Notes on “A first course in General Relativity” (Schutz, 2009)

Robert B. Scott,[1,2*]

[1]Institute for Geophysics, Jackson School of Geosciences, The University of Texas at Austin, Austin, Texas, USA

[2]Currently at the National Oceanography Centre, University of Southampton, Southampton, UK

*To whom correspondence should be addressed; E-mail: rscott@ig.utexas.edu.

April 17, 2010

Chapter 1

Special Relativity

1.1 Fundamental principles of special relativity (SR) theory

In the footnote 4 on p. 3, I believe the answer to the first question is "no, the soup is unaffected by a acceleration experienced by an astronaut in orbit." This would appear to also cause problems for SR, since how do we know that an observer is in an inertial frame? The acceleration cannot, necessarily, be measured locally. And there's no special reference frame with which to measure ones acceleration.

1.2 Definition of an inertial observer in SR

Gives a "geometrical" definition of an inertial reference frame, or coordinate system.

Notes that gravity makes it impossible to construct such an inertial coordinate system.

1.3 New units

Introduces what Misner et al. (1973) called "geometric units", wherein time is measured in distance of light travel.

They claim the motivation is that $c = 3 \times 10^8$m/s in SI, a "ridiculous value". I disagree, since then a $1/3s$ becomes the ridiculously large 10^5km!

Perhaps more useful motivation comes from velocity becoming a dimensionless parameter, space-time diagrams having the same units on all axes, and the world lines of light paths having unit slope.

1.4 Spacetime diagrams

Fig. 1.4: $\mathbf{v}$ is of course a vector, so one should replace this with $v = |\mathbf{v}|$.

1.5 Construction of the coordinates used by another observer

This is an extremely important section. Unfortunately he doesn't explain why the angle of the $\overline{x}-$axis to the $x-$axis is $\phi = \arctan(v)$, where $v = |\mathbf{v}|$ is the magnitude of the velocity of $\overline{\mathcal{O}}$ along the $x-$axis axis. Rather this result appears in Fig. 1.5 without explanation, nor even delegating it as an "exercise for the student". The result does follow from the construction of the $\overline{x}-$axis , but the steps involved are not trivial. Below is my attempt at a proof.

Call the unknown angle between the $\overline{x}-$axis and the $x-$axis α. Extend the line from $\mathcal{P}$ to $\mathcal{R}$ all the way to the $t-$axis, and call this intersection $\mathcal{Q}$. Draw two lines parallel to the $x-$axis, one through $\mathcal{R}$ and where it crosses the $t-$axis, call this $\mathcal{U}$. The other through $\mathcal{P}$ and where it crosses the $t - axis$, call this $\mathcal{T}$. The events $\xi, \mathcal{P}, \mathcal{R}$ form a right triangle, with hypotenuse $\xi\mathcal{R} = 2\,a$. We need the angle at $\mathcal{R}\xi\mathcal{P}$, which turns out to be $\chi = \pi/4 - \phi$. (Call angle $\mathcal{ORQ}$ γ. Then $\phi + \gamma + \pi/4 = \pi$, and angle $\xi\mathcal{RP}$, which is $\pi - \gamma = \pi/4 + \phi$. It follows that $\chi = \pi/4 - \phi$.) So now we can compute the length $\mathcal{RP} = 2a\sin(\chi) = a\sqrt{2}(\cos(\phi) - \sin(\phi))$. $\mathcal{UR} = a\sin(\phi)$. Then $\mathcal{QR} = \mathcal{UR}/\sin(\pi/4) = a\sin(\phi)\sqrt{2}$. Summing the two lengths $\mathcal{QP} = \mathcal{QR} + \mathcal{RP} = a\sqrt{2}\cos(\phi)$. We're now after $\mathcal{OT} = \mathcal{OQ} - \mathcal{TQ}$. But $\mathcal{OQ} = \mathcal{OU} + \mathcal{UQ}$, with $\mathcal{UQ} = \mathcal{UR} = a\sin(\phi)$, and $\mathcal{OU} = a\cos(\phi)$. Also, $\mathcal{TQ} = \mathcal{QP}\cos(\pi/4) = a\cos(\phi)$ So $\mathcal{OT} = a(\sin(\phi) + \cos(\phi)) - a\cos(\phi) = a\sin(\phi)$. Note that the sought after angle α satisfies $\tan(\alpha) = \mathcal{OT}/\mathcal{TP} = \mathcal{OT}/\mathcal{TQ} = a\sin(\phi)/a\cos(\phi)$, so $\alpha = \phi$, the desired result.

1.6 Invariance of the interval

This section purports to provide a proof of the invariance of the interval. But it assumes that the relationship between coordinates in different frames is linear, see discussion before Eqn. 1.2. But that's not a general proof then! What if they were not linear, e.g.

$$\Delta\overline{x}^{\alpha} = M0_{\alpha\beta}\Delta x^{\beta} + M1_{\alpha\beta}(\Delta x^{\beta})^{2}$$

Then $M1$ would have units of 1/length.

Eqn. 1.2 is a bit sneaky! He claims that "the numbers $\Delta\bar{t}, \Delta\overline{x}, \Delta\overline{y}, \Delta\overline{z}$ as linear combinations of their unbarred counterparts." One would expect then, that

$$\Delta\overline{x}^{\alpha} = \mathbf{A}^{\alpha\beta}x^{\beta}$$

where the coefficients of the matrix $\mathbf{A}^{\alpha\beta}$ give the linear transformation. And then the interval would be written

$$\Delta\overline{s}^{2} = \eta^{\alpha\beta}\Delta\overline{x}^{\alpha}\Delta\overline{x}^{\beta}$$

where $\eta^{\alpha\beta}$ is the metric tensor. And so,

$$\Delta\overline{s}^{2} = \eta^{\alpha\beta}\mathbf{A}^{\alpha\gamma}\Delta x^{\gamma}\mathbf{A}^{\beta\mu}\Delta x^{\mu}$$

But instead he wrote down his Eqn. 1.2, which is much less general. I claim he's being sneaky because he sounds like he's starting very general, talking about linear combinations, yet he sneaks in a much more restricted relationship!

He reduces the relationship between the interval in one frame and another to a function of the relative velocities of their origins, see Eqn. 1.5 on p. 10,

$$\Delta\overline{s}^{2} = -\mathbf{M}_{00}\Delta s^{2} = \phi(\mathbf{v})\Delta s^{2}.$$

To show that $\phi(\mathbf{v})$ depends only upon direction he considers the case of a rod (or two events A and B at the ends of the rod) lying along the $y-$axis. A and B are simultaneous in $\mathcal{O}$, and he argues that they are therefore also simultaneous in $\overline{\mathcal{O}}$, by constructing the $\overline{y}$-axis as he did in Fig. 1.3. But now the velocity of $\overline{\mathcal{O}}$ is orthogonal to the constructed axis, so of course the simultaneity of events is not changed by the coordinate transform. The intermediate result is that the space-time interval between A and B in either frame is just the square of the length, so their ratio is the sought-after $\phi(\mathbf{v})$. Now the subtle point is that he then claims that this ratio cannot depend on the direction of the velocity, because the rod is perpendicular to it and there are no preferred directions?! So what??

I think the solution is that $\mathbf{v}$ could be in an arbitrary direction in the $x-z$ plane. The ratio of lengths should not depend upon this direction, because then there would be preferred directions. But as far as I can tell, this only shows that the direction of the component orthogonal to the $y-$axis cannot influence $\phi(\mathbf{v})$.

1.7 Invariant hyperbolae

At the end of the section it's stated that

> "The lesson of Fig. 1.12b is that tangent to a hyperbola at any event $\mathcal{P}$ is line of simultaneity of the Lorentz frame whose time axis joins $\mathcal{P}$ to the origin. If this frame has velocity $\mathbf{v}$, the tangent has slope v."

The above is stated without proof or even hint that there's some calculation involved. Fortunately it proceeds straightforwardly. We seek the slope of the tangent to a hyperbola. Differentiate any time-like hyperbola wrt $\overline{x}$, to obtain in general

$$\frac{d\overline{t}}{d\overline{x}} = \frac{\overline{x}}{\overline{t}}.$$

At some point $\mathcal{P}$ the slope of the tangent wrt the $\overline{x}-$axis is $\frac{\overline{x_p}}{\overline{t_p}}$. Now if the $t-$axis is chosen to go through the origin and $\mathcal{P}$ its slope wrt the $\overline{t}-$axis will also be $\frac{\overline{x_p}}{\overline{t_p}}$, corresponding to $\tan(\phi) = v = \frac{\overline{x_p}}{\overline{t_p}}$. But we know from Fig. 1.5 that the corresponding $x-$axis will have slope v relative to the $\overline{x}-$axis. That is, the tangent is parallel the $\overline{x}-$axis, and is therefore a line of simultaneity for $\mathcal{O}$. QED.

1.8 Particularly important results

Time dilation This was straightforward once one uses the invariant hyperbolae. The event $x_{\mathcal{B}}$ was constructed so that it had $\overline{t} = 1$. The corresponding event in $\mathcal{O}$ is obtained by tracing the point back to the $t-$axis along the hyperbola with the same interval, $\Delta s^2 = -1$,

$$-t^2 + x^2 = -1$$

One must also note that the equation for the $\bar{t}-$axis is $t = x/v$. Substituting this into the hyperbola,

$$-t_B^2 + x_B^2 = -1 \tag{1.1}$$
$$-t_B^2 + (t_B v)^2 = -1 \tag{1.2}$$
$$t_B = \frac{1}{\sqrt{1-v^2}} \tag{1.3}$$

This gives Eqn. 1.8.

Lorentz contraction I still don't see how he came up with

$$x_{\mathcal{C}} = \frac{l}{\sqrt{1-v^2}}$$

But I obtain the same end result using instead the invariance of the interval, which in $\overline{\mathcal{O}}$ is

$$\Delta s^2_{\mathcal{AC}} = -\Delta\bar{t}^2 + \Delta\overline{x}^2 = -\bar{t}^2 + \overline{x}^2 = 0 + l^2 = l^2.$$

and therefore must also be in $\mathcal{O}$. I also used the equation for the $\overline{x}-$axis, $t = vx$. This was confusing at first since the units look wrong! But it's clear when you go back to Fig. 1.5 and note that $\tan(\phi) = v$, which was obvious for the $\bar{t}-$axis since the observer $\overline{\mathcal{O}}$ is moving along the $x-$axis at speed v. That the $\overline{x}-$axis was also inclined at the same angle ϕ was more complicated. One also needs a three relation, which is simply $x_{\mathcal{C}} - x_{\mathcal{B}} = vt_{\mathcal{C}}$. A little algebra gives the Lorentz contraction:

$$x_{\mathcal{B}} = l\sqrt{1-v^2}.$$

1.9 Lorentz transformation

The first step, substituting the equations for the $\overline{\mathcal{O}}$ axes proceeds immediately to

$$\bar{t} = \alpha(t - v\,x) \tag{1.4}$$
$$\overline{x} = \sigma(x - v\,t) \tag{1.5}$$

I had trouble seeing how $\alpha = \sigma$ from the path of a light ray, so I used the invariant hyperbolae instead. Substituting (1.4 and 1.5) into the equation for the interval from the origin, gives,

$$\Delta s^2 = -t^2 + x^2 = \Delta \bar{s}^2 = -\bar{t}^2 + \bar{x}^2$$

The cross term on the RHS involving $x\,t$ must be zero, giving that $\alpha^2 = \sigma^2$. Equating either of the other terms gives the Lorentz factor,

$$\alpha^2 = \frac{1}{\sqrt{1 - v^2}}$$

As Schutz (2009, p. 22) points out, the positive root is selected so that the coordinates are not inverted when $v = 0$.

