Chapter 1

Special relativity

1.1 Fundamental principles of special relativity (SR) theory

Special relativity can be summarized by two fundamental postulates:

1. The principle of relativity (Galileo), which states that no experiment may measure the absolute velocity of an observer.

2. The universality of the speed of light (Einstein), which states that the speed of light is constant when measured from any inertial reference frame.

1.2 Definition of an inertial observer in SR

When we say "observer", what we really mean is a coordinate system. Thus an inertial observer is a coordinate system that meets the following 3 criteria:

1. The distance between two spatial points P_1 and P_2 is independent of time.

2. Time is synchronized and moves at the same rate at all spatial points.

3. At any constant time, space is Euclidean.

It follows from these criteria that the observer must be **unaccelerated**.

1.3 New units

The speed of light, c, is approximately $3.00 \times 10^8 \, \text{ms}^{-1}$ in SI units. However, these units predate relativity, and are very inconvenient. Life becomes easier if we define our units around c, such that $c \equiv 1$.

This can be done by repurposing the meter as a measure of time as well. We thereby define the meter as "the time it takes light to travel 1 meter". Thus the speed of light becomes

$$c = \frac{1\,\mathrm{m}}{1\,\mathrm{m}}.$$

Indeed, it turns out in SR that time is most conveniently measured in distance ($c = 3.00 \times 10^{10}\,\mathrm{cm}$), and in GR mass is as well ($G/c^{-2} = 7.425 \times 10^{-29}\,\mathrm{cm\,g^{-1}}$).

1.4   Spacetime diagrams

1.5   Construction of the coordinates used by another observer

1.6   Invariance of the interval

For two nearby events, we can define the **invariant interval**, defining a 4D Minskowski spacetime:

$$\mathrm{d}s^2 = -(c\,\mathrm{d}t)^2 + \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2\,,$$

or when we set $c \equiv 1$:

$$\mathrm{d}s^2 = -\,\mathrm{d}t^2 + \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2\,. \tag{Schutz 1.1}$$

This notation can be simplified be defining

$$\eta_{\mu\nu} = \mathrm{diag}(-1,1,1,1) = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}; \quad \mathrm{d}s^2 = \sum_{\mu=0}^{3}\sum_{\nu=0}^{3} \eta_{\mu\nu}\,\mathrm{d}x^{\mu}\,\mathrm{d}x^{\nu}$$

When we want to find $\mathrm{d}\bar{s}^2$, we can consider the fact that each of its components, $\mathrm{d}\bar{x}^{\alpha}$, is a linear combination of the components of $\mathrm{d}s^2$,

$$\mathrm{d}\bar{x}^{\alpha} = \sum_{\beta=0}^{3} a_{\alpha\beta}x^{\beta}.$$

Now, when we consider the square of $\mathrm{d}\bar{x}^{\alpha}$, the cross terms make it a quadratic function. Since the sum of four quadratics (the four $\mathrm{d}\bar{x}^{\alpha}$'s) is also a quadratic, we can write $\mathrm{d}\bar{s}^2$ as

$$\mathrm{d}\bar{s} = \sum_{\alpha=0}^{3}\sum_{\beta=0}^{3} M_{\alpha\beta}(\mathrm{d}x^{\alpha})(\mathrm{d}x^{\beta}) \tag{Schutz 1.2}$$

If are talking about light, $\mathrm{d}s^2 = 0$, and so we can say

$$\mathrm{d}s^2 = 0 = -\mathrm{d}t^2 + \mathrm{d}r^2 \implies \mathrm{d}t = \mathrm{d}r$$

Now by looking at Exercise 8 in Section 1.14, we see that

$$\mathrm{d}\bar{s}^2 = M_{00}(\mathrm{d}r)^2$$
$$+ 2\left(\sum_{i=1}^{3} M_{0i}\,\mathrm{d}x^i\right)\mathrm{d}r$$
$$+ \sum_{i=1}^{3}\sum_{j=1}^{3} M_{ij}\,\mathrm{d}x^i\,\mathrm{d}x^j\,, \qquad\qquad \text{(Schutz 1.3)}$$

where

$$M_{0i} = 0 \qquad\qquad \text{(Schutz 1.4a)}$$

and

$$M_{ij} = -(M_{00})\delta_{ij}, \qquad\qquad \text{(Schutz 1.4b)}$$

where δ_{ij} is the Kronecker delta.

