatrix} u & S \\ S & \sigma \end{pmatrix} \begin{pmatrix} 1 & -v \\ -v & 1 \end{pmatrix}.$$

Expanding to first order in v, we have

$$S' = (u + \sigma)v + O(v^2).$$

The momentum of the photons is found by integrating S', giving

$$p_\gamma = \int_0^{L/\gamma} S' \, dx = L(u + \sigma)v + O(v^2).$$

The first term, Luv, is just what we would naively expect by transforming the four-momentum of the photons as a four-vector, and it's the answer we find in part (b). The pressure exerted by the walls yields the additional contribution $vL\sigma$. The energy density in the rest frame is simply $u = N\hbar\omega_0/L$, while the pressure exerted by the walls is $\sigma = N\hbar\omega_0/L$. Summing the terms gives

$$p_\gamma = 2Nv\hbar\omega_0$$

just as we found more directly in part (c).

Now we're in a position to see where the extra momentum is. The walls of the box cause a constant current of x-momentum to flow rightward through the photons. Hence the internal forces of the box must have an equal and opposite current of x-momentum leftward. Thus, by the same argument as above, in the primed frame p_{box} contains a contribution $-L\sigma v$ which precisely cancels the unwanted $L\sigma v$ contribution in the photons. Hence the total momentum is indeed

$$p_{\text{tot}} = Mv + N\hbar\omega_0 v$$

as it must be. For a similar setup, see this paper, which considers a capacitor containing an electromagnetic field, modeled classically instead of in terms of photons.

Remark

In Newtonian mechanics, we know that for an isolated system, $\mathbf{p}_{\text{tot}} = M_{\text{tot}}\mathbf{v}_{\text{CM}}$. In relativity, however, the idea of a "center of mass" no longer makes any sense. For example, suppose a particle with mass m decays into two photons. Each of the photons has no mass, so the center of mass is no longer defined! You can always define the mass of an overall system as

$\sqrt{E_{\text{tot}}^2 - p_{\text{tot}}^2}$, and this quantity remains equal to m, but it's no longer the sum of the masses of the individual parts. Since you can't break the mass of the system into parts, you can't sum over the parts to define a center of mass.

However, you can still define a "center of energy",

$$\mathbf{x}_{\text{CE}} = \frac{\sum_i \mathbf{x}_i E_i}{\sum_i E_i}$$

where E_i is the energy of particle i. It turns out that in relativity, we always have

$$\mathbf{p}_{\text{tot}} = \frac{E_{\text{tot}}}{c^2} \mathbf{v}_{\text{CE}}$$

which is called the "center of energy theorem". (Specifically, it comes from applying Noether's theorem to the symmetry of Lorentz boosts.) Of course, this reduces to $\mathbf{p}_{\text{tot}} = M_{\text{tot}} \mathbf{v}_{\text{CM}}$ in the nonrelativistic limit, since in that case almost all the energy is rest energy, $E = mc^2$.

3 Optimal Collisions

These collision problems are conceptually simple, but somewhat more mathematically challenging.

Idea 4

The minimum energy configuration of a system of particles with fixed total momentum is the one where they all move with the same velocity. This is easiest to show by boosting to the center of mass frame (i.e. the frame with zero total momentum) and then boosting back.

Example 3: KK 14.3

A high energy photon (γ ray) collides with a proton at rest. A neutral pi meson is produced according to the reaction

$$\gamma + p \to p + \pi^0.$$

What is the minimum energy the γ ray must have for this reaction to occur? The rest mass of a proton is 938 MeV and the rest mass of a neutral pion is 135 MeV.

Solution

The total four-momentum is $(E + m_p, E)$ where E is the energy of the γ ray in the lab frame. This four-momentum has norm $2Em_p + m_p^2$. Crucially, the norms of four-momenta don't change upon changing frames, so the total four-momentum in the center of mass frame is

$$\left(\sqrt{2Em_p + m_p^2}, 0\right)$$

because the total spatial momentum vanishes by definition. On the other hand, we also know that the reaction can just barely happen when both the proton and pion are produced at rest

in the center of mass frame, with a final four-momentum of $(m_p + m_\pi, 0)$. Hence we have

$$\sqrt{2Em_p + m_p^2} = m_p + m_\pi$$

and plugging in the numbers gives $E = 145\,\text{MeV}$. As expected, this is a little bit more than the mass-energy of the pion, because the final system inevitably has some kinetic energy too.

Example 4

Two photons of frequencies ω_1 and ω_2 collide head-on. Under what conditions can an electron-positron pair be created?

Solution

The naive answer is to say the energy present must exceed the rest energy,

$$\hbar\omega_1 + \hbar\omega_2 \geq 2m_e.$$

However, this is incorrect because the electron and positron will inevitably have kinetic energy, since the photons initially have a net momentum. The lowest total kinetic energy is achieved when the electron and positron come out with the same velocity, which is the velocity of the center of mass frame of the photons.

The total four-momentum of the photons is

$$(\hbar(\omega_1 + \omega_2), \hbar(\omega_1 - \omega_2))$$

in the lab frame, and $(E_{\text{cm}}, 0)$ in the center of mass frame. Therefore,

$$E_{\text{cm}}^2 = \hbar^2((\omega_1 + \omega_2)^2 - (\omega_1 - \omega_2)^2) = 4\hbar^2\omega_1\omega_2.$$

In the center of mass frame, the electron and positron can be produced at rest, so the condition is $E_{\text{cm}} \geq 2m_e$, which means

$$\hbar\sqrt{\omega_1\omega_2} \geq m_e.$$

[3] **Problem 12.** In a particle collider, a proton of mass m is given kinetic energy E and collided with an initially stationary proton.