In retrospect, it is clear how the path of a light ray gives $\alpha = \sigma$. Simply note that the world line of line ray has $\Delta x = \pm \Delta t$ and $\Delta \bar{x} = \pm \Delta \bar{t}$. Substitution into (1.4 and 1.5) gives,

$$\frac{\sigma}{\alpha} \left(\frac{\Delta x - v \Delta t}{\Delta t - v \Delta x} \right) = \frac{\Delta x}{\Delta t} = 1.$$

So $\sigma = \alpha$.

The Lorentz transformation is often said to reduce to the Galilean transformation in the limit $v \ll 1$, but that's not strictly true. Unlike for the Galilean transformation, in the Lorentz transformation time is affected at large distances even for small velocities.

1.10 Velocity composition law

1.11 Paradoxes and physical intuition

1.12 Further reading

A more thoughtful look at fundamentals, Bohm (2008).

1.13 Appendix: the twin paradox dissected

1.14 Exercises

1.14.1 Convert to geometric units

a)

$$10\,\mathrm{J} = 10\mathrm{N\ m} = 10\mathrm{kg\ m^2/s^2} = 10/9\times 10^{16}\mathrm{kg} = 1.1\times 10^{-16}\mathrm{kg}.$$

b)

$$100\mathrm{W} = 100\mathrm{J/s} = 1.1\times 10^{-15}\mathrm{kg/s} = 1.1\times 10^{-15}/3\times 10^{8}\mathrm{kg/m} = 0.3\times 10^{-23}\mathrm{kg/m}$$

c)

$$\hbar = 1.05\times 10^{-34}\mathrm{J\ s} = \frac{1.05\times 10^{-34}\mathrm{J\ s}}{3\times 10^{8}\mathrm{m/s}} = 0.33\times 10^{-42}\mathrm{kg\ m}$$

d) Car velocity [108 km/hr]

$$v = 30\mathrm{m/s} = 10^{-7}$$

e) Car momentum

$$p = 30\mathrm{m/s}\times 1000\mathrm{kg} = 10^{-4}\mathrm{kg}$$

f) Atmospheric pressure,

$$1\mathrm{bar} = 10^{5}\mathrm{N\ m^{-2}} = \frac{10^{5}\mathrm{kg\ m\ s^{-2}}}{9\times 10^{16}\mathrm{m^4 s^{-2}}} = 1.1\times 10^{-12}\mathrm{kg\ m^{-3}}$$

g) water density

$$10^{3}\mathrm{kg}m^{-3}$$

h) Luminosity flux

$$10^{6}\mathrm{J\ s^{-1}cm^{-2}} = 10^{10}\mathrm{J\ s^{-1}m^{-2}} = \frac{10^{10}\mathrm{kg\ s^{-3}m^{-1}}}{3^3\times 10^{24}\mathrm{m^3 s^{-3}}} \approx 4\times 10^{-16}\mathrm{kg\ m^{-4}}$$

1.14.6 Show that Eq. (1.2) contains only $M_{\alpha\beta} + M_{\beta\alpha}$ when $\alpha \neq \beta$, not $M_{\alpha\beta}$ and $M_{\beta\alpha}$ independently. Argue that this allows us to set $M_{\alpha\beta} = M_{\beta\alpha}$ without loss of generality.

$$\Delta\bar{s}^2 = \sum_{\alpha=0}^{3}\sum_{\beta=0}^{3} M_{\alpha\beta}(\Delta x^\alpha)(\Delta x^\beta)$$

Pick a pair of indices, $\alpha = \alpha'$ and $\beta = \beta'$ say, where $\alpha' \neq \beta'$, and $\alpha' \in \{0\ldots 3\}$ and $\beta' \in \{0\ldots 3\}$. So $\Delta\bar{s}^2$ contains a term like,

$$M_{\alpha'\beta'}(\Delta x^{\alpha'})(\Delta x^{\beta'}).$$

But $\Delta\bar{s}^2$ also contains a term like,

$$M_{\beta'\alpha'}(\Delta x^{\beta'})(\Delta x^{\alpha'}) = M_{\beta'\alpha'}(\Delta x^{\alpha'})(\Delta x^{\beta'}).$$

The equality follows because of course the product does not depend upon the order of the factors. So we can group these two terms and factor out the $(\Delta x^{\alpha'})(\Delta x^{\beta'})$ leaving,

$$(\Delta x^{\alpha'})(\Delta x^{\beta'})(M_{\alpha'\beta'} + M_{\beta'\alpha'})$$

Because the off-diagonal terms always appear in pairs as above, we could without changing the interval (and therefore without loss of generality) replace them with their mean value

$$\tilde{M}_{\alpha\beta} \equiv (M_{\alpha\beta} + M_{\beta\alpha})/2$$

Thus the new tensor $\tilde{M}_{\alpha\beta}$ is by construction symmetric.

1.14.7 In the discussion leading up to Eq. (1.2), assume that the coordinates of $\overline{\mathcal{O}}$ are given as the following linear combinations of those of $\mathcal{O}$:

$$\bar{t} = \alpha t + \beta x, \quad (1.6)$$
$$\bar{x} = \mu t + \upsilon x, \quad (1.7)$$
$$\bar{y} = a y, \quad (1.8)$$
$$\bar{z} = b z, \quad (1.9)$$

where $\alpha, \beta, \mu, \upsilon, a,$ and b may be functions of the velocity $\mathbf{v}$ of $\overline{\mathcal{O}}$ relative to $\mathcal{O}$, but they do not depend on the coordinate. Find the numbers $\{M_{\alpha\beta}, \alpha, \beta = 0, \ldots 3\}$ of Eq. (1.2) in terms of $\alpha, \beta, \mu, \upsilon, a,$ and b.

First note that the origins of the two coordinate systems line up, and that $\Delta t = t$ etc. Then the result follows from straightforward substitution of (1.6) to (1.9) into Eq. (1.1)

$$\Delta\bar{s}^2 = -\Delta\bar{t}^2 + \Delta\bar{x}^2 + \Delta\bar{z}^2 + \Delta\bar{z}^2 \tag{1.10}$$

$$= -(\alpha\Delta t + \beta\Delta x)^2 + (\mu\Delta t + \upsilon\Delta x)^2 + (a\Delta y)^2 + (b\Delta z)^2 \tag{1.11}$$

Grouping terms we find that $(-\alpha^2 + \mu^2)$ multiplies Δt^2, so $M_{00} = (-\alpha^2 + \mu^2)$. Similarly, the term multiplying Δx^2 is $M_{11} = -\beta^2 + \upsilon^2$. The cross terms give $M_{01} = M_{10} = -\alpha\beta + \mu\upsilon$, and the remaining diagonal terms are $M_{22} = a^2$, $M_{33} = b^2$.

1.14.8 a) Derive Eq. (1.3) from Eq. (1.2) for general $M_{\alpha\beta}$.

Start with Eq. (1.2)

$$\Delta\bar{s}^2 = M_{\alpha\beta}\Delta x^\alpha \Delta x^\beta.$$

Substituting

$$\Delta\bar{s}^2 = M_{00}\Delta t^2 + M_{0i}\Delta x^i \Delta t + M_{i0}\Delta x^i \Delta t + M_{ij}\Delta x^i \Delta x^j$$

Note that $M_{i0} = M_{0i}$ (problem 6). Consider case $\Delta s^2 = 0$, so from Eq. (1.1), $\Delta t = \Delta r = \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2}$. Then,

$$\Delta\bar{s}^2 = M_{00}\Delta r^2 + 2M_{0i}\Delta x^i \Delta r + M_{ij}\Delta x^i \Delta x^j$$

which is Eq. (1.3).

b) Since $\Delta\bar{s}^2 = 0$ in Eq. (1.3) for any $\{\Delta x^i\}$, replace Δx^i by $-\Delta x^i$ in Eq. (1.3) and subtract the resulting equations from Eq. (1.3) to establish that $M_{0i} = 0$ for $i = 1, 2, 3$.

We have set $\Delta s^2 = 0$ and it followed, based upon the universality of the speed of light, that $\Delta\bar{s}^2 = 0$. Note that changing Δx^i to $-\Delta x^i$ does not change Δr nor Δs. So that's why $\Delta\bar{s}^2 = 0$ in Eq. (1.3).

The only term in Eq. (1.3) to change sign when changing Δx^i to $-\Delta x^i$ is the $2M_{0i}\Delta x^i \Delta r$ term. The final term doesn't because changing Δx^i to

$-\Delta x^i$ also changes Δx^j to $-\Delta x^j$; the i is just a dummy index. So when we subtract from Eq. (1.3) the following

$$\Delta \bar{s}^2 = M_{00}\Delta r^2 - 2M_{0i}\Delta x^i \Delta r + M_{ij}\Delta x^i \Delta x^j$$

we're left with

$$0 = 4M_{0i}\Delta x^i \Delta r.$$

This must be true for arbitrary Δx^i so $M_{0i} = 0$. QED.

c) Derive Eq. (1.4)b

Required to show:

$$M_{ij} = -M_{00}\delta_{ij}, \quad (i, j = 1, 2, 3).$$

Adding to Eq. (1.3) the following

$$0 = \Delta \bar{s}^2 = M_{00}\Delta r^2 - 2M_{0i}\Delta x^i \Delta r + M_{ij}\Delta x^i \Delta x^j$$

gives,

$$0 = M_{00}\Delta r^2 + M_{ij}\Delta x^i \Delta x^j \tag{1.12}$$

Suppose, $\Delta x = \Delta r, \Delta y = \Delta z = 0$. Substituting into (1.12) then gives $M_{00} = -M_{11}$. Or, when $\Delta y = \Delta r, \Delta x = \Delta z = 0$, we see that $M_{00} = -M_{22}$. Similarly, $M_{00} = -M_{33}$. To see that the off-diagonal terms are zero, note that it's also possible that $\Delta x = \Delta y = \Delta r/\sqrt{2}$ and $\Delta z =$. Substitution into (1.12) gives that

$$0 = (M_{12} + M_{21})\Delta r/2 = \Delta r M_{12} = 0$$

Similarly, $M_{13} = 0 = M_{23}$. In summary,

$$M_{ij} = -M_{00}\delta_{ij}, \quad (i, j = 1, 2, 3).$$

which is Eq. (1.4)b. QED.

1.14.18 a) Show that velocity parameters add linearly, b) apply to a specific problem

Define the velocity parameter W through $w = \tanh(W)$.