1.7 Invariant hyperbolae

1.8 Particularly important results

1.9 The Lorentz transformation

1.10 The velocity-composition law

1.11 Paradoxes and physical intuition

1.12 Further reading

1.13 Appendix: The twin 'paradox' dissected

Consider two twins, Joe and Ed. Joe goes off in a straight line traveling at a speed of $(24/25)c$ for 7 years, as measured on his clock, then instantaneously reverses and returns at half the speed. Ed remains at home. When they return, what is the difference in ages between Joe and Ed?

$\tau_1 = 7\,\text{yr}$. $t_1 = \tau_1\gamma_1$, where $\gamma_1 = \left[1 - \left(\frac{24}{25}\right)^2\right]^{-1/2}$. So $t_1 = 25\,\text{yr}$.

$t_2 = 2t_1 = 50\,\text{yr}$.

$\tau_2 = t_2\gamma_2^{-1}$, where $\gamma_1 = \left[1 - \left(\frac{12}{25}\right)^2\right]^{-1/2}$. So $\tau_2 = 2\sqrt{481}\,\text{yr} \approx 44\,\text{yr}$. Finally, $\tau = \tau_1 + \tau_2 \approx 51\,\text{yr}$, and $t = t_1 + t_2 = 75\,\text{yr}$, so Ed ages $t - \tau \approx 24\,\text{years}$ more than Joe.

1.14   Exercises

1 Convert the following to units in which $c = 1$, expressing everything in terms of m and kg.

(Note that $c = 1 \implies 1 \approx 3 \times 10^8\,\text{m}\,\text{s}^{-1} \approx (3 \times 10^8)^{-1}\text{m}^{-1}\,\text{s}$

(a) $10\,\text{J}$

$$10\,\text{J} = 10\,\text{N}\,\text{m} = 10\,\text{kg}\,\text{m}^2\,\text{s}^{-2} \approx 10\,\text{kg}\,\text{m}^2\,\text{s}^{-2} \cdot ((3 \times 10^8)^{-1}\text{m}^{-1}\,\text{s})^2$$

$$\approx 10\,\text{kg}(3 \times 10^8)^{-2} = 10\,\text{kg}\left(\frac{1}{9} \times 10^{-16}\right) \approx 1.11 \times 10^{-16}\,\text{kg}$$

(b) $100\,\text{W}$

$$100\,\text{W} = 100\,\text{kg}\,\text{m}^2\,\text{s}^{-3} \approx 100\,\text{kg}\,\text{m}^2\,\text{s}^{-3} \cdot ((3 \times 10^8)^{-1}\text{m}^{-1}\,\text{s})^3$$

$$\approx 100\,\text{kg}\,\text{m}^{-1}(3^{-3} \times 10^{-24}) = \frac{100}{27} \times 10^{-24}\text{kg}\,\text{m}^{-1} \approx 3.7 \times 10^{-24}\,\text{kg}\,\text{m}^{-1}$$

2 Convert the following from natural units ($c = 1$) to SI units:

(a) A velocity $v = 10^{-2}$.

$$v = 10^{-2} = 10^{-2}c = 10^{-2}3 \times 10^8\,\text{m}\,\text{s}^{-1} = 3 \times 10^6\,\text{m}\,\text{s}^{-1}$$

(b) Pressure $P = 10^{19}\text{kg}\,\text{m}^{-3}$.

$$P = 10^{19}\text{kg}\,\text{m}^{-3} \approx 10^{19}\text{kg}\,\text{m}^{-3}(3 \times 10^8\,\text{m}\,\text{s}^{-1})^2$$

$$\approx 10^{19}\text{kg}\,\text{m}^{-3}(9 \times 10^{16}\,\text{m}^2\,\text{s}^{-2}) = 9 \times 10^{35}\,\text{N}\,\text{m}^2$$

3 Draw the t and x axes of the spacetime coordinates of an observer $\mathcal{O}$ and then draw:

(a) The world line of $\mathcal{O}$'s clock at $x = 1\,\text{m}$.