(a) What is the minimum E required to produce a proton-antiproton pair, $p + p \to p + p + p + \overline{p}$?

(b) How about N proton-antiproton pairs, where $N = 1$ in part (a)?

The scaling behavior of the answer you found in part (b) is the reason many particle colliders use two beams going in opposite directions, even though managing two beams precisely enough to collide them at the desired points is technically challenging.

Solution. Part of the point of this problem is to test if you can avoid getting confused with using E to mean kinetic energy, when it's usually the total energy.

(a) Let p be the momentum of the proton. The total four momentum is then

$$p^\mu = (E + 2m, p).$$

We end up with four particles of mass m. From the idea above, the threshold energy is minimized when all of these particles have the same velocity, so they each have $p_i = p/4$. Then the final four-momentum is

$$p^\mu = 4(\sqrt{m^2 + p^2/16}, p/4).$$

Setting the energies equal, we have

$$\sqrt{16m^2 + p^2} = E + 2m$$

and using $E^2 = p^2 + m^2$ and simplifying gives

$$2Em = 12m^2, \quad E = 6m.$$

With this calculation in mind, the Bevatron at Berkeley was designed to accelerate protons to a kinetic energy of about $6.2m$. It discovered the antiproton in 1955, winning the 1959 Nobel prize.

(b) Now we have $2N + 2$ particles of mass m at the end, which have $p_i = p/(2N + 2)$. Now we instead have

$$p^\mu = (2N + 2)(\sqrt{m^2 + (p/(2N + 2))^2}, p/(2N + 2))$$

and setting the energies equal again gives

$$\sqrt{(2N + 2)^2 m^2 + E^2 + 2Em} = E + 2m$$

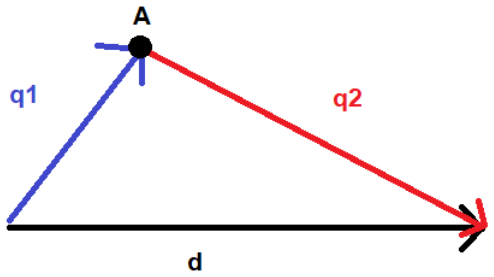
and solving gives

$$E = (2N^2 + 4N)m.$$

In other words, the energy required scales up quadratically in the mass-energy of the stuff you want to create!

[3] **Problem 13** (MPPP 196)**.** Two ultrarelativistic particles with negligible rest mass collide with oppositely directed momenta p_1 and p_2 elastically. Find the minimum possible angle between their velocities after the collision.

Solution. Let $\mathbf{q}_1, \mathbf{q}_2$ be the two new momenta of the new (still ultra-relativistic) particles. We see that $\mathbf{q}_1 + \mathbf{q}_2 = (p_1 - p_2)\hat{\mathbf{x}} \equiv \mathbf{d}$ and $q_1 + q_2 = p_1 + p_2 \equiv 2a$ (energy).



We see that the point A lies on an ellipse with foci at the endpoints of $\mathbf{d}$. The obtuse angle in the picture is actually the supplement of what we want, so we want to maximize the obtuse angle. This happens when A is on the perpendicular bisector of $\mathbf{d}$.

Let 2θ be the obtuse angle. We see that $\sin\theta = (d/2)/a = (p_1 - p_2)/(p_1 + p_2)$, giving a maximum possible angle of

$$\theta_{\max} = \pi - 2\sin^{-1}\left(\frac{p_1 - p_2}{p_1 + p_2}\right) = 2\cos^{-1}\left(\frac{p_1 - p_2}{p_1 + p_2}\right).$$

[3] **Problem 14.** IPhO 2003, problem 3A.

[4] **Problem 15.** APhO 2007, problem 3B. A comprehensive relativistic dynamics problem.

4 Relativistic Dynamics

In this section we'll consider some genuinely dynamic situations involving relativistic particles.

Idea 5

In relativity, the force four-vector is defined as

$$f^\mu = \frac{dp^\mu}{d\tau} = ma^\mu.$$

There's a bit of a subtlety here. In relativity, the invariant mass of a system can change when it absorbs energy, even if it doesn't exchange any particles with its environment. For example, putting a system on the stove gives it energy but not momentum, thereby changing $m = \sqrt{E^2 - p^2}$. That's a perfectly valid four-force, but it feels strange to call it a "force". Therefore, we often restrict to four-forces that don't change the invariant mass, and since

$$\frac{dm^2}{d\tau} = \frac{d}{d\tau}(p \cdot p) = 2mu \cdot f$$

that corresponds to demanding $f \cdot u = 0$.

Idea 6

There's also a second way to define force in special relativity, with three-vectors. The first subtlety here is that you could define it as $d\mathbf{p}/dt$ or $m\mathbf{a}$, but the two differ in relativity. Since accelerations transform in a rather nasty way, as we saw in **R1**, the usual choice is to define

$$\mathbf{F} = \frac{d\mathbf{p}}{dt}.$$

The second subtlety is that, whenever we define forces as three-forces, we usually implicitly assume that they fix the invariant mass m, i.e. we automatically rule out "put it on a stove" forces. Otherwise, there wouldn't be any way to tell how the energy changes over time.