Want to show the velocity addition law,

$$w' = \frac{u + w}{1 + wu}$$

implies linear addition of velocity parameters. Simply substitute the definition of velocity parameter,

$$w' = \frac{\tanh(U) + \tanh(W)}{1 + \tanh(U)\tanh(W)} \tag{1.13}$$

$$= \frac{(\tanh(U) + \tanh(W))\cosh(W)\cosh(U)}{\cosh(W)\cosh(U) + \sinh(U)\sinh(W)} \tag{1.14}$$

The numerator can be written as,

$$N = \sinh(W)\cosh(U) + \cosh(W)\sinh(U)$$

so that

$$w' = \frac{\sinh(W)\cosh(U) + \cosh(W)\sinh(U)}{\cosh(W)\cosh(U) + \sinh(U)\sinh(W)}$$

The following identities are useful:

$$\begin{aligned}
\cosh(a)\cosh(b) &= \left(\frac{\exp(a) + \exp(-a)}{2}\right)\left(\frac{\exp(b) + \exp(-b)}{2}\right) \\
&= \frac{\exp(a+b) + \exp(-(a+b))}{4} + \frac{\exp(a-b) + \exp(-(a-b))}{4} \\
&= \frac{\cosh(a+b)}{2} + \frac{\cosh(a-b)}{2} \qquad (1.15)
\end{aligned}$$

$$\begin{aligned}
\sinh(a)\sinh(b) &= \left(\frac{\exp(a) - \exp(-a)}{2}\right)\left(\frac{\exp(b) - \exp(-b)}{2}\right) \\
&= \frac{\exp(a+b) + \exp(-(a+b))}{4} - \frac{\exp(a-b) + \exp(-(a-b))}{4} \\
&= \frac{\cosh(a+b)}{2} - \frac{\cosh(a-b)}{2} \qquad (1.16)
\end{aligned}$$

$$\begin{aligned}
\sinh(a)\cosh(b) &= \left(\frac{\exp(a) - \exp(-a)}{2}\right)\left(\frac{\exp(b) + \exp(-b)}{2}\right) \\
&= \frac{\exp(a+b) - \exp(-(a+b))}{4} + \frac{\exp(a-b) - \exp(-(a-b))}{4} \\
&= \frac{\sinh(a+b)}{2} - \frac{\sinh(a-b)}{2} \qquad (1.17)
\end{aligned}$$

Using (1.15) and (1.16) the denominator above simplifies to $D = \cosh(U + W)$. Using (1.17) the numerator simplifies to $N = \sinh(U + W)$. So,

$$w' = \tanh(U + W)$$

which reveals that we can linearly add velocity parameters, then apply tanh to reduce the final parameter to the final velocity.

b Velocity of 2nd star relative to first, $u_2 = 0.9$. Velocity of nth star relative to (n-1)th, $u_n - u_{n-1} = 0.9$. So the nth star relative to the first is,

$$u'_N = \tanh[(N-1)U]$$

where $0.9 = \tanh(U)$.

1.14.19 a) Lorentz Transformation using velocity parameter

$$\begin{aligned} \bar{t} &= \gamma t - \gamma v x \qquad (1.18) \\ \overline{x} &= -\gamma v t + \gamma x \\ \overline{y} &= y \\ \overline{z} &= z \end{aligned}$$

Let, $v = \tanh(V)$. Note that the Lorentz factor also simplifies,

$$\begin{aligned} \gamma &\equiv \frac{1}{\sqrt{1-v^2}} \\ &= \left(1 - \tanh^2(V)\right)^{-1/2} \\ &= \left(\frac{\cosh^2(V)}{\cosh^2(V) - \sinh^2(V)}\right)^{1/2} \\ &= \pm\cosh(V) \qquad (1.19) \end{aligned}$$

I'm not sure why we always take the positive root in the Lorentz factor.

The final equality follows from the following identity, which is stated without proof in b).

$$\begin{aligned} \cosh^2(V) - \sinh^2(V) &= \left(\frac{\exp(V) + \exp(-V)}{2}\right)^2 - \left(\frac{\exp(V) - \exp(-V)}{2}\right)^2 \\ &= \left(\frac{\exp(2V) + \exp(-2V) + 2}{4}\right) - \left(\frac{\exp(2V) + \exp(-2V) - 2}{4}\right) \\ &= 1 \qquad (1.20) \end{aligned}$$

Substituting $v = \tanh(V)$ and (1.19) into (1.18) gives the desired result,

$$\begin{aligned} \bar{t} &= \cosh(V)\, t - \sinh(V)\, x \\ \bar{x} &= -\sinh(V)\, t + \cosh(V)\, x \\ \bar{y} &= y \\ \bar{z} &= z \end{aligned} \tag{1.21}$$

1.14.19 b) invariance of the interval using velocity parameter

The given identity is derived above (1.20). Invariance of the interval follows from straightforward substitution into (1.21).

$$\begin{aligned} \Delta\bar{s}^2 &= -\Delta\bar{t}^2 + \Delta\bar{x}^2 + \Delta\bar{y}^2 + \Delta\bar{z}^2 \\ &= -(\cosh(V)\Delta t - \sinh(V)\Delta x)^2 + (-\sinh(V)\Delta t + \cosh(V)\Delta x)^2 + \Delta y^2 + \Delta z^2 \\ &= \Delta s^2 \end{aligned} \tag{1.22}$$

In the final equality, the cross terms cancelled directly while the squared terms simplified with the identity (1.20).

1.14.19 c) analogy between Lorentz transformation using velocity parameter and Euclidean coordinate transformation

Hyperbolic trigonometric functions replace regular trigonometric functions, *but* the sign changes for the sine term in the Euclidean coordinate transformation and not the sinh term of the Lorentz transformation.

The analog to the interval Δs^2 is the squared distance to the origin.

The analog to the invariant hyperbolae are circles. These could be used to calibrate axes of the rotated Euclidean frame.

1.14.20 Lorentz tranformation in matrix form

$$\bar{\mathbf{x}} = \mathbf{A}x$$

where

$$\overline{\mathbf{x}} = \begin{bmatrix} \bar{t} \\ \overline{x} \\ \overline{y} \\ \overline{z} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} t \\ x \\ y \\ z \end{bmatrix}$$

and

$$\mathbf{A} = \begin{bmatrix} \cosh(V) & -\sinh(V) \\ -\sinh(V) & \cosh(V) \end{bmatrix}$$

1.14.21 a) Timelike separated events can be transformed to occur at the same point.

1.15 Additional thoughts

I think it's worth mentioning that the Lorentz transformation, which is linear by construction, transforms lines to lines. This is easily verified by substituting the equation for a line in $\mathcal{O}$ and confirming that it's also a line in $\overline{\mathcal{O}}$.

It's also worth pointing out that a tangent line to a curve in $\mathcal{O}$ remains a tangent line in $\overline{\mathcal{O}}$. Of course it would be quite strange if this were not true, but on the other hand it was not immediately obvious to me that it holds.

Chapter 2

Vector Analysis in Special Relativity

2.1 Definition of a vector

Buried in footnote 2 on p. 35 is an important notational point.

2.2 Vector algebra

Eq. (2.10) introduces a strange notational twist. Apparently enclosing the vectors $\vec{e_\alpha}$ with parentheses and writing a superscript β implies that we are forming a tensor from the set of these vectors?

$$(\vec{e_\alpha})^\beta = \delta^\beta_\alpha$$

There's no comment to explain this. Earlier the author explained that the superscript notation will become clear when he introduces differential geometry. For now I just note that the RHS is the Kronecker delta, which is a second-rank tensor.

Eq. (2.18) is described as a key formula. Exercise 2.11c is to verify it.

2.4 The four-momentum

Typo on p. 42, in the example, $p^1 = m\mathbf{v}(1 - v^2)^{-1/2}$ should be $p^1 = mv(1 - v^2)^{-1/2}$.

2.9 Exercises

2 a) α is the dummy index. One equation.
b) ν is the dummy index. $\overline{\mu}$ is the free index. Four equations.

d) ν and μ are free indices, and there are 16 equations. Although the indices are repeated, they're not repeated in the same factor, and one is not superscript.

3 Prove Eq. (2.5). There's nothing to prove really. It follows immediately from the definition and notation conventions. In particular, the LHS involves a sum over all values of the dummy index $\beta \in \{0, 1, 2, 3\}$, see p. 34. The RHS merely spells this out, with the convention that Roman indices like i take all values $i \in \{1, 2, 3\}$.

4 a) $-6\vec{A} \rightarrow_{\mathcal{O}} (-30, 6, 0, -6)$

5 a) Show that the basis vectors are linearly independent. Start with a general linear combination a_μ,

$$0 = a_\mu (\vec{e}_\alpha)^\mu = a_\mu \delta^\mu_\alpha$$

Start with the first component, $\alpha = 0$. The equation above is $0 = a_0 \times 1$, so a_0 must be zero. Similarly for the other components. Since this trivial solution is the only solution, the basis vectors must be linearly independent.

5 b) The given set is not linearly independent, since the linear combination $(-5, -3, +2, 1)$ gives the zero vector.

6 As in Fig. 1.5, the $\bar{t}$ and $\overline{x}$ axes are tilted at an angle ϕ relative to their $\mathcal{O}$ frame counterparts and toward the world line of the line ray $t = x$. The basis vectors are parallel to these $\overline{\mathcal{O}}$ axes. Here $\tan(\phi) = 0.6$. For the $\overline{\overline{\mathcal{O}}}$ the axes will be tilted even further toward $t = x$. The angle of this basis vectors θ can be computed as

$$\tan(\theta) = \tanh(2 \text{arctanh}\,(0.6))$$

7 a) Verify Eq. (2.10). As mentioned above, this is a strange notational twist. If we write the basis vectors as row vectors as in Eq. 2.9, then the set form a matrix, and the matrix element is unity when row and column numbers are equal, and zero otherwise, *i.e.* the identity matrix. The RHS of Eq. 2.10 can of course be written as the identity matrix.

7 b) I've always thought of Eq. 2.11 as the definition of the vector, so it seems to me a tautology, rather than something to prove. Perhaps it's worth stating the result in words. If you use the components of the vector, $\vec{A}$ to form the linear combination of basis vectors $\vec{e}$, *i.e.*

$$A^\alpha \vec{e}_\alpha$$

then you, of course, recover the vector $\vec{A}$. In particular, for the first component, $\alpha = 0$, the first component A^0 multiplies all the basis vectors, but only the first one $\vec{e}_0$ contributes since the other basis vectors are all zero in the first component. Similarly for the other components.

8 a) The zero vector has the same components in all reference frames. This follows immediately from the use of a linear transformation to go between reference frames. See p. 35, and Eq. 2.7 for the definition of the general (4-) vector and the linear transformation.

8 b) If two vectors have equal components in one frame their components are equal in all frames. My first thought is that if their components are equal in a given frame, then they're the same vector. By the definition of a vector, they are invariant under coordinate transformation. So their components are equal in all other frames. But that doesn't use 8a.

Using 8a, one could subtract the two equal vectors, giving the zero vector in that frame. Under coordinate transformation, this difference vector remains they zero vector. Thus their components must be equal in any other frame.

9 There are 16 terms to write out, which is too much work. It seems convincing enough to me to note that for each term in the sum on the LHS, there is a corresponding term on the RHS. In general these terms look like,

$$\Lambda^{\bar{\alpha}}_{\beta} A^\beta \vec{e}_{\bar{\alpha}}$$

Of course the order of summation doesn't matter for a finite sum. Substituting specific values for the dummy indices might make this more clear, say $\overline{\alpha} = 0, \beta = 1$.

10 Prove Eq. (2.13) from

$$A^{\alpha}(\Lambda^{\overline{\beta}}_{\alpha}\vec{e}_{\overline{\beta}} - \vec{e}_{\alpha}) = 0$$

Choosing any A^{α} with only one non-zero entry, like $(1, 0, 0, 0)$, or $(10, 0, 0, 0)$, shows straight away that

$$\Lambda^{\overline{\beta}}_{0}\vec{e}_{\overline{\beta}} = \vec{e}_{0}$$

and similarly $(0, 1, 0, 0)$, or $(0, 2, 0, 0)$, shows straight away

$$\Lambda^{\overline{\beta}}_{1}\vec{e}_{\overline{\beta}} = \vec{e}_{1}.$$

So repeating this argument gives the result for the other two basis vectors.