4 Write out all the terms of the following sums, substituting the coordinate names (t, x, y, z) for (x^0, x^1, x^2, x^3):

(a) $\sum_{\alpha=0}^{3} V_\alpha\,\text{d}x^\alpha = V_0\,\text{d}t + V_1\,\text{d}x + V_2\,\text{d}y + V_3\,\text{d}z$.

(b) $\sum_{i=1}^{3}(\text{d}x^i)^2 = \text{d}x^2 + \text{d}y^2 + \text{d}z^2 = \text{d}r^2$.

5

(a) Use the spacetime diagram of an observer $\mathcal{O}$ to describe the following experiment performed by $\mathcal{O}$. Two bursts of particles of speed $v = 0.5$ are emitted from $x = 0$ at $t = -2\,\text{m}$, one traveling in the $+x$ direction and the other in the $-x$ direction. These encounter detectors located at $x = \pm 2\,\text{m}$. After a delay of $0.5\,\text{m}$ of time, the detectors send signals back to $x = 0$ at speed $v = 0.75$.

See figure below
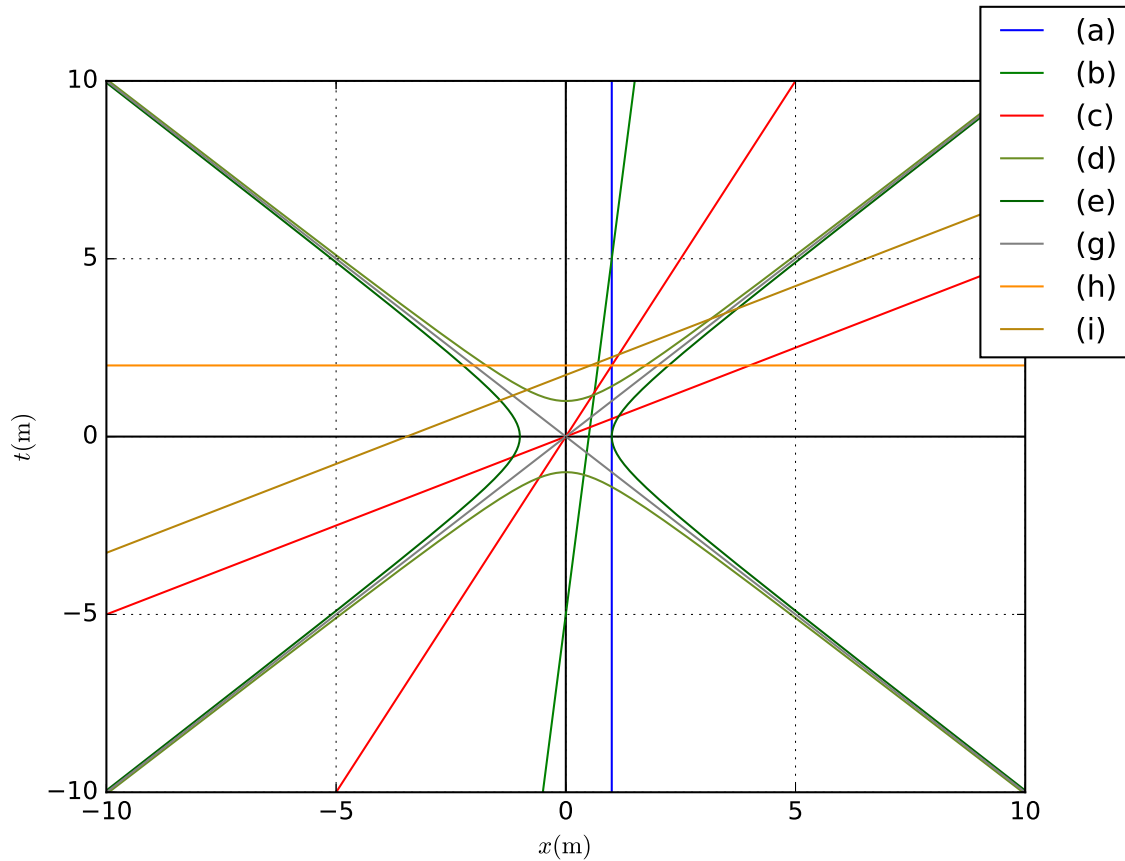
Exercise 3

(b) The signals arrive back at $x = 0$ at the same event. (Make sure your spacetime diagram shows this!) From this the experimenter concludes that the particle detectors did indeed send out their signals simultaneously, since he knows they are equal distances from $x = 0$. Explain why this conclusion is valid.

Assuming he knows the signals traveled with equal speeds, and the detectors are an equal distance away, then they must have been emitted simultaneously, in order for them to arrive at $x = 0$ simultaneously.

(c) A second observer $\bar{\mathcal{O}}$ moves with speed $v = 0.75$ in the $-x$ direction relative to $\mathcal{O}$. Draw the spacetime diagram of $\bar{\mathcal{O}}$ and in it depict the experiment performed by $\mathcal{O}$. Does $\bar{\mathcal{O}}$ conclude that particle detectors sent out their signals simultaneously? If not, which signal was sent first.

See the diagram below. On it, I have drawn lines $\bar{t}_{\text{left}}$ and $\bar{t}_{\text{right}}$ (note that they are parallel to the $\bar{x}$ axis). As you can see from the plot, the left emission occurs *before* the right emission.

(d)

Using $\mathcal{O}$, the distance is

$$\Delta s^2 = \Delta x^2 = 16\,\text{m}^2.$$

Using $\bar{\mathcal{O}}$, we first need to find $\bar{x}_{\{a,b\}}$ and $\bar{t}_{\{a,b\}}$. We use the Lorentz transformation to do this.