[4] **Problem 16.** In this problem, we'll derive some properties of the three-force and four-force. For reference, see section 12.5 of Morin.

(a) Show that for a particle traveling along the $\hat{\mathbf{x}}$ direction,

$$\mathbf{F} = m(\gamma^3 a_x, \gamma a_y, \gamma a_z).$$

This is the relativistic three-vector analogue of $\mathbf{F} = m\mathbf{a}$, but it implies that force is no longer parallel to acceleration, which will be important in the problems below.

(b) Now let S' be the momentary rest frame of that particle. In this frame, since the particle is at rest, the nonrelativistic expression $\mathbf{F}' = m\mathbf{a}'$ holds. By using the transformation of acceleration derived in **R1**, show that

$$\mathbf{F} = (F_x', F_y'/\gamma, F_z'/\gamma).$$

That is, transverse forces are redshifted in relativity, while longitudinal forces are unchanged.

(c) Show that the components of the four-force are

$$f^\mu = \left(\gamma \frac{dE}{dt}, \gamma \mathbf{F}\right).$$

Use the relativistic transformation of the four-force to rederive the result of part (b).

(d) The four-impulse is defined as

$$\Delta p^\mu = \int f^\mu \, d\tau.$$

But you can also consider the Lorentz scalar

$$\int f^\mu \, dx_\mu.$$

This ought to be something nice and simple that you already know about. What is it?

Solution. (a) Using the chain rule and the definition of $\mathbf{p}$,

$$\mathbf{F} = \frac{d\mathbf{p}}{dt} = \gamma m \mathbf{a} + m\mathbf{v}\frac{d\gamma}{dt}.$$

Thus, the y and z components in the desired expression are correct, while the x component (i.e. the part parallel to $\mathbf{v}$ itself) has an extra contribution due to the second term. We have

$$\frac{d\gamma}{dt} = \frac{d\gamma}{dv}\frac{dv}{dt} = \gamma^3 v a_x$$

using a result from **R1**, so

$$F_x = \gamma m a_x (1 + \gamma^2 v^2) = m\gamma^3 a_x$$

as desired.

(b) We see that

$$\mathbf{F} = m(\gamma^3 a_x, \gamma a_y, \gamma a_z) = m(\gamma^3 a_x'/\gamma^3, \gamma a_y'/\gamma^2, \gamma a_z/\gamma^2) = (F_x', F_y'/\gamma, F_z'/\gamma)$$

where we used $\mathbf{F}' = m\mathbf{a}'$ in the last step.

(c) We just note that

$$\frac{d}{d\tau} = \frac{dt}{d\tau}\frac{d}{dt} = \gamma\frac{d}{dt}$$

which gives

$$f^\mu = \frac{dp^\mu}{d\tau} = \gamma\frac{dp^\mu}{dt} = \left(\gamma\frac{dE}{dt}, \gamma\frac{d\mathbf{p}}{dt}\right) = \left(\gamma\frac{dE}{dt}, \gamma\mathbf{F}\right).$$

In the primed frame of part (b), the components are

$$f^{\mu'} = \left(0, \mathbf{F}'\right).$$

Applying a Lorentz transformation to the original frame, we have

$$f^x = \gamma F_x', \quad f^y = F_y', \quad f^z = F_z'.$$

Since we know that $f^i = \gamma F_i$, we find

$$F_x = F_x', \quad F_y = F_y'/\gamma, \quad F_z = F_z'/\gamma$$

as desired.

(d) Using the chain rule, we have

$$I = \int f^\mu \frac{dx_\mu}{d\tau}\, d\tau = \int f \cdot u\, d\tau = \int \frac{1}{2m}\frac{dm^2}{d\tau}\, d\tau = \Delta m$$

so I gives the change in rest mass, which is of course a scalar, and just zero in most cases.

Remark

In popular science books and some older textbooks, relativistic dynamics is introduced using the idea of relativistic mass, $m_r = \gamma m$. This definition implies the simple results $E = m_r c^2$ and $\mathbf{p} = m_r\mathbf{v}$, so these books often say that relativistic dynamics is just like ordinary dynamics, except that moving objects have more mass. This picture is misleading because it breaks down once you go beyond one dimension: in problem 16, you showed that $\mathbf{F}$ is not even parallel to $\mathbf{a}$, so there's no definition of mass that recovers Newtonian mechanics. You instead need separate "transverse" and "longitudinal" relativistic masses,

$$\mathbf{F} = m_\perp \mathbf{a}_\perp + m_\parallel \mathbf{a}_\parallel, \quad m_\perp = \gamma m, \quad m_\parallel = \gamma^3 m.$$

I think this picture is honestly more confusing than helpful, though. It's better to avoid talking about mass and acceleration too much, and focus more on momentum and energy.

You'll also see arguments that relativistic mass is useful when thinking about gravity. In general relativity, all energy produces gravity equally. If you have a box with n particles bouncing around, which all have Lorentz factor γ and rest mass m, then the energy of the box is the same as that of n particles at rest, with mass m_r. So it looks like the gravity sourced by the particles is described by their relativistic mass. Unfortunately, this argument is *also* wrong, because in general relativity pressure also produces gravity. In the limit $\gamma \to \infty$, describing a gas of ultrarelativistic particles, the pressure contribution means we get twice as much gravitational attraction as would be predicted from the energy alone.