Perhaps more instructive is to note that this result works for more general situations. The quantity inside the parentheses is a set of 4 different vectors $\vec{v}_{\alpha}$,

$$(\Lambda^{\overline{\beta}}_{\alpha}\vec{e}_{\overline{\beta}} - \vec{e}_{\alpha}) = \vec{v}_{\alpha}$$

Then view the components of A^{α} as the components of a linear combination of this vector $\vec{v}_{\alpha}$. Now it's clear that the RHS is not just the number zero, but the 4-vector $(0, 0, 0, 0)$. The linear combination of the set of $\vec{v}_{\alpha}$ must sum to the zero vector for arbitrary components of the linear combination. If the first three led to a non-zero vector,

$$\sum_{\alpha=0}^{2} \vec{v}_{\alpha} = (2, 4, 6, 8)$$

then A^3 would have to be chosen so bring this to zero. For example, if $\vec{v}_3 = (1, 2, 3, 4)$ one would have to choose $A^3 = -2$. But since A^{α} was arbitrary so then choosing $A^3 = +2$ would violate the equality. So this means that the only way it could work is if

$$\sum_{\alpha=0}^{2} \vec{v}_{\alpha} = 0$$

and $\vec{v}_3 = 0$. One can now repeat this argument for the $\sum_{\alpha=0}^{1} \vec{v}_\alpha$ etc. and show that all $\vec{v}_\alpha$ are the zero 4-vector. And the result Eq. (2.13) holds.

11 (a) Matrix of $\Lambda^{\nu}_{\bar{\mu}}(-\mathbf{v})$. Exercise 1.20 was to put the Lorentz transformation in matrix form. Note that $\sinh(-V) = -\sinh(V)$, $\cosh(-V) = \cosh(V)$. So we only have to change the sign of the $\sinh(V)$ elements,

$$\mathbf{\Lambda} = \begin{bmatrix} \cosh(V) & \sinh(V) & 0 & 0 \\ \sinh(V) & \cosh(V) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

where $v = \tanh(V)$.

(b) $A^{\bar{\alpha}}$ for all $\bar{\alpha}$.

$$A^{\bar{0}} = \cosh(V)A^0 - \sinh(V)A^1 \tag{2.1}$$

$$A^{\bar{1}} = -\sinh(V)A^0 + \cosh(V)A^1 \tag{2.2}$$

$$A^{\bar{2}} = A^2 \tag{2.3}$$

$$A^{\bar{3}} = A^3 \tag{2.4}$$

(c) Verify Eq. (2.18). Written out in matrix form Eq. (2.18) becomes,

$$\begin{bmatrix} \cosh(V) & \sinh(V) & 0 & 0 \\ \sinh(V) & \cosh(V) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cosh(V) & -\sinh(V) & 0 & 0 \\ -\sinh(V) & \cosh(V) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

To show this it's useful to use the hyperbolic function identity,

$$\cosh^2(x) - \sinh^2(x) = 1.$$

Eq. (2.18) follows immediately from matrix multiplication. This identity is easy to derive, and can be found at `http://en.wikipedia.org/wiki/Hyperbolic_function#Similarities_to_circular_trigonometric_functions` along with other properties.

(d) The Lorentz transformation matrix from $\overline{\mathcal{O}}$ to $\mathcal{O}$ is just the matrix in (a). Since $\overline{\mathcal{O}}$ is moving to toward increasing x with velocity v with respect to $\mathcal{O}$, then from $\overline{\mathcal{O}}$ point of view $\mathcal{O}$ is moving toward increasing x with velocity $-v$.

(e) $A^{\overline{\alpha}}$ for all $\overline{\alpha}$.

$$A^0 = \cosh(V)A^{\overline{0}} + \sinh(V)A^{\overline{1}} = A^0 \tag{2.5}$$
$$A^1 = +\sinh(V)A^{\overline{0}} + \cosh(V)A^{\overline{1}} = A^1 \tag{2.6}$$
$$A^2 = A^{\overline{2}} = A^2 \tag{2.7}$$
$$A^3 = A^{\overline{3}} = A^3 \tag{2.8}$$

Relation to Eq. (2.18): Multiplying the vector $\vec{A}$ on the left by the Lorentz transformation matrix $\mathbf{\Lambda}(\mathbf{v})$ gives the components in the $\overline{\mathcal{O}}$ frame, $A^{\overline{\alpha}} = \Lambda^{\overline{\alpha}}_{\beta}(\mathbf{v})A^{\beta}$. Multiplying this vector on the right by the Lorentz transformation matrix $\mathbf{\Lambda}(-\mathbf{v})$ should return the vector to the $\mathcal{O}$ frame. And indeed it does, when we use Eq. (2.18) in the final step below:

$$\Lambda^{\nu}_{\overline{\alpha}}(-\mathbf{v})A^{\overline{\alpha}} = A^{\nu} \tag{2.9}$$
$$\Lambda^{\nu}_{\overline{\alpha}}(-\mathbf{v})\Lambda^{\overline{\alpha}}_{\beta}(\mathbf{v})A^{\beta} = A^{\nu} \tag{2.10}$$
$$\delta^{\nu}_{\beta}A^{\beta} = A^{\nu} \tag{2.11}$$

(f) Verify that the order applying the transformations doesn't matter. Physically we know this must be true. Mathematically it works out because if we repeat (c) with the matrices in the opposite order, we get the same result:

$$\begin{bmatrix} \cosh(V) & -\sinh(V) & 0 & 0 \\ -\sinh(V) & \cosh(V) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cosh(V) & \sinh(V) & 0 & 0 \\ \sinh(V) & \cosh(V) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

(g) Establish that

$$\vec{e}_{\alpha} = \delta^{\nu}_{\alpha}\vec{e}_{\nu}$$

I find this a rather strange question. From the definition of the Kronecker delta function, Eq. (1.4c), the result is immediately obvious. Another way to see this is that the Kronecker delta can be written as the identity matrix. And of course, writing the vector on the RHS as a column vector, multiplying by the identity matrix, gives back the original vector.

12 (b) Remember not to add the velocities linearly, but to use the Einstein law of composition of velocities Eq. (1.13), or use the velocity parameters introduced in Exercise 1.18.

(c) Note that the definition of the magnitude of the vector is analogous to the interval introduced in Chapter 1, see Eq. (2.24).

$$\vec{A}^2 = -0^2 + (-2)^2 + 3^2 + 5^2 = 38.$$

(d) The magnitude should be independent of the reference frame, because of the invariance of the interval.

13

(a) Transformation of coordinates from $\mathcal{O}$ to $\overline{\overline{\mathcal{O}}}$ is can be constructed in two steps. First transform to $\overline{\mathcal{O}}$,

$$A^{\overline{\gamma}} = \Lambda^{\overline{\gamma}}(\mathbf{v})_\mu A^\mu.$$

Then transform from $\overline{\mathcal{O}}$ to $\overline{\overline{\mathcal{O}}}$,

$$A^{\overline{\overline{\alpha}}} = \Lambda^{\overline{\overline{\alpha}}}_{\overline{\gamma}}(\mathbf{v}')(\Lambda^{\overline{\gamma}}(\mathbf{v})_\mu A^\mu).$$

So the Lorentz transformation from $\mathcal{O}$ to $\overline{\overline{\mathcal{O}}}$ is

$$\Lambda^{\overline{\overline{\alpha}}}_\mu = \Lambda^{\overline{\overline{\alpha}}}_{\overline{\gamma}}(\mathbf{v}')\ \Lambda^{\overline{\gamma}}(\mathbf{v})_\mu.$$

(b) I thought we just did show that Eq. (2.41) was the matrix product of the two individual Lorentz transformations. Maybe he means write it out in matrix form? I'm not sure what he's looking for.

(c) The was an important exercise for me because I learned that the Lorentz transformation matrix did not have to be symmetric when there are velocity components in two directions.

$$\Lambda^{\overline{\overline{\alpha}}}_{\mu} = \begin{bmatrix} \gamma(v)\gamma(v') & -\gamma(v)v & -\gamma(v)\gamma(v')v' & 0 \\ -\gamma(v)v\gamma(v') & \gamma(v) & \gamma(v)v\gamma(v')v' & 0 \\ -\gamma(v')v' & 0 & \gamma(v') & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

(d) Show that the interval is invariant under the above transformation.

(e) Show that the order matters in constructing the Lorentz transformation as in (a), *i.e.*

$$\Lambda^{\overline{\overline{\alpha}}}_{\overline{\gamma}}(\mathbf{v})\ \Lambda^{\overline{\gamma}}(\mathbf{v}')_{\mu} \neq \Lambda^{\overline{\overline{\alpha}}}_{\overline{\gamma}}(\mathbf{v}')\ \Lambda^{\overline{\gamma}}(\mathbf{v})$$

Using the example from (c), the LHS of the above would be,

$$\text{LHS} = \begin{bmatrix} \gamma(v')\gamma(v) & -\gamma(v')v' & -\gamma(v')\gamma(v)v & 0 \\ -\gamma(v')v'\gamma(v) & \gamma(v') & \gamma(v')v'\gamma(v)v & 0 \\ -\gamma(v)v & 0 & \gamma(v) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \neq \Lambda^{\overline{\overline{\alpha}}}_{\mu}$$

Comparison with the matrix in (c) shows it's different. This is surprising if we think in a Galilean way. However, mathematically we know in general that matrix multiplication is not commutative, `http://en.wikipedia.org/wiki/Matrix_multiplication#Common_properties`. Physically we know that the Lorentz transformation results in the axes tilting toward the $t = x$ line, as in Fig. 1.5. The order of rotations matters. For example, rotating the globe 90° to the east about the polar axis, then 45° clockwise about the axis through the Equator and 90°W and 90°E, puts the coordinates $0°\text{N}, 0°\text{E}$ where the South Indian Ocean used to be. But performing the same rotations in the opposite order leaves the coordinates $0°\text{N}, 0°\text{E}$ on the old Equatorial plane.

14 (a) $v = -3/5$ in the positive z direction. The off-diagonal term gives the direction, $-v\gamma = 0.75$, and the diagonal term gives $\gamma = 1.25$. One can confirm that $\gamma = 1/\sqrt{1 - v^2}$, once v is found.

(b) Since it's a Lorentz transformation, the inverse should be obtained by from the Lorentz transformation from $\overline{\mathcal{O}}$ back to $\mathcal{O}$.

$$\Lambda(-\mathbf{v}) = \begin{bmatrix} 1.25 & 0 & 0 & -0.75 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -0.75 & 0 & 0 & 1 \end{bmatrix}$$

And matrix multiplication confirms this is the inverse.

(c)

$$\begin{bmatrix} 1.25 & 0 & 0 & -0.75 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -0.75 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1.25 \\ 2 \\ 0 \\ -0.75 \end{bmatrix}$$

15 (a) The particle 3-velocity is $\mathbf{v} = (v, 0, 0)$. In the frame moving with the particle, the 4-velocity is $\vec{e}_0$, so $\vec{A} \to_{\overline{\mathcal{O}}} (1, 0, 0, 0)$. The Lorentz transformation back to the $\mathcal{O}$ frame is

$$\Lambda(-\mathbf{v}) = \begin{bmatrix} \gamma(v) & v\,\gamma(v) & 0 & 0 \\ v\,\gamma(v) & \gamma(v) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

So $\vec{A}$ in the $\mathcal{O}$ frame has components $\vec{A} \to_{\mathcal{O}} (\gamma(v), v\gamma(v), 0, 0)$.