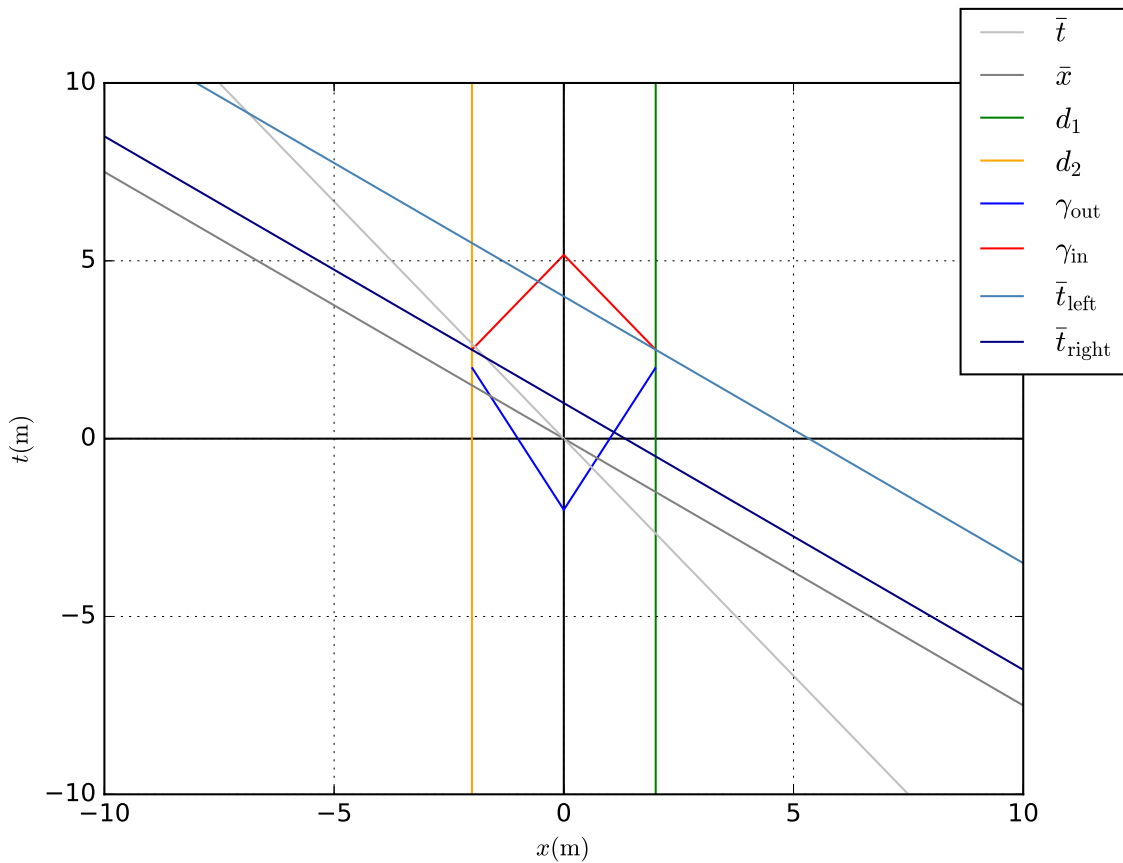$$\bar{t} = \gamma(t - vx)$$

$$\bar{x} = \gamma(x - vt)$$

Using this, we find

$$\bar{t}_a = \frac{16\sqrt{7}}{7} \qquad\qquad \bar{t}_b = \frac{4\sqrt{7}}{7}$$

$$\bar{x}_a = \frac{-31\sqrt{7}}{14} \qquad\qquad \bar{x}_b = \frac{\sqrt{7}}{14}$$

This gives us a distance of

$$\Delta \bar{s}^2 = -(\Delta \bar{t})^2 + (\Delta \bar{x})^2 = 16 \, \mathrm{m}^2,$$

which is of course what we expect.

Exercise 5

6 Show that Equation (Schutz 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 enables us to set $M_{\alpha\beta} = M_{\beta\alpha}$ without loss of generality.

When we expand the summation in (Schutz 1.2), there is no point where

$$d\bar{s}^2 = \ldots + M_{\alpha\alpha}(dx^\alpha)^2 + M_{\alpha\alpha}(dx^\alpha)^2 + \ldots$$

occurs, because a double summation only contains $M_{\alpha\alpha}$ once. If it did, we could absorb the two $M_{\alpha\beta}$ terms into a single one. Therefore we can assert the first point.

Now we consider the second point. If we expand the summation, assuming now that an $M_{\alpha\beta}$ and $M_{\beta\alpha}$ term only occur when $\alpha \neq \beta$, then we see

$$\begin{aligned}
d\bar{s}^2 &= \ldots + M_{\alpha\beta}(dx^\alpha)(dx^\beta) + M_{\beta\alpha}(dx^\beta)(dx^\alpha) + \ldots \\
&= \ldots + (M_{\alpha\beta} + M_{\beta\alpha})\left[(dx^\alpha)(dx^\beta)\right] + \ldots \\
&= \ldots + \mathbf{X}\left[(dx^\alpha)(dx^\beta)\right] + \ldots.
\end{aligned}$$

Now, what really matters in this summation is the value of $\mathbf{X} = M_{\alpha\beta} + M_{\beta\alpha}$, not the individual values of $M_{\alpha\beta}$ and $M_{\beta\alpha}$. Therefore we can *choose*, without loss of generality, $M_{\alpha\beta} = M_{\beta\alpha} = \mathbf{X}/2$, thereby asserting the second point.