Example 5

A circular pendulum consists of a mass m attached to a string of length ℓ, with the other end fixed. Suppose the mass rotates in a small circle of radius $r \ll \ell$. Find the angular frequency of the oscillations in the lab frame, and in a frame where the entire setup moves vertically with a relativistic speed v.

Solution

In the lab frame, this is a standard rotational mechanics problem. By the small angle approximation, the horizontal component of the three-force is $F_\perp = mgr/L$. This is equal to

$$F_\perp = ma_\perp = m\omega^2 r$$

from which we immediately conclude $\omega = \sqrt{g/L}$. We can use the results of problem 16 to find the answer in the other frame. The two effects are that the transverse force is redshifted, and the force's relation with acceleration is different,

$$F_\perp = \frac{mgr}{\gamma L}, \quad F_\perp = \gamma m a_\perp = \gamma m \omega^2 r.$$

Combining these results, we find

$$\omega = \frac{1}{\gamma}\sqrt{\frac{g}{L}}.$$

Of course, γ is just the usual time dilation factor. We knew this had to be the answer, because time dilation follows directly from the postulates of relativity, but now we can explicitly show this is the right answer in this specific example. (With similar reasoning, you can show that a mass-spring system oscillates slower, too.)

Remark

It's important not to misunderstand the meaning of the above example. Like many old physicists, Oleg Jefimenko decided one day that relativity had to be completely wrong. His argument was along the lines of the previous example: he showed that length contraction and time dilation could be derived dynamically in some simple cases, without the need to switch frames. Therefore, they can't be "real".

This argument doesn't make sense. It's like saying that energy can't be real because you can solve many mechanics problems with just $F = ma$, without needing to invoke energy conservation. (Though amazingly, some people actually do spend years arguing whether force or energy is "more real", in a debate that resembles rival high school cheerleading squads, when it's better to realize that they're both wonderful tools with complementary uses.)

Furthermore, it actually turns out to be *extremely* difficult to derive the core results of relativistic dynamics (such as the "transverse" and "longitudinal" masses, already measured by the turn of the 20^{th} century) without using relativistic assumptions. In the early 1900s, many physicists tried to explain the dynamics of the electron solely in terms of its

electromagnetic fields. Since the field energy and field momentum of a moving point charge are infinite, it was necessary to take a model of the electron with finite size, but there were many possibilities, leading to many different expressions for the transverse mass, as well as persistent issues like the 4/3 problem mentioned in **E7**.

Relativity circumvents all of these issues. If you accept the postulates of relativity, you don't need to care whether the electron is shaped like a sphere, an ellipsoid, a torus, or a dumbbell: as long as its dynamics obey Lorentz symmetry, its four-momentum is a four-vector, and the usual results follow. And that's just as well, because with the advent of quantum mechanics, we learned that the electron is not like *any* of these classical models. But the relativistic result still holds, because our quantum theories obey the postulates of relativity too. This flexibility comes about because, like thermodynamics, relativity isn't so much a physical theory, as it is a framework within which many theories can be formulated.

Example 6

In the preceding example, how large of a force does the pendulum bob experience?

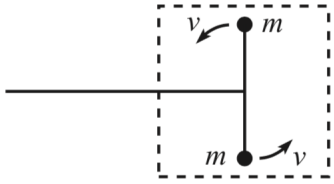
Solution

Here, we aren't referring to the three-force of the bob in the lab frame, which we already know is $m\omega^2 r$. We're referring to the proper force, i.e. the three-force in the bob's own frame. We could calculate this by Lorentz transforming into that frame, but there's an easier way. Note that $f \cdot f$ is a Lorentz scalar which, in the bob's frame, is $-|\mathbf{F}|^2$. Therefore, the proper force is $\sqrt{-f \cdot f}$, and evaluating this in the lab frame gives

$$\sqrt{(\gamma F_\perp)^2} = \gamma^2 m\omega^2 r.$$

That's a factor of γ^2 larger than the nonrelativistic expectation, so quite violent!

[3] **Problem 17** (Morin 12.8)**.** Consider a dumbbell made of two equal masses, m. The dumbbell spins around, with its center pivoted at the end of a stick.



If the speed of the masses is v, then the energy of the system is $2\gamma m$. Treated as a whole, the system is at rest. Therefore, the mass of the system must be $2\gamma m$. (Imagine enclosing it in a box, so that you can't see what's going on inside.) Convince yourself that the system does indeed behave like a mass of $M = 2\gamma m$, by pushing on the stick (when the dumbbell is in the "transverse" position shown in the figure) and showing that $F = dp/dt = Ma$.

Solution. Consider speeding up the system by dv to the left. The relativistic velocity addition

formula for u plus dv becomes

$$\frac{u+dv}{1+\frac{u\,dv}{c^2}} = (u+dv)(1-udv/c^2) = u + dv(1-u^2/c^2).$$

Let γ_u be $1/\sqrt{1-u^2}$. Let γ_u' be the gamma factor for $u + dv(1-u^2)$. One can easily check that $\gamma_u' = \gamma(1+u\,dv)$. Thus, the change in momentum due to the extra dv is

$$\gamma m(1+u\,dv)(u+dv(1-u^2)) - \gamma mu = \gamma m\,dv,$$

which is surprisingly what one would naively expect. Thus, the total change in momentum of the system is simply $dp = 2\gamma m\,dv$, so $dp/dt = M\,dv/dt$, as desired.