(b) For general particle 3-velocity is $\mathbf{v} = (u, v, w)$. Let's start with a slightly less general 3-velocity is $\mathbf{v} = (u, v, 0)$ to make the algebra easier. One could rotate through an angle θ to a frame where $\mathbf{v} = (|\mathbf{v}|, 0, 0)$. Here θ is such that

$$\begin{bmatrix} u' \\ v' \end{bmatrix} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} |\mathbf{v}| \\ 0 \end{bmatrix}$$

Now we have the situation as in (a) so we can apply the Lorentz transformation back to the $\mathcal{O}$ frame

$$\Lambda(-\mathbf{v}) = \begin{bmatrix} \gamma(|\mathbf{v}|) & |\mathbf{v}|\,\gamma(|\mathbf{v}|) & 0 & 0 \\ |\mathbf{v}|\,\gamma(|\mathbf{v}|) & \gamma(|\mathbf{v}|) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

So $\vec{A}$ in a frame moving with the $\mathcal{O}$ frame but rotated through θ has components $\vec{A} \to_{\mathcal{O}} (\gamma(|\mathbf{v}|), |\mathbf{v}|\gamma(|\mathbf{v}|), 0, 0)$. Finally we rotate through $-\theta$ to obtain $\vec{A}$ in the $\mathcal{O}$ frame

$$\vec{A} \to_{\mathcal{O}} (\gamma(|\mathbf{v}|), |\mathbf{v}|\gamma(|\mathbf{v}|)\cos(\theta), |\mathbf{v}|\gamma(|\mathbf{v}|)\sin(\theta), 0) \quad (2.12)$$
$$= (\gamma(|\mathbf{v}|), u\gamma(|\mathbf{v}|), \gamma(|\mathbf{v}|)v, 0) \quad (2.13)$$

Finally, there's no reason for the z component to behave differently, so we can generalize this. For general particle 3-velocity is $\mathbf{v} = (u, v, w)$, the 4-velocity is

$$\vec{A} \to_{\mathcal{O}} (\gamma(|\mathbf{v}|), |\mathbf{v}|\gamma(|\mathbf{v}|)\cos(\theta), |\mathbf{v}|\gamma(|\mathbf{v}|)\sin(\theta), 0) \quad (2.14)$$
$$= (\gamma(|\mathbf{v}|), u\gamma(|\mathbf{v}|), v\gamma(|\mathbf{v}|), w\gamma(|\mathbf{v}|)) \quad (2.15)$$

where

$$|\mathbf{v}| = \sqrt{u^2 + v^2 + w^2}.$$

(c) Starting with the 4-velocity components $\{U^\alpha\}$, one can write the 3-velocity,

$$\mathbf{v} = (U^1/\gamma, U^2/\gamma, U^3/\gamma)$$

where $\gamma \equiv 1/\sqrt{1 - \mathbf{v}\cdot\mathbf{v}} = U^0$.

(d) Applying the above formula, if the 4-velocity is given as $(2, 1, 1, 1)$ then the 3-velocity is $\mathbf{v} = (1/2, 1/2, 1/2)$. Note the magnitude of the 4-velocity is $-4 + 3 = -1$, making it a legitimate example.

16 Particle moves with speed w, say along the $x-$axis, in a reference frame $\overline{\mathcal{O}}$ moving along the $x-$axis with speed v. Deriving Einstein's velocity addition law from a Lorentz transformation of the particle's 4-velocity.

The particle's 4-velocity in reference frame $\overline{\mathcal{O}}$, $U \to_{\overline{\mathcal{O}}} (\gamma(w), \gamma(w)\, w, 0, 0)$. Lorentz transformation from $\overline{\mathcal{O}}$ to $\mathcal{O}$

$$\Lambda(-v) = \begin{bmatrix} \gamma(v) & v\,\gamma(v) & 0 & 0 \\ v\,\gamma(v) & \gamma(v) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

So the 4-velocity is, $U \rightarrow_{\mathcal{O}} (\gamma(w)\gamma(v)+vw\gamma(w)\gamma(v), v\gamma(v)\gamma(w)+w\gamma(v)\gamma(w), 0, 0)$. Converting this to the 3-velocity using the formula in 15c,

$$v^x = \frac{U^x}{U^0} \tag{2.16}$$
$$= \frac{\gamma(v)\,\gamma(w)(v+w)}{\gamma(v)\,\gamma(w)(1+v\,w)} \tag{2.17}$$
$$= \frac{(v+w)}{(1+v\,w)} \tag{2.18}$$

17 (a) Prove that any timelike vector $\vec{U}$ for which $U^0 > 0$ and $\vec{U}\cdot\vec{U} = -1$ is the four-velocity of some world line.

The four-velocity is the $\vec{e}_0$ in the MCRF. If $\vec{U}$ is some world line's four-velocity, then there exists a Lorentz transformation for which $\vec{U} \rightarrow_{\mathcal{O}} (1, 0, 0, 0)$. Let's see if that's possible for the given vector $\vec{U}$.

The coordinate system can be rotated so that $U^\alpha = (U^0, u, 0, 0)$, just to make the algebra simpler. Now apply an arbitrary Lorentz transformation

$$\begin{bmatrix} \gamma(v) & -v\,\gamma(v) & 0 & 0 \\ -v\,\gamma(v) & \gamma(v) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} U^0 \\ u \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

for some v and $\gamma(v)$. Thus we require

$$1 = \gamma(U^0 - v\,u) \tag{2.19}$$
$$0 = \gamma(u - U^0 v). \tag{2.20}$$

But in general we require $\gamma \geq 1$, so the second equation (2.20) requires $v = u/U^0$. We know $U^0 > 0$ (given) and it follows from the fact that $\vec{U}$ is timelike that $U^0 > u$. So thus $v < 1$. Thus $\gamma(v) > 1$, and most importantly, $\gamma(v) \in \Re$, *i.e.* the Lorentz transformation is possible. Does this required Lorentz transformation also bring the time component to unity?

The algebra can get messy, but simplifies if we use the fact that $\vec{U}\cdot\vec{U} = -1$. Eliminate v in the first equation (2.19) gives

$$\frac{1}{\gamma(v)} = U^0 - \frac{u^2}{U^0} = \frac{1}{U^0}((U^0)^2 - u^2) = \frac{1}{U^0}$$

So

$$\gamma = U^0$$

(b) Use this to prove that for any timelike vector $\vec{V}$ there is a Lorentz frame in which $\vec{V}$ has zero spatial components.

The magnitude of a vector is the interval between the origin and the coordinates of the vector. For a timelike interval the vector is timelike, and vice versa. Timelike intervals can be transformed via a Lorentz transformation to have zero spatial part, see Exercise 1.21. The corresponding vector will have zero spatial components.

If you haven't done Exercise 1.21, you can construct a proof using part 17(a). We are no longer required to make the time part unity; we only require the space part to be zero, *i.e.* (2.20), $0 = \gamma(v)(u - V^0 v)$, where u is now $V^i V_i = u^2$. We no longer have $V^0 > 0$, but that doesn't matter. Because it's a timelike vector we have

$$(V^0)^2 > V^i V_i = u^2$$

So (2.20) implies now that

$$|v| < 1$$

and again, $\gamma(v) \in \Re$, *i.e.* the Lorentz transformation is possible.

18 (a) Sum of two spacelike orthogonal vectors is spacelike.

By definition, orthogonal vectors have $\vec{A} \cdot \vec{B} = 0$, so

$$(\vec{A} + \vec{B}) \cdot (\vec{A} + \vec{B}) = \vec{A} \cdot \vec{A} + \vec{B} \cdot \vec{B} + 2\vec{A} \cdot \vec{B} \tag{2.21}$$

$$= \vec{A} \cdot \vec{A} + \vec{B} \cdot \vec{B} > 0. \tag{2.22}$$

Spacelike vectors have positive magnitude, $\vec{A} \cdot \vec{A} > 0$. So $(\vec{A} + \vec{B})$ is also spacelike.

(b) Timelike vector and null vector cannot be orthogonal.

Timelike vector $\vec{A}$. Let's keep the algebra simple and rotate to a coordinate frame such that the spacepart of the null vector $\vec{N}$ is all in one component,

$$A \to_{\mathcal{O}} (A^0, A^1, 0, 0)$$

The null vector $\vec{N}$ has unknown coordinates in this frame, but

$$A \cdot N = -A^0 N^0 + A^1 N^1$$

19 Stuck!

20 The particle moves in a circle in the $x - y$ plane of radius b, in a clockwise sense when viewed in the direction of decreasing z. The circle translates along the $x-$axis at speed a. It's stated that $|b\omega| < 1$, but the requirement for a realistic particle is actually that $|a+b\omega| < 1$. The 3-velocity is computed directly by differentiating the given equations, $\mathbf{v} \rightarrow_{\mathcal{O}} (\dot{x}, \dot{y}, 0)$, where

$$\dot{x} = a + \omega b \sin(\omega t) \tag{2.23}$$

$$\dot{y} = -\omega b \cos(\omega t) \tag{2.24}$$

The 4-velocity is obtained from the 3-velocity using the formula derived in problem 1.15b.

$$\vec{V} \rightarrow_{\mathcal{O}} (\gamma(v), \dot{x}\gamma(v), \dot{y}\gamma(v), 0)$$

where $v = |\mathbf{v}| = \sqrt{(a + \omega b \sin(\omega t))^2 + (\omega b \cos(\omega t))^2} = \sqrt{a^2 + 2a\omega b \sin(\omega t) + \omega^2 b^2}$.

To obtain the 4-acceleration we require the 4-velocity as a function of proper time, τ, not t, the time in the inertial frame. But remember that the proper time is the time measured by a clock at, say, the origin of the MCRF. Call this frame $\overline{\mathcal{O}}$, and then $\bar{t} = \tau = \overline{x}^{\overline{0}}$. And $t = \Lambda(-v)^0_{\overline{\alpha}} \overline{x}^{\overline{\alpha}}$. For simplicity we choose the MCRF with origin at the particle location, so $\overline{x}^{\overline{\alpha}} \rightarrow_{\overline{\mathcal{O}}} (\tau, 0, 0, 0)$, and $t = \gamma(-v)\tau = \gamma(v)\tau$. Then we obtain the 4-acceleration from the given equations in t and the chain rule,

$$\vec{a} \equiv \frac{d\vec{U}}{d\tau} = \frac{d\vec{U}}{dt}\frac{dt}{d\tau} = \gamma(v)\frac{d\vec{U}}{dt}$$

We now confront the question as to whether or not to let $\gamma(v)$ in this derivative! Stuck!

Let's retreat to safer ground and compute the 4-velocity from the position as a function of t.

21 The motion is hyperbolic in frame $\mathcal{O}$,

$$x^2 - t^2 = a^2 \cosh^2\left(\frac{\lambda}{a}\right) - a^2 \sinh^2\left(\frac{\lambda}{a}\right) = a^2$$

and therefore hyperbolic in all reference frames, $-\bar{t}^2 + \bar{x}^2 = a^2$. The velocity is obtained by differentiating with respect to λ,

$$v = \frac{dx}{dt} = \frac{dx}{d\lambda}/\frac{dt}{d\lambda} = \tanh\left(\frac{\lambda}{a}\right).$$

So we notice that $\left(\frac{\lambda}{a}\right)$ is a velocity parameter for v, see problem 1.18.

The Lorentz transformation to the MCRF can be written in a simple form with the velocity parameter, see problem 1.20:

$$\Lambda = \begin{bmatrix} \cosh\left(\frac{\lambda}{a}\right) & -\sinh\left(\frac{\lambda}{a}\right) \\ -\sinh\left(\frac{\lambda}{a}\right) & \cosh\left(\frac{\lambda}{a}\right) \end{bmatrix}$$

Thus we find the points transform to

$$\begin{bmatrix} \bar{t}(\lambda) \\ \bar{x}(\lambda) \end{bmatrix} = \begin{bmatrix} \cosh\left(\frac{\lambda}{a}\right) & -\sinh\left(\frac{\lambda}{a}\right) \\ -\sinh\left(\frac{\lambda}{a}\right) & \cosh\left(\frac{\lambda}{a}\right) \end{bmatrix} \begin{bmatrix} a\sinh\left(\frac{\lambda}{a}\right) \\ a\cosh\left(\frac{\lambda}{a}\right) \end{bmatrix} = \begin{bmatrix} 0 \\ a \end{bmatrix}$$

The particle always ends up on the $\bar{x}-$axis.