7 In the discussion leading up to Equation (Schutz 1.2), assume that the coordinates of $\bar{\mathcal{O}}$ are given as the following linear combinations of those $\mathcal{O}$:

$$\begin{aligned}
\bar{t} &= \alpha t + \beta x, \\
\bar{x} &= \mu t + \nu x, \\
\bar{y} &= ay, \\
\bar{z} &= bz,
\end{aligned}$$

where α, β, μ, ν, a, and b may be functions of the velocity $\vec{v}$ of $\bar{\mathcal{O}}$ relative to $\mathcal{O}$, but they do not depend on the coordinates. Find the values of $M_{\alpha\beta}$ of Equation (Schutz 1.2).

$$\begin{aligned}
d\bar{s}^2 &= -(d\bar{t})^2 + (d\bar{x})^2 + (d\bar{y})^2 + (d\bar{z})^2 \\
&= -(\alpha\,dt + \beta\,dx)^2 + (\mu\,dt + \nu\,dx)^2 + (a\,dy)^2 + (b\,dz)^2 \\
&= -\alpha^2\,dt^2 - \alpha\beta\,dt\,dx - \beta^2\,dx^2 + \mu^2\,dt^2 + \mu\nu\,dt\,dx + \nu^2\,dx^2 + a^2\,dy^2 + b^2\,dz^2 \\
&= (\mu^2 - \alpha^2)\,dt^2 + (\mu\nu - \alpha\beta)\,dt\,dx + (\nu^2 - \beta^2)\,dx^2 + a^2\,dy^2 + b^2\,dz^2
\end{aligned}$$

$$M_{00} = \mu^2 - \alpha^2$$

$$M_{01} = M_{10} = \frac{\mu\nu - \alpha\beta}{2}$$

$$M_{11} = \nu^2 - \beta^2$$

$$M_{22} = a^2$$

$$M_{33} = b^2,$$

and all other $M_{\alpha\beta} = 0$.

8

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

Equation (Schutz 1.3) is just an expansion of the summation in (Schutz 1.2).

We start by taking out the dt^2 term, which corresponds to $\alpha = \beta = 0$, which gives us

$$d\bar{s}^2 = M_{00}(dt)^2 + \dots,$$

now we use the equivalence of dt and dr to make the substitution

$$d\bar{s}^2 = M_{00}(dr)^2 + \dots.$$

For the middle terms, we use the fact that $M_{\alpha\beta} = M_{\beta\alpha}$, and look at only the terms where *one* of α and β is zero. The symmetry means we can write $M_{0i} = M_{i0}$, and pull out a 2 because there are twice as many terms, giving us

$$d\bar{s}^2 = M_{00}(dr)^2$$
$$+ 2\left(\sum_{i=1}^{3} M_{0i}(dx^i)(dt)\right)$$
$$+ \dots.$$

Now we use the equivalence of dt and dr once again, and pull the term out of the sum, giving us

$$d\bar{s}^2 = M_{00}(dr)^2$$
$$+ 2\left(\sum_{i=1}^{3} M_{0i}\, dx^i\right)dr$$
$$+ \dots.$$

Finally, we simply include the terms which have not yet been accounted for, which are all the *spacial-only* terms, which arrives us back at Equation (Schutz 1.3):

$$d\bar{s}^2 = M_{00}(dr)^2$$
$$+ 2\left(\sum_{i=1}^{3} M_{0i}\, dx^i\right)dr$$
$$+ \sum_{i=1}^{3}\sum_{j=1}^{3} M_{ij}\, dx^i\, dx^j.$$

(b) Since $d\bar{s}^2 = 0$ in Equation (Schutz 1.3), for *any* dx^i, replace dx^i with $-dx^i$, and subtract that result

from the original equation. This will establish that $M_{0i} = 0$.