Idea 7

The Lorentz force is a three-force as defined in problem 16. That is, we have

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) = \frac{d\mathbf{p}}{dt}$$

and the force keeps the invariant mass fixed.

Example 7

A point charge q of mass m is initially at rest, and experiences a uniform electric field E. What time t does it take the object to move a distance x?

Solution

In **R1**, we found $x(t)$ for a uniformly accelerated rocket, which assumed a constant three-force in the momentarily comoving frame. By contrast, here we have a constant three-force $F = qE$ in the *lab* frame. However, we showed in problem 16 that forces along the direction of motion are the same in both frames, so these two problems are actually identical!

So we already know the answer to the problem, but it turns out that in the lab frame perspective, there's a slick alternative derivation that yields the result in one step. Consider the energy and momentum. Recall from problem 1 that the three-force F obeys $F = dp/dt$ and $F = dE/dx$. Therefore, when the object reaches its destination,

$$E = m + Fx, \quad p = Ft.$$

But we also know that $E^2 = p^2 + m^2$, so plugging the results in and solving for t gives

$$t(x) = \sqrt{x^2 + \frac{2mx}{F}}$$

which is compatible with our earlier expression for $x(t)$. The reason this was so easy is that momentum and energy behave simply in relativity, while position and velocity don't.

Example 8

The LHC accelerates protons to an energy of $E = 7\,\text{TeV}$, and is a tunnel of radius $R = 4.3\,\text{km}$. If the protons are kept in a circular orbit in the tunnel by a magnetic field of magnitude B, find the required value of B. If the value of B is kept constant, what would be the radius of a future collider which accelerates protons to an energy of 20 TeV?

Solution

The centripetal force required is

$$F = \left|\frac{d\mathbf{p}}{dt}\right| = \omega p$$

where ω is the angular velocity. The speed of the protons is very close to c, so the angular velocity is $\omega \approx c/R$, and the momentum is $p \approx E/c$. The deflecting force is $qvB \approx qcB$, so

$$qcB \approx \omega p \approx \frac{E}{R}.$$

Therefore, we have

$$B = \frac{E}{qcR} = \frac{7 \times 10^{12}}{(3 \times 10^8)(4.3 \times 10^3)}\,\text{T} = 5.4\,\text{T}.$$

This is slightly lower than what is actually used, because magnets don't take up the entire tunnel. Since $R \propto E$, the future collider would need a radius of

$$R' = \frac{20\,\text{TeV}}{7\,\text{TeV}} R = 12\,\text{km}.$$

Remark

You might be wondering how to write the Lorentz force as a four-force. It certainly should be possible, since we know electromagnetism is compatible with relativity (indeed, it led us to relativity in the first place), but it seems challenging because electromagnetism is so naturally written in terms of three-vectors. It turns out that the proper way to express the electromagnetic field in relativity is to join the electric and magnetic fields together, making them the components of an antisymmetric rank 2 tensor,

$$F_{\mu\nu} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -B_z & B_y \\ -E_y & B_z & 0 & -B_x \\ -E_z & -B_y & B_x & 0 \end{pmatrix}$$

called the field strength tensor. Then the four-force is

$$f^\mu = q u_\nu F^{\mu\nu}$$

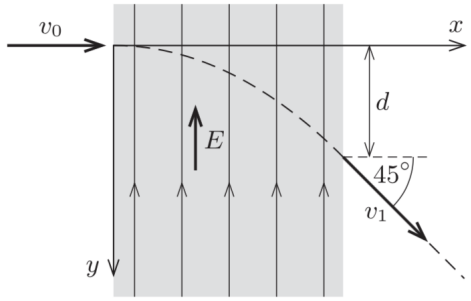
where u_ν is the four-velocity. Note that this ensures the rest mass of the particle is fixed, as

$$f \cdot u = q u_\mu u_\nu F^{\mu\nu} = -q u_\mu u_\nu F^{\nu\mu} = -f \cdot u$$

using the antisymmetric property, so $f \cdot u = 0$. (In fact, the requirement to keep the rest mass fixed is quite restrictive, so this is one of the simplest relativistic force laws.)

[2] **Problem 18.** USAPhO 2013, problem A3. A warmup question using the above facts.

[3] **Problem 19** (MPPP 192)**.** An electron moving with speed $v_0 = 0.6c$ enters a homogeneous electric field that is perpendicular to its velocity.



When the electron leaves the field, its velocity makes an angle 45° with its initial direction.

(a) Find the speed v_1 of the electron after it has crossed the electric field.

(b) Find the distance d shown above, if the strength of the electric field is $E = 510\,\text{kV/m}$.

Note that the rest energy of an electron is 510 keV.