To show that the parameter λ is the proper time, we show that

$$\frac{d\bar{t}}{d\lambda} = 1$$

for a MCRF and any λ. This is a bit subtle, because we want to hold the Lorentz transformation fixed (so hold $\lambda = \lambda_{\text{MCRF}}$ fixed), so that the MCRF is inertial. But we want to let λ vary about $\lambda = \lambda_{\text{MCRF}}$ so we can take the derivative of $\bar{t}(\lambda)$ wrt λ. I've written out this dependence explicitly below:

$$\bar{t}(\lambda) = \cosh\left(\frac{\lambda_{\text{MCRF}}}{a}\right) a\sinh\left(\frac{\lambda}{a}\right) - \sinh\left(\frac{\lambda_{\text{MCRF}}}{a}\right) a\cosh\left(\frac{\lambda}{a}\right)$$

Now differential wrt λ, and evaluate at $\lambda = \lambda_{\text{MCRF}}$ giving,

$$\frac{d\bar{t}}{d\lambda} = \cosh^2\left(\frac{\lambda}{a}\right) - \sinh^2\left(\frac{\lambda}{a}\right) = 1$$

The 4-velocity is

$$\vec{U} \to_{\mathcal{O}} \left(\cosh\left(\frac{\lambda}{a}\right), \sinh\left(\frac{\lambda}{a}\right), 0, 0\right)$$

The 4-acceleration is easy for this problem because we have the 4-velocity as a function of proper time!

$$\vec{a} \equiv \frac{d\vec{U}}{d\tau} = \frac{d\vec{U}}{d\lambda} \to_{\mathcal{O}} \left(\frac{1}{a}\sinh\left(\frac{\lambda}{a}\right), \frac{1}{a}\cosh\left(\frac{\lambda}{a}\right), 0, 0\right)$$

We can check if it's orthogonal to the 4-velocity, as it should be.

$$\vec{U}\cdot\vec{a} = -\frac{1}{a}\sinh\left(\frac{\lambda}{a}\right)\cosh\left(\frac{\lambda}{a}\right) + \frac{1}{a}\sinh\left(\frac{\lambda}{a}\right)\cosh\left(\frac{\lambda}{a}\right) = 0.$$

Is it uniformly accelerating?

$$\vec{a}\cdot\vec{a} = -\frac{1}{a^2}\sinh^2\left(\frac{\lambda}{a}\right) + \frac{1}{a^2}\cosh^2\left(\frac{\lambda}{a}\right) = \frac{1}{a^2}.$$

And a was given as constant, and it's always pointing in the $x-$direction, so it is uniformly accelerating (see definition in problem 2.19).

22 (a) Given 4-momentum, $\vec{p} \to_{\mathcal{O}} (4, 1, 1, 0)$ kg. Find:

Energy in $\mathcal{O}$: In general $\vec{p} \to_{\mathcal{O}} (E, p^1, p^2, p^3)$, so $E = 4$ kg.

3-velocity in $\mathcal{O}$: In general $m\vec{U} = \vec{p}$, where m is the rest mass and $\vec{U}$ is the 4-velocity. And the 3-velocity is related to the 4-velocity as inferred in problem 2.15b, $U^\alpha = \Lambda(-|\mathbf{v}|)^\alpha_0$. So $\vec{p} \to_{\mathcal{O}} (m\gamma, mu\gamma, mv\gamma, mw\gamma)$, where $\mathbf{v} \to_{\mathcal{O}} (u, v, w)$ are the components of the 3-velocity. Note that $E = m\gamma$, and simply dividing through by E gives $\mathbf{v} \to_{\mathcal{O}} (1/4, 1/4, 0)$.

Rest mass:

$$\gamma = \frac{1}{\sqrt{1 - \mathbf{v}\cdot\mathbf{v}}} = \frac{4}{\sqrt{14}}.$$

From which it follows from $E = m\gamma = 4$ that $m = \sqrt{14}$.

(b) We must apply the law of conservation of 4-momentum.

$$\vec{p}_I = \vec{p}_1 + \vec{p}_2 \to_{\mathcal{O}} (5, 0, 1, 0) \text{ kg}$$

By conservation of 4-momentum,

$$\vec{p}_F = \vec{p}_I = \vec{p}_3 + \vec{p}_4 + \vec{p}_5,$$

so

$$\vec{p}_5 = \vec{p}_I - \vec{p}_3 - \vec{p}_4 =\rightarrow_{\mathcal{O}} (3, -1/2, 1, 0) \text{ kg}.$$

Now, like in problem (a), we know the 4-momentum. From an analysis just like in (a), we find the 5th particle has in this same reference frame: $E_5 \rightarrow_{\mathcal{O}} 3$, and $\mathbf{v_5} \rightarrow_{\mathcal{O}} (-1/6, 1/3, 0)$. Finally, the rest mass is $m = \sqrt{31}/2$.

The CM frame is found by finding the Lorentz transformation that transforms the $\vec{p}_F$ to have only a time component,

$$\Lambda^{\bar{\alpha}}_{\beta} p^{\beta} = (\vec{e}_0)^{\bar{\alpha}}$$

This gives the equation for the y-direction,

$$-v\gamma 5 + \gamma = 0$$

So CM has 3-velocity $\mathbf{v} \rightarrow_{\mathcal{O}} (0, 1/5, 0)$.

23 Find the energy given the 3-velocity and rest mass.

First find the 4-momentum, $\vec{p} = m\vec{U} = m\gamma(1, u, v, w)$. And the energy is the time-part of the 4-momentum,

$$E = m\gamma$$

We can find an approximate value of γ from the binomial series, `http://en.wikipedia.org/wiki/Binomial_series`. This is just a Taylor series about $x = 0$. Let $x = \mathbf{v} \cdot \mathbf{v} = v^2$, and $\alpha = -1/2$, so we obtain,

$$\gamma = 1 + \frac{1}{2}v^2 - \frac{3}{8}v^4 + \ldots$$

So

$$E \approx m(1 + \frac{1}{2}v^2 - \frac{3}{8}v^4 + \ldots)$$

i.e. the rest mass, plus the classical kinetic energy, plus a correction of order $O(v^4)$. The correction is 1/2 the kinetic energy when,

$$v = \sqrt{2/3}$$

24 Show that it's impossible for a positron and an electron to annihilate and produce a single $\gamma-$ray.

Apparently particles come and go, but 4-momentum is conserved. Line up the coordinates such that the $x-$axis is aligned with the direction of propagation of the $\gamma-$ray. Then conservation of 4-momentum,

$$\vec{p}_{e^+} + \vec{p}_{e^-} = \vec{p}_\gamma,$$

gives two equations. The time part looks like conservation of energy,

$$p^0_{e^+} + p^0_{e^-} = p^0_\gamma,$$

while the spatial part looks like traditional conservation of momentum,

$$p^1_{e^+} + p^1_{e^-} = p^1_\gamma.$$

It's important to realize that they are not independent, since in a reference frame wherein the electron and positron move with velocities v_{e^-} and v_{e^+}, we have

$$m(\gamma(v_{e^+}) + \gamma(v_{e^-})) = h\nu \tag{2.25}$$

$$m(\gamma(v_{e^+})\, v_{e^+} + \gamma(v_{e^-})\, v_{e^-}) = h\nu, \tag{2.26}$$

where m is the rest mass of the electron and positron and ν is the frequency of the $\gamma-$ray. The only mathematical solution is then $v_{e^-} = v_{e^+} = 1$, which is physically impossible because of their non-zero rest mass. Nothing moves at the speed of light, except electromagnetic radiation and possibly gravity waves if they exist.

It's possible to produce two $\gamma-$rays. Suppose they are travel in opposite directions with equal and opposite momentum in some frame of reference. Then the final total 4-momentum is the null vector. To satisfy momentum conservation we only require that the positron and electron have equal and opposite momentum in the same frame of reference, so $v_{e^+} = -v_{e^-}$ with arbitrary v_{e^+}, which can obviously be satisfied.

25 Doppler shift.

In frame $\mathcal{O}$ photon has 4-momentum

$$\vec{p} \rightarrow_{\mathcal{O}} (h\nu, h\nu\cos(\theta), h\nu\sin(\theta), 0)$$

Transforming to the frame $\overline{\mathcal{O}}$ moving at speed v along the $x-$axis, we apply the Lorentz transformation

$$\Lambda(v) = \begin{bmatrix} \gamma(v) & -v\,\gamma(v) & 0 & 0 \\ -v\,\gamma(v) & \gamma(v) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

to obtain

$$\vec{p} \rightarrow_{\overline{\mathcal{O}}} \begin{bmatrix} \gamma h\nu - v\gamma(v)h\nu\cos(\theta) \\ -v\gamma(v)h\nu + \gamma(v)h\nu\cos(\theta) \\ h\nu\sin(\theta) \\ 0 \end{bmatrix}$$

So the Doppler shift is obtained from the time component, *i.e.* the first component, and can be expressed as,

$$\frac{\overline{\nu}}{\nu} = \gamma(v)(1 - v\cos(\theta)) = \frac{1}{\sqrt{1-v^2}}(1 - v\cos(\theta))$$

as given.

(b)
No Doppler shift occurs when

$$\frac{\overline{\nu}}{\nu} = 1 = \frac{1}{\sqrt{1-v^2}}(1 - v\cos(\theta))$$

or

$$\theta = \arccos\left(\frac{1-\sqrt{1-v^2}}{v}\right)$$

Extra questions: Does this have solutions? For $|v| \ll 1$ use the binomial series to see that $\theta \approx \pi/2$. What's the maximum angle of no Doppler shift? As $v \rightarrow 1, \theta \rightarrow 0$. Show that at $v = 1/2, \theta \approx 74.5°$.

(c)
(2.35) is the frame-invarient expression for energy $\overline{E}$ relative to observer moving with velocity $\vec{U}_{\text{obs}}$,

$$-\vec{p} \cdot U_{\text{obs}} = \overline{E}$$

and (2.38) was just $E = h\nu$. This calculation ends up being exactly the same as above, but allows one to focus on the relevant parts, *i.e.* just the time component. Since

$$\vec{U}_{\text{obs}} \rightarrow_{\mathcal{O}} (\gamma(v), \gamma(v)v, 0, 0)$$

and recall

$$\vec{p} \rightarrow_{\mathcal{O}} (h\nu, h\nu\cos(\theta), h\nu\sin(\theta), 0)$$

so we can immediately find

$$\overline{E} = \gamma(v)h\nu - v\gamma(v)h\nu\cos(\theta).$$

which was the time component of the $\vec{p} \rightarrow_{\overline{\mathcal{O}}}$ found in (a) above.