$$\mathrm{d}\bar{s}^2 = M_{00}(\mathrm{d}r)^2$$

$$- 2 \left(\sum_{i=1}^{3} M_{0i} \, \mathrm{d}x^i \right) \mathrm{d}r$$

$$+ \sum_{i=1}^{3} \sum_{j=1}^{3} M_{ij} \, \mathrm{d}x^i \, \mathrm{d}x^j \, .$$

$$\mathrm{d}\bar{s}^2 - \mathrm{d}\bar{s}^2 = 0 = \underline{0 \, M_{00}(\mathrm{d}r)^2}$$

$$+ 4 \left(\sum_{i=1}^{3} M_{0i} \, \mathrm{d}x^i \right) \mathrm{d}r$$

$$+ 0 \sum_{i=1}^{3} \sum_{j=1}^{3} M_{ij} \, \mathrm{d}x^i \, \mathrm{d}x^j.$$

$$0 = \cancel{4} \left(\sum_{i=1}^{3} M_{0i} \, \mathrm{d}x^i \right) \cancel{\mathrm{d}r}$$

Now there are two possibilities. In one case, $\mathrm{d}x^i \equiv 0$, but that is a trivial solution and in general is not true. The other case is that $M_{0i} \equiv 0$, which means we can simplify Equation (Schutz 1.3) to

$$\mathrm{d}\bar{s}^2 = M_{00}(\mathrm{d}r)^2$$

$$+ \sum_{i=1}^{3} \sum_{j=1}^{3} M_{ij} \, \mathrm{d}x^i \, \mathrm{d}x^j \, .$$

(c) Use the result of part (b) with $\mathrm{d}\bar{s}^2 = 0$ to establish Equation (Schutz 1.4b).

$$\mathrm{d}\bar{s}^2 = 0 = M_{00}(\mathrm{d}r)^2 + \sum_{i=1}^{3} \sum_{j=1}^{3} M_{ij} \, \mathrm{d}x^i \, \mathrm{d}x^j$$

$$\implies -M_{00}(\mathrm{d}r)^2 = \sum_{i=1}^{3} \sum_{j=1}^{3} M_{ij} \, \mathrm{d}x^i \, \mathrm{d}x^j \, ,$$

now if we expand $(\mathrm{d}r)^2$, we see that there can only be non-zero M_{ij} when $i = j$, and so

$$-M_{00} \left((\mathrm{d}x^2) + (\mathrm{d}y^2) + (\mathrm{d}z^2) \right) = \sum_{i=1}^{3} M_{ii}(\mathrm{d}x^i)^2$$

$$\implies -(M_{00})\delta_{ij} = M_{ij},$$

which is simply Equation (Schutz 1.4b).

9 Explain why the line $\mathcal{PL}$ in Figure 1.7 is drawn in the manner described in the text.

10 For the pairs of events whose coordinates (t, x, y, z) in some frame are given below, classify their separations as timelike, spacelike, or null.

(a) $(0,0,0,0)$ and $(-1,1,0,0)$:

$$\mathrm{d}s^2 = -(0+1)^2 + (0-1)^2 + (0-0)^2 + (0-0)^2 = -1+1+0+0 = 0 \implies \text{null}$$

(b) $(1,1,-1,0)$ and $(-1,1,0,2)$:

$$\mathrm{d}s^2 = -(1+1)^2 + (1-1)^2 + (-1-0)^2 + (0-2)^2 = -4+0+1+4 = 1 \implies \text{spacelike}$$

(c) $(6,0,1,0)$ and $(5,0,1,0)$:

$$\mathrm{d}s^2 = -(6-5)^2 + (0-0)^2 + (1-1)^2 + (0-0)^2 = -1+0+0+0 = -1 \implies \text{timelike}$$

(d) $(-1,1,-1,1)$ and $(4,1,-1,6)$:

$$\mathrm{d}s^2 = -(-1-4)^2 + (1-1)^2 + (-1+1)^2 + (1-6)^2 = -25+0+0+25 = 0 \implies \text{null}$$

11 Show that the hyperbolae $-t^2 + x^2 = a^2$ and $-t^2 + x^2 = -b^2$ are asymptotic to the lines $t = \pm x$, regardless of a and b.