Solution. (a) Since we are working with three-forces here, we use $\mathbf{F} = d\mathbf{p}/dt$. This tells us that the component of momentum p_x is unchanged. Since the velocity is at a 45° angle, so is the momentum, so $p_y = p_x$. Thus, the momentum increases by a factor of $\sqrt{2}$. The momentum per mass started at $0.6/0.8 = 3/4$, so its now $\frac{3}{4}\sqrt{2}$. Thus,

$$\frac{v_1}{\sqrt{1-v_1^2}} = \frac{3\sqrt{2}}{4} \implies \frac{v_1^2}{(1-v_1^2)} = \frac{9}{8} \implies v_1 = \frac{3c}{\sqrt{17}}.$$

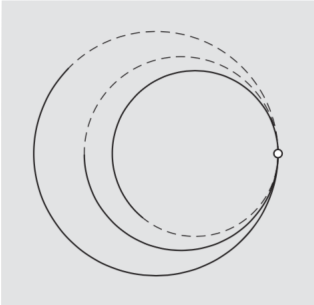
Note that this implies that v_x has *decreased*, even though the electric 3-force had no x-component. As we warned above, this is a manifestation of the fact that $\mathbf{F}$ is no longer parallel to $\mathbf{a}$ in relativity.

(b) As we showed in problem 1, the basics of work still work the same in relativity. The amount of work done on the electron is eEd, while the energy change is $m\Delta\gamma$, where

$$\Delta\gamma = \frac{1}{\sqrt{1-9/17}} - \frac{1}{\sqrt{1-9/25}} = \frac{\sqrt{17}}{\sqrt{8}} - \frac{5}{3}.$$

Plugging in the numbers gives $d = 20.8\,\text{cm}$.

[3] **Problem 20** (MPPP 194)**.** The trajectories of charged particles, moving in a homogeneous magnetic field, can be seen by observing the tracks they leave in cloud chambers. Because the particles are moving quickly, it is impossible to see the tracks being formed; instead, one must infer what happened from the shapes of the tracks. Is it possible that, when a charged particle decays into two other charged particles, the trail segments close to the decay point (before the particles have started to slow down significantly) are arcs of circles that touch each other, as shown?



If so, identify which track belongs to the original particle. If not, explain why not.

Solution. Number the three tracks as 1, 2, and 3 starting from the inside, and let their radii be $r_1 < r_2 < r_3$. We know that even for relativistic motion, the momentum of a particle is $p = qBr$. We can then use conservation of momentum and conservation of charge to investigate each case.

Case 1: Particle 1 decays, implying that a particle comes in along track 1, and particles leave along tracks 2 and 3. The curvatures of the tracks imply

$$q_1 > 0, \quad q_2 > 0, \quad q_3 > 0.$$

Conservation of charge and momentum imply

$$q_1 = q_2 + q_3, \quad q_1 r_1 = q_2 r_2 + q_3 r_3.$$

By combining these equations, we may solve for r_1 to find

$$r_1 = \frac{q_2 r_2 + q_3 r_3}{q_2 + q_3}.$$

However, this is impossible because we know r_1 is smaller than both r_2 and r_3.

Case 2: Particle 2 decays, which implies

$$q_1 < 0, \quad q_2 < 0, \quad q_3 > 0.$$

Conservation of charge and momentum imply

$$q_2 = q_1 + q_3, \quad |q_2 r_2| = |q_1 r_1| - |q_3 r_3|.$$

Being careful with minus signs, momentum conservation implies

$$-q_2 r_2 = -q_1 r_1 - q_3 r_3.$$

Again solving for r_1, we find

$$r_1 = \frac{q_3 r_3 + (-q_2) r_2}{q_3 + (-q_2)}$$

which is a contradiction for the same reason as in case 1.

Case 3: Particle 3 decays, which implies

$$q_1 < 0, \quad q_2 > 0, \quad q_3 < 0.$$

Conservation of charge and momentum imply

$$q_3 = q_1 + q_2, \quad |q_3 r_3| = |q_1 r_1| - |q_2 r_2|.$$

Again being careful with minus signs, momentum conservation implies

$$-q_3 r_3 = -q_1 r_1 - q_2 r_2.$$

Again solving for r_1, we find

$$r_1 = \frac{q_2 r_2 + (-q_3) r_3}{q_2 + (-q_3)}$$

which is again a contradiction. Thus, the series of tracks shown is impossible.

[3] **Problem 21.** USAPhO 2006, problem A4.

[3] **Problem 22.** USAPhO 2022, problem B2. A nice problem on deriving the time dilation formula for an electrostatic "clock".

[3] **Problem 23.** Consider a particle at the origin at time $t = 0$, with initial x-momentum p_0 and total energy E_0. A constant three-force F acts on the particle in the $-y$ direction.

(a) Calculate $y(t)$. (Hint: don't write down any equations containing γ, because it depends on $v_x(t)$, which we don't know yet.)

(b) Calculate $x(t)$.

(c) Combine these results to get $y(x)$. This is the path of a relativistic projectile.

Solution. We use the technique of example 7, setting $c = 1$ throughout.

(a) By the definition of three-force and the work-energy theorem,

$$p_x = p_0, \quad p_y = -Ft, \quad E = E_0 - Fy.$$

To find $y(t)$, we use the fact that $v_y = p_y/E$, so

$$\frac{dy}{dt} = -\frac{Ft}{E_0 - Fy}.$$

Separating and integrating, then using the initial condition gives

$$y^2 - \frac{2E_0}{F} y = t^2.$$

Solving the quadratic in y gives

$$y(t) = \frac{E_0}{F} - \sqrt{\frac{E_0^2}{F^2} + t^2}.$$

(b) Similarly, we have

$$\frac{dx}{dt} = \frac{p_x}{E} = \frac{p_0}{E_0 - Fy} = \frac{p_0}{\sqrt{E_0^2 + F^2 t^2}}$$