26 Energy required to accelerate an object with rest mass m from v to δv to first order in δv.

$$E = m\gamma(v) = m\frac{1}{\sqrt{1-v^2}}$$

so the change in energy is just

$$\delta E = m(\gamma(v+\delta v) - \gamma(v)).$$

When $v \ll 1$ the problem is easy. Just differentiate γ wrt v to get the Taylor series approximation

$$\gamma(v+\delta v) - \gamma(v) = \gamma'(v)\delta v + \frac{1}{2}\gamma''(v)\delta v^2 + \ldots$$

where

$$\gamma' = \frac{d\gamma}{dv} = v\gamma^3 \qquad (2.27)$$

$$\gamma'' = \gamma^3 + v3\gamma^2\gamma' \qquad (2.28)$$

So

$$\gamma(v+\delta v) - \gamma(v) = v\gamma^3\delta v + O(\delta v^2)\ldots$$

And so the change in energy is,

$$\delta E \approx mv\gamma^3\delta v.$$

A subtlety arises when v is not small. The coefficient γ'' become large relative to γ', so ignoring the $O(\delta v^2)$ term becomes misleading. The author should have instructed us to check this. In particular,

$$\frac{\gamma''}{\gamma'} = \frac{1}{v} + 3v\gamma^2$$

When $v \ll 1$ we can replace

$$\frac{1}{2}\gamma''\delta v^2 \approx \gamma'\delta v\left(\frac{\delta v}{2v}\right) \ll \gamma'(v)\delta v$$

since we're given that $(\frac{\delta v}{v}) \ll 1$. So we're still justified in ignoring the 2nd term in the Taylor series. But when v is not small we need another approach.

The above argument is not formally correct when v is not small because the higher order terms in the Taylor series can no longer be ignored. Here is one approach.

Write $v = 1 - \epsilon$ where $0 < \epsilon \ll 1$, so we're close to the speed of light. Use $\epsilon \ll 1$ and the Binomial series to simplify γ,

$$\gamma(v) = \frac{1}{\sqrt{\epsilon(2-\epsilon)}} \approx \frac{1}{\sqrt{2\epsilon}}(1 + \epsilon/4),$$

and

$$\gamma(v + \delta v) = \frac{1}{\sqrt{(\epsilon - \delta\epsilon)(2 - \epsilon + \delta\epsilon)}}$$

where $\delta\epsilon = -\delta v$. To simplify the latter we need to consider the case where $|\delta\epsilon| \ll \epsilon$. But this is not so restrictive. Then

$$\gamma(v + \delta v) = \frac{1}{\sqrt{(\epsilon - \delta\epsilon)(2 - \epsilon + \delta\epsilon)}} \approx \frac{1}{\sqrt{2\epsilon}}\left(1 + \frac{\delta\epsilon}{2\epsilon}\right)\left(1 + \frac{\epsilon - \delta\epsilon}{4}\right)$$

To find the perturbation in energy we take the difference,

$$\gamma(v + \delta v) - \gamma(v) \approx \frac{1}{\sqrt{2\epsilon}}\left(\frac{\delta\epsilon}{4}\right)$$

It's clear that as $\epsilon \to 0$, so $v \to 1$,

A simpler and better solution: Write $v = 1 - \epsilon$ where $0 < \epsilon \ll 1$, so we're close to the speed of light.

$$\gamma(v) = \frac{1}{\sqrt{\epsilon(2-\epsilon)}}.$$

Now expand this in a Taylor Series in ϵ:

$$\gamma(v+\delta v)-\gamma(v)=\frac{d\gamma}{d\epsilon}(-\delta\epsilon)+\frac{1}{2}\frac{d^2\gamma}{d\epsilon^2}(-\delta\epsilon)^2+\ldots$$

And

$$\frac{d\gamma}{d\epsilon}=\frac{-\left(1-\frac{\epsilon}{4}\right)}{(2\epsilon)^{3/2}(1-\epsilon/2)^{3/2}}\approx\frac{-\left(1-\frac{\epsilon}{4}\right)\left(1+\frac{3\epsilon}{4}\right)}{(2\epsilon)^{3/2}}$$

where the approximation exploits $0<\epsilon\ll 1$ with the Binomial Series approximation. It's important to check the size of the 2nd derivative relative to the first. We find, again using the Binomial Series,

$$\frac{d^2\gamma}{d\epsilon^2}\approx\frac{-3}{2\epsilon}\frac{d\gamma}{d\epsilon}$$

so we're only justified in ignoring the 2nd term if $|\delta\epsilon|\ll\epsilon$. In this case, the change in energy is

$$\delta E\approx m\frac{1}{(2\epsilon)^{3/2}}\delta v=m\gamma^3\delta v+O(\epsilon)$$

This actually agrees with the result we would have obtained from using the simply Taylor Series above.

We're asked to show that the energy becomes infinite when $v\to 1$. This is easily obtained by noting that γ is finite for $0\le v<1$. However,

$$\lim_{v\to 1}\gamma(v)\to\infty.$$

27 Increasing temperature increases the rest mass.

Object has rest mass, $m(T_0)=10$[kg]. Increasing temperature from T_0 to T by heat flux $\delta Q=100$ J. This must be reflected in an increase in rest mass, since in the MCRF of the object, $U^0=1$ and $mU^0=p^0=E$. So

$$m(T)=m(T_0)[\mathrm{kg}]+\delta Q[\mathrm{J}]/c^2[\mathrm{m}^2/\mathrm{s}^2]=10+1.1\times 10^{-15}[\mathrm{kg}]$$

This problem is interesting to look at from a thermodynamics point of view. The heat flux increases the temperature and enthalpy of the object, which is reflected on a microscopic scale by an increase in the motion, relative to the centre of mass of the object, of the elements (atoms or molecules or sea of electrons depending on the material) composing the object. This motion increases the effective mass of the elements. Say an element has rest mass m_i, then when it has thermal speed v_i it has "relativistic mass"

$$m_{i,\mathrm{rel}} = m_i \gamma(v_i).$$

I found this website, which expands on these ideas `http://en.wikipedia.org/wiki/Massenergy_equivalence`.

28 Boring.

29

$$\frac{d}{d\tau}(\vec{U}\cdot\vec{U}) = \frac{d}{d\tau}\left(-(U^0)^2 + (U^1)^2 + (U^2)^2 + (U^3)^2\right) \tag{2.29}$$

$$= -2U^0\frac{dU^0}{d\tau} + 2U^i\frac{dU^i}{d\tau} \tag{2.30}$$

$$= 2\vec{U}\cdot\frac{d\vec{U}}{d\tau} \tag{2.31}$$

Q.E.D.

30 Four velocity of rocket ship,

$$\vec{U} \rightarrow_{\mathcal{O}} (2, 1, 1, 1)$$

High-velocity cosmic ray with 4-momentum,

$$\vec{P} \rightarrow_{\mathcal{O}} (300, 299, 0, 0) \times 10^{-27}\mathrm{kg}$$

(a) Transform to MCRF of rocket ship. We know from Ex. 2.15, that for general particle 3-velocity is $\mathbf{v} = (u, v, w)$, the 4-velocity is

$$\vec{A} \rightarrow_{\mathcal{O}} (\gamma(|\mathbf{v}|), |\mathbf{v}|\gamma(|\mathbf{v}|)\cos(\theta), |\mathbf{v}|\gamma(|\mathbf{v}|)\sin(\theta), 0) \tag{2.32}$$

$$= (\gamma(|\mathbf{v}|), u\gamma(|\mathbf{v}|), v\gamma(|\mathbf{v}|), w\gamma(|\mathbf{v}|)) \tag{2.33}$$

where

$$|\mathbf{v}| = \sqrt{u^2 + v^2 + w^2}.$$

Inspection of $\vec{U}$ reveals that

$$\begin{aligned} \gamma &= 2 \\ u &= 1/2 \\ v &= 1/2 \\ w &= 1/2 \end{aligned}$$

and $|\mathbf{v}| = \sqrt{u^2 + v^2 + w^2} = \sqrt{3}/2$. Now we need the Lorentz transformation for a reference frame moving with 3-velocity with more than one non-zero component. Up to this point we haven't learned this, and I'm a bit surprised Schutz has thrown this at us now. To lead one through the steps to construct a general Lorentz transformation, I've created supplementary problem R.1 in section 2.10. Here we note that we actually only need the first row of the Lorentz transformation matrix, since we only require $P^{\bar{0}} = E$. This first row must be such that it transforms $\vec{U} \to_{\bar{\mathcal{O}}} (1, 0, 0, 0)$. Thus it must be related to the components of $\vec{U}$ as follows:

$$\Lambda^0_0 = U^0 \quad \Lambda^0_i = -U^i.$$

Applying Λ^0_α to the given $\vec{P}$ gives, $E = 301 \times 10^{-27}$kg in rocket ship frame.

(b)

$$-\vec{P} \cdot \vec{U}_{\text{obs}} = E_{\text{obs}} = 10^{-27}[300\ 299\ 0\ 0] \begin{bmatrix} 2 \\ 1 \\ 1 \\ 1 \end{bmatrix} = 301 \times 10^{-27}\text{kg}$$

(c) Of course (b) was faster. The same computations were performed to get the answer, but in (b) we *only* did the necessary computations.

31 Photon reflects off mirror without changing frequency ν. Angle of incidence is θ.

This appears to be a straightforward application of conservation of 4-momentum, but it fun because it gets us thinking about all 4 components.

Let the mirror lie in the $y-z$ plane, with photon travelling initially in the $x-y$ plane, with angle θ to the $x-$axis. Then the initial 4-momentum of the photon is written

$$\vec{P}_i = (h\nu, \cos(\theta)h\nu, \sin(\theta)h\nu, 0).$$

First let's construct the 4-momentum of the reflected photon $\vec{P}_r$. Since the photon frequency doesn't change, we know instantly the time component,

$$P_r^0 = P_i^0 = h\nu.$$

For a smooth mirror we assume that the momentum transferred is only in the $x-$direction. So then we can also construct the components,

$$P_r^2 = P_i^2 = \sin(\theta)h\nu, \quad P_r^3 = P_i^3 = 0$$

Recall from Eq. (2.37) that the 4-momentum of a photon is orthogonal to itself. This along gives us two possibilities for $P_r^1 = \pm\cos(\theta)h\nu$. For the *reflected* photon, we choose the minus sign. In summary,

$$\vec{P}_r = (h\nu, -h\nu\cos(\theta), h\nu\sin(\theta), 0).$$

By conservation of 4-momentum, we see that the momentum transferred to the mirror must be $\Delta P_m^1 = 2h\nu\cos(\theta)$ in the $x-$direction. How did the mirror acquire $x-$direction momentum without gaining energy? See Supplementary Problem R2 in section 2.10.

If the photon is absorbed, then the momentum transferred to the mirror has three components,

$$\Delta\vec{P}_m = (\Delta E_m, \Delta P_m^1, \Delta P_m^2, 0) = (h\nu, h\nu\cos(\theta), h\nu\sin(\theta), 0),$$

How did the mirror acquire the extra energy $\Delta E_m = h\nu$? See Supplementary Problem R2 in section 2.10.

32 Derive the Compton scattering relationship Eq. 2.43.

Initially the 4-momentum in the particle's initial rest frame $\mathcal{O}$ is

$$\vec{P} \to_{\mathcal{O}} (h\nu_i, h\nu_f, 0, 0) + (m, 0, 0, 0)$$

After the scattering event,

$$\vec{P} \to_{\mathcal{O}} (h\nu_f, h\nu_f \cos(\theta), h\nu_f \sin(\theta), 0) + m(\gamma, v\gamma \cos(\phi), v\gamma \sin(\phi), 0)$$

where v and ϕ are the speed and the angle of the particle's scattered trajectory in the $x-y$ plane relative to the initial direction of the incident photon. Equating the three nonzero components of 4-momentum gives 3 equations for the 3 unknowns ν_f, v, ϕ. In principle one can then solve for ν_f in terms of the other two unknowns, but I found it too tedious to do so.

33 Compton scattering of a cosmic microwave background radiation photon off a cosmic ray (high-energy proton). What's the max frequency of scattered photon?

Very nice problem. At first appears very challenging, but the extreme differences in energy between the two particles simplifies things.

First we note that in the rest frame of the particle, Compton scattering only reduces the frequency and more so for less massive particles (see also supplementary problem R.2 below). So how can Compton scattering increase the energy of the photon?? The increase in energy is revealed via the Doppler shift.

The key simplification in this problem is that the Compton scattering in the frame of the particle has very little effect on frequency.

$$\frac{1}{h\nu_i} = 5000\text{eV}^{-1} \gg \frac{1}{m_p} = 10^{-9}\text{eV}^{-1}.$$

So the angle of the Compton scattering has very little effect on the finally frequency *in the particles initial rest frame.* So in considering the effect of the angle, we need only consider its effect on the Doppler shift.

Now the problem is easy. The Doppler shift in frequency is given in general by Eq. 2.42. Obviously to maximize the frequency in the cosmic ray frame, $\bar{\nu}_i$, we want the photon and cosmic ray traveling in a line in opposite directions, *i.e.* $\theta = \pi$ radians, for which Eq. 2.42 gives

$$h\overline{\nu}_i = h\nu_i \frac{1}{\sqrt{1-v^2}}(1+v) \approx h\nu_i \frac{2}{\sqrt{1-v^2}} = h\nu_i 2 \times 10^9 = 4 \times 10^5\text{eV}.$$

The Doppler shift has made a tremendous increase in frequency! The Compton scattering will make very little difference, so to maximize the scattered frequency in the Sun's frame, choose the Compton scattering angle to maximize the Doppler shift. That is, choose the scattering angle to be π. Eq. 2.43 gives

$$\frac{1}{h\bar{\nu}_f} = \frac{1}{h\bar{\nu}_i} + \frac{2}{m_p} = 0.25 \times 10^{-5} + 2 \times 10^{-9} \approx 0.25 \times 10^{-5} [\mathrm{eV}]^{-1}.$$

Compton scattered caused negligible decrease in energy in the proton's frame. The proton, like the mirror in problem 31, is massive enough to cause little change in frequency of the photon in the proton's frame. See also Supplementary problem R.2. Now Lorentz transform back to the Sun's frame. The photon again gains tremendously from the Doppler shift (that's why we choose the scattering angle to be complete reflection).

$$h\nu_f \approx h\bar{\nu}_f \; 2 \times 10^9 \approx 8 \times 10^{14}\mathrm{eV}.$$

This is a very hard $\gamma-$ray. A pair of 511 keV photons arising from annihilation of an electron and positron are considered to be $\gamma-$rays. This more than a billion times more energetic.

34 These are quite trivial. For example, expand out the dot product in terms of components using the definition in Eq. 2.26, and use the linearity property given by Eq. 2.8,

$$\begin{aligned}(\alpha\vec{A}) \cdot \vec{B} &= -\alpha A^0 B^0 + \alpha A^1 B^1 + \alpha A^2 B^2 + \alpha A^3 B^3 \\ &= \alpha(-A^0 B^0 + A^1 B^1 + A^2 B^2 + A^3 B^3) \\ &= \alpha(\vec{A} \cdot \vec{B}) \end{aligned} \quad (2.34)$$

35 Show that $\vec{e}_{\bar{\beta}}$ obtained from Eq. 2.15,

$$\vec{e}_{\bar{\mu}} = \Lambda^{\nu}_{\bar{\mu}}(-\mathbf{v})\vec{e}_{\nu},$$

obey

$$\vec{e}_{\bar{\alpha}} \cdot \vec{e}_{\bar{\beta}} = \eta_{\bar{\alpha}\bar{\beta}}$$

$$\begin{aligned}\vec{e}_{\bar{\alpha}} \cdot \vec{e}_{\bar{\beta}} &= \Lambda^{\nu}_{\bar{\alpha}}(-\mathbf{v})\vec{e}_{\nu} \cdot \Lambda^{\mu}_{\bar{\beta}}(-\mathbf{v})\vec{e}_{\mu} \\ &= \Lambda^{\nu}_{\bar{\alpha}}\,\Lambda^{\mu}_{\bar{\beta}}\;\vec{e}_{\nu} \cdot \vec{e}_{\mu} \\ &= \Lambda^{\nu}_{\bar{\alpha}}\,\Lambda^{\mu}_{\bar{\beta}}\;\eta_{\nu\mu}\end{aligned}$$

The LHS is a vector expression, and it shouldn't depend upon the orientation of the coordinate axes. So let's rotate the axes so that $\mathbf{v}$ is oriented along the $x-$axis. Then

$$\mathbf{\Lambda}(\mathbf{v}) = \begin{bmatrix} \gamma & -v\gamma & 0 & 0 \\ -v\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Note that $\mathbf{\Lambda}$ is symmetric so we can interchange indices on one without effect,

$$\vec{e}_{\bar{\alpha}} \cdot \vec{e}_{\bar{\beta}} = \Lambda^{\bar{\beta}}_{\mu} \Lambda^{\nu}_{\bar{\alpha}}\, \eta_{\nu\mu}$$

For given $\bar{\alpha} = \bar{\beta}$, the RHS looks like the product of a row of $\Lambda^{\bar{\beta}}_{\mu}$ times a column $\Lambda^{\nu}_{\bar{\alpha}}$. It's easy to see that the result is -1 for $\bar{\alpha} = \bar{\beta} = 0$ and $+1$ for $\bar{\alpha} = \bar{\beta} > 0$. When $\bar{\alpha} \neq \bar{\beta}$, the RHS $= 0$. Q.E.D.

2.10 Rob's supplemental problems

R.1 Suppose the 4-velocity of rocket ship is $\vec{U} \rightarrow_{\mathcal{O}} (2, 1, \sqrt{2}, 0)$ in some reference frame $\mathcal{O}$.

(a) Show that the given $\vec{U}$ is a legitimate 4-velocity. Show that $\vec{V} \rightarrow_{\mathcal{O}} (2, 1, 1, 0)$ is not possible.

(b) Find the 3-velocity in $\mathcal{O}$. Hint: see Ex. 2.15. (You'll need this for (c)).

(c) Find the matrix that rotates of spatial coordinates such that the 3-velocity has only one non-zero component, in say the $x-$direction. What's the matrix that rotates the 4-velocity to have only one nonzero spatial component?

(d) Find the inverse rotation matrices for above. Hint: Think physically and check mathematically, *i.e.* $\mathbf{R_4^{-1} R_4} = \mathbf{I}$

(e) Find the Lorentz transformation from $\mathcal{O}$ to the MCRF of the rocket ship. Confirm that it has the correct effect applied to $\vec{U}$ itself. Hint: The problem here is that we have so far only seen the Lorentz transformation when the 3-velocity has only one non-zero component. Use your rotation matrix from above and its inverse.

Solution:

(a)

$$\vec{U}\cdot\vec{U} = -2^2 + 1^2 + \sqrt{2}^2 = -1$$

which is consistent with Eq. (2.28). On the other hand,

$$\vec{V}\cdot\vec{V} = -2^2 + 1^2 + 1^2 = -2$$

which is inconsistent with Eq. (2.28).

(b) See solution to Ex. 2.15:

$$\mathbf{v} \rightarrow_{\mathcal{O}} (1/2, \sqrt{2}/2, 0)$$

(c) Rotating anticlockwise through angle $\theta = \arccos(1/\sqrt{3})$ aligns the $x-$axis with the 3-velocity. This is accomplished with the matrix $\mathbf{R_3}$,

$$\mathbf{R_3} = \begin{bmatrix} \cos(\theta) & \sin(\theta) & 0 \\ -\sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

For the 4-velocity

$$\mathbf{R_4} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\theta) & \sin(\theta) & 0 \\ 0 & -\sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

(d) To find the inverse of the rotation matrix just change the sign of the angle!

$$\mathbf{R_4^{-1}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) & 0 \\ 0 & \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

(e) The Lorentz transformation for the case $\Lambda(u, v, 0)$ can be built from the above tools. Consider transforming a vector, $\vec{U}$.

$$\begin{aligned} \vec{\bar{U}} &= \mathbf{\Lambda}\vec{U} \\ &= \mathbf{R_4^{-1}\Lambda'(u', 0, 0)R_4}\vec{U} \end{aligned}$$

where

$$\mathbf{\Lambda}'(u',0,0)=\begin{bmatrix}\gamma(u') & -u'\gamma(u') & 0 & 0\\ -u'\gamma(u') & \gamma(u') & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}$$

So this defines the desired Lorentz transformation $\mathbf{\Lambda}(u,v,0)$,

$$\mathbf{\Lambda}(u,v,0)=\begin{bmatrix}\gamma(|\mathbf{v}|) & -u\gamma(|\mathbf{v}|) & -v\gamma(|\mathbf{v}|) & 0\\ -u\gamma(|\mathbf{v}|) & \gamma(|\mathbf{v}|)\cos^2(\theta)+\sin^2(\theta) & (\gamma(|\mathbf{v}|)-1)\cos(\theta)\sin(\theta) & 0\\ -v\gamma(|\mathbf{v}|) & (\gamma(|\mathbf{v}|)-1)\cos(\theta)\sin(\theta) & \gamma(|\mathbf{v}|)\sin^2(\theta)+\cos^2(\theta) & 0\\ 0 & 0 & 0 & 1\end{bmatrix} \tag{2.35}$$

where $|\mathbf{v}|=\sqrt{u^2+v^2}$ and $\theta=\arctan(v/u)$. It's straightforward, albeit a bit tedious, to show that

$$\mathbf{\Lambda}(u,v,0)\vec{U}=\begin{bmatrix}1\\0\\0\\0\end{bmatrix}.$$

R.2 (a)How did the mirror in problem 2.31 acquire $x-$direction momentum without acquiring energy when the photon was reflected?

(b) How did it acquire the energy when the photon was absorbed?

Solution:

(a) The change in 4-momentum is related to the change in 4-velocity of a massive object,

$$\Delta\vec{P}_m=m\Delta\vec{U}=m(\Delta\gamma,\Delta(u\gamma),0,0)=m(\gamma-1,u\gamma,0,0),$$

where the 2nd equality assumes the mirror is initially at rest. Thus the ratio of

$$\frac{\Delta P_m^0}{\Delta P_m^1}=\frac{\Delta E_m}{\Delta(mU^1)}=\frac{1}{u}(1-\sqrt{1-u^2})\approx\frac{u}{2}.$$

The approximation applies in the limit $u\ll 1$ using the binomial series. So the change in energy can be arbitrarily small for a given change in momentum if the change in velocity is correspondingly small. This corresponds to intuition that a more massive mirror would rebound less for a given momentum transfer. I suspect the imposition of "reflection without change in frequency" is an idealization applicable for massive "mirrors". Indeed the next problem, 2.32 covers Compton scattering, wherein a photon reflects off a particle of mass m. In Eq. 2.43 we see that for

$$\frac{m}{h} \gg \nu_i$$

where ν_i is the incident frequency of the photon, the reflected frequency $\nu_f \approx \nu_i$.

(b) For a massive mirror, the energy must have become mostly thermal energy. For a less massive mirror the energy, more the energy would go into the translational kinetic energy of the rebound.

Bibliography

Misner, C. W., K. S. Thorne, and J. A. Wheeler, 1973: *Gravitation*. W. H. Freeman and company. 1279 + XXVI pp.

Schutz, B., 2009: *A first course in General Relativity*. Cambridge University Press. 2nd ed., 393 + XV pp.