where we used the result of part (a). Separating and integrating,

$$x = \int_0^t \frac{p_0 \, dt}{\sqrt{E_0^2 + F^2 t^2}}.$$

Nondimensionalizing the integral, it can be performed with the hyperbolic trigonometric substitution $t = (E_0/F) \sinh \theta$, giving

$$x(t) = \frac{p_0}{F} \sinh^{-1} \frac{Ft}{E_0}.$$

(c) To get $y(x)$, we invert the above to get $t(x)$ and plug it into our expression for $y(t)$. We have

$$\frac{Ft}{E_0} = \sinh \frac{Fx}{p_0}$$

and plugging this in gives

$$y(x) = \frac{E_0}{F}(1 - \cosh(Fx/p_0c))$$

where we restored c in the last step. In other words, relativistic projectile motion follows an inverted catenary! To check the nonrelativistic limit, we just note that

$$\cosh u = 1 + \frac{u^2}{2} + \dots$$

which tells us that

$$y(x) \approx -\frac{1}{2}\frac{E_0}{F}\left(\frac{Fx}{p_0c}\right)^2 \approx -\frac{1}{2}\frac{mF}{p_0^2}x^2 \approx -\frac{1}{2}\frac{F}{mv_0^2}x^2$$

which is indeed the usual parabola.

[5] **Problem 24.** IPhO 1994, problem 1. Print out the custom answer sheets before starting.

Remark

The setup of problem 24 is a nice model for mesons, particles composed of two quarks. And it's not just something made up for an Olympiad; it is a simple version of the MIT "bag model", which was one of the most important advances in the field in the 1970s. In fact, if you look at the original paper, which has thousands of citations, you'll find the answer to the IPhO question in figure 3!

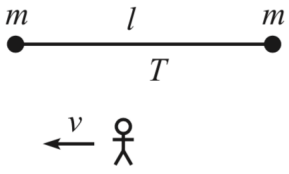
Idea 8

In string theory, strings carry a constant tension T, in the sense that the force $\mathbf{F} = d\mathbf{p}/dt$ exerted on one piece of string by its neighbors is T in the momentary rest frame of that piece. The strings may stretch or shrink freely, and have zero mass when they have zero length.

[3] **Problem 25** (Morin 12.16)**.** A simple exercise involving relativistic string.

(a) Two masses m are connected by a string of length ℓ and constant tension T. The masses are released simultaneously, and they collide and stick together. What is the mass, M, of the resulting blob?

(b) Consider this scenario from the point of view of a frame moving to the left at speed v.



The energy of the resulting blob must be γMc^2. Show that you obtain the same result by computing the work done on the two masses.

Solution. (a) The total work done on the masses is ℓT, so by energy conservation this must manifest as rest energy in the final blob, $M = 2m + \ell T/c^2$.

(b) Let $c = 1$. The initial energy is $2\gamma m$, so we need to show that the work done is $\gamma \ell T$.

At first glance, this is puzzling, because the initial distance between the masses in this frame is ℓ/γ. Therefore, naively applying $W = \int F\,dx$, we have

$$W = \int T\,dx_1 - \int T\,dx_2 = T\int dx_1 - dx_2 = T\ell/\gamma$$

which is wrong. The resolution is that we have assumed the masses are released simultaneously in the original frame, which means they aren't released simultaneously in this frame.

The mass on the left will start accelerating first, and after some time, the mass on the right will accelerate. In the original frame, these two events have $\Delta x = \ell$ and $\Delta t = 0$. Thus, applying the Lorentz transformation,

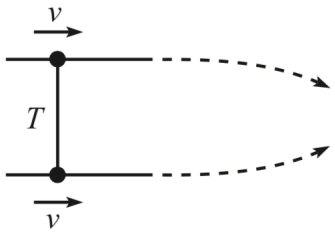
$$\Delta x' = \gamma \Delta x = \gamma \ell.$$

Suppose that after it starts experiencing the tension, the left mass moves a distance x_0 before it collides with the right mass. Then the above calculation shows that after the right mass starts experiencing the tension, it moves a distance $x_0 - \Delta x'$ until collision. Thus,

$$W = T(x_0 - (x_0 - \Delta x')) = \gamma \ell T$$

as desired.

[3] **Problem 26** (Morin 12.37)**.** Two equal masses are connected by a relativistic string with tension T. The masses are constrained to move with speed v along parallel lines, as shown.



The constraints are then removed, and the masses are drawn together. They collide and make one blob which continues to move to the right. Is the following reasoning correct?

> The forces on the masses point in the y direction. Therefore, there is no change in the momentum of the masses in the x direction. But the mass of the resulting blob is greater than the sum of the initial masses (because they collide with some relative speed). Therefore, the speed of the resulting blob must be less than v (to keep p_x constant), so the whole apparatus slows down in the x direction.

If your answer is "no," exactly what's wrong about the reasoning above?

Solution. The reasoning is incorrect. To see this, we can consider working in the initial rest frame of the system. In this frame, the masses just approach each other and collide, ending up at rest. So in the original frame, the whole apparatus must keep going at the same speed as before.

There are two ways to see what's going on. First, consider just the top mass, and work throughout in the original frame. Then the incorrect statement is the very first sentence: the three-force on the top mass is not always in the y direction. Recall the relativistic transformation of the three-force derived in problem 16. This tells us that if we align the x' axis with the *instantaneous* motion of the particle, then

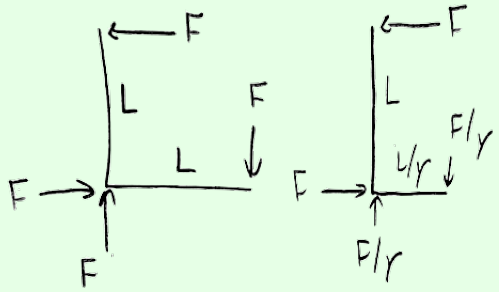
$$\mathbf{F} = (F'_{x'}, F'_{y'}/\gamma, F'_{z'}/\gamma).$$

Once the top mass gets moving, it has velocity components along both x and y, so the x' axis must be tilted accordingly. Upon applying this formula (i.e. redshifting the y' component of the force), we end up with a nonzero x component of the force, so the logic above fails.

Alternatively, we can consider the entire system, of the masses and string. In this case, the statement that fails is the second parenthetical, "to keep p_x constant". The issue here is that the string itself has a linear mass density of T/c^2, due to the energy stored in it in the stretching process, and hence also carries momentum. This needs to be accounted for in the momentum conservation equation, and gives the "missing" momentum we need. Note that this is totally compatible with the previous paragraph; in that case we focused on precisely how this string momentum ends up transferred to the masses.

Example 9: Right Angle Lever Paradox

In 1909, Lewis and Tolman found one of the first relativistic paradoxes. Consider a rigid lever in static equilibrium, with both arms of length L, experiencing the forces shown at left.



In a frame where the lever moves to the right with speed v, one of the lever arms will be contracted to L/γ, as shown at right. In addition, by the results of problem 16, the vertical external forces will be redshifted to F/γ. This implies a net torque of

$$\tau = FL - \frac{F}{\gamma}\frac{L}{\gamma} = FLv^2.$$

The paradox is, given that $\boldsymbol{\tau} = d\mathbf{L}/dt$, why doesn't the lever rotate?

Solution

The resolution is that, in the frame shown at right, the angular momentum of the lever is constantly increasing. The horizontal forces are continually doing equal and opposite work on the lever, resulting in a upward flow of energy of rate Fv in the vertical arm. As explained in **E7** and earlier in this problem set, in relativity, energy flow is equal to momentum density,

so the total upward momentum in the vertical arm is FLv. Therefore,

$$\frac{dL}{dt} = \frac{dx}{dt}(FLv) = FLv^2$$

exactly as expected.

Remark

The resolution of the right angle lever paradox is very controversial, with dozens of papers written on the subject, so we should discuss what it even means to "resolve" a paradox. As long as we believe relativity is self-consistent, we already know what's going to happen: the lever won't rotate. Everything the lever does is determined by $\mathbf{F} = d\mathbf{p}/dt$ alone, so if it looks like angular momentum considerations give a different answer, that just means we haven't formulated the latter correctly. The reason there are so many different resolutions out there is just that people choose different ways to define torque and angular momentum.

The solution above is the standard one, and its implicit definition of angular momentum can be motivated by Noether's theorem. That's a reasonable choice, since it's a specific output of a useful and general theorem, and we thereby know for sure that it's conserved for isolated systems. Unfortunately, explaining the definition takes some advanced math.

We define the angular momentum density tensor

$$M^{\mu\nu\rho}(x) = x^\mu T^{\nu\rho}(x) - x^\nu T^{\mu\rho}(x)$$

where the right-hand side contains the stress-energy tensor, from the solution to problem 11. The total angular momentum density is an antisymmetric rank 2 tensor,

$$J^{\mu\nu}(t) = \int d\mathbf{x}\, M^{\mu\nu\rho}(x).$$

Noether's theorem states that it is this quantity that is conserved for an isolated system, due to symmetry under rotations and boosts. (More specifically, the spatial components correspond to what we call angular momentum in nonrelativistic physics, $J^{xy} = L^z$, while the J^{0i} components have to do with the center of mass motion.) If there is an external four-force per unit proper volume $f^\mu(x)$, which in terms of the stress-energy tensor implies $\partial_\mu T^{\mu\nu} = f^\nu$, the rate of change of angular momentum is

$$\frac{dJ^{\mu\nu}}{dt} = \tau^{\mu\nu}, \quad \tau^{\mu\nu} = \int d\mathbf{x}\, x^\mu f^\nu(x) - x^\nu f^\alpha(x)$$

which looks quite similar to the Newtonian expression. The component of this equation relevant to this paradox is $dJ^{xy}/dt = \tau^{xy}$, where

$$J^{xy} = \int d\mathbf{x}\, xT^{y0} - yT^{x0}, \quad \tau^{xy} = \sum_k x^{(k)} F_y^{(k)} - y^{(k)} F_x^{(k)}$$

where the index k sums over the four forces, and the T^{i0} stand for the density of momentum in the i direction. From this point on, the solution proceeds as above.

There is something a bit strange here, though. In the lever's rest frame, the angular momentum is zero, so if $J^{\mu\nu}$ were a tensor, it would have to be zero in all frames, but instead it rises to arbitrarily high values in the other frame. The reason is that when there are external torques, $J^{\mu\nu}$ isn't a tensor at all, just like how the four-momentum wasn't a four-vector in the solution to problem 11. That's one of the reasons there's a controversy: there just doesn't exist any definition that has all the nice properties one might want.