We will generalize a and $-b$ with a new constant, $\alpha \in \mathbb{R}$, and so we have: $-t^2 + x^2 = \alpha^2$. Now if we solve for t, we get $t = \pm\sqrt{x^2 - \alpha^2}$.

Now take the limit of t as $x \to \infty$ (or $-\infty$, they are equivalent since x is real and squared), which gives us:

$$\lim_{x\to\infty} t = \lim_{x\to\infty} \pm\sqrt{x^2 - \alpha^2} = \pm\sqrt{x^2} = \pm x.$$

Note that we dropped the α^2 term in the limit, as it was being subtracted from a number approaching infinity, and was therefore negligible.

12

(a) Use the fact that the tangent to the hyperbola $\mathcal{DB}$ in Figure 1.14 is the line of simultaneity for $\bar{\mathcal{O}}$ to show that the time interval $\mathcal{AE}$ is shorter than the time recorded on $\bar{\mathcal{O}}$'s clock as it moved from $\mathcal{A}$ to $\mathcal{B}$.

If we look at the figure, we see that $\mathcal{AD}$ and $\mathcal{AB}$ lie along the same hyperbola. This means that when $\mathcal{O}$ measures $\mathrm{d}t = \mathcal{AD}$, and $\bar{\mathcal{O}}$ measures $\mathrm{d}\bar{t} = \mathcal{AB}$, the two measurements are the same. Since $\mathrm{d}t = \mathcal{AE}$ is clearly shorter than $\mathrm{d}t = \mathcal{AD}$, then $\mathrm{d}t = \mathcal{AD} < \mathrm{d}\bar{t} = \mathcal{AB}$.

(b) Calculate that

$$(\mathrm{d}s^2)_{\mathcal{AC}} = (1 - v^2)(\mathrm{d}s^2)_{\mathcal{AB}}$$

$$
\begin{aligned}
(\mathrm{d}s^2)_{\mathcal{AC}} &= -(\mathrm{d}t)^2_{\mathcal{AC}} \\
(\mathrm{d}s^2)_{\mathcal{AB}} &= (\mathrm{d}\bar{s}^2)_{\mathcal{AB}} \\
&= -(\mathrm{d}\bar{t})^2_{\mathcal{AB}} \\
&= -(\gamma(\mathrm{d}t - v\,\mathrm{d}x))^2 = -(\gamma(\mathrm{d}t - v \cdot 0))^2 = -(\gamma\,\mathrm{d}t)^2 = \gamma^2[-(\mathrm{d}t)^2] \\
&= \gamma^2(\mathrm{d}s^2)_{\mathcal{AC}} = \frac{(\mathrm{d}s^2)_{\mathcal{AC}}}{1 - v^2}
\end{aligned}
$$

$$\implies (ds^2)_{\mathcal{AC}} = (1 - v^2)(ds^2)_{\mathcal{AB}}$$

13 The Half-life of the elementary particle called the π-meson (or pion) is 2.5×10^{-8} s when the pion is at rest relative to the observer measuring its decay time. Show, by the principle of relativity, that pions moving at speed $v = 0.999$ must have a half-life of 5.6×10^{-7} s, as measured by an observer at rest.

$$dt = \gamma \, d\bar{t} = \frac{2.5 \times 10^{-8} \, \text{s}}{\sqrt{1 - 0.999^2}} \approx 5.59 \times 10^{-7} \, \text{s}$$

14 Suppose the velocity $\mathbf{v}$ of $\bar{\mathcal{O}}$ relative to $\mathcal{O}$ is small, $|\mathbf{v}| \ll 1$. Show that the time dilation, Lorentz contraction, and velocity-addition formulae can be approximated by respectively:

(a) $dt \approx (1 + \frac{1}{2}v^2) \, d\bar{t}$

$$\gamma = (1 - v^2)^{-1/2} = \sum_{k=0}^{\infty} \binom{-1/2}{k} x^k = 1 + (-1/2)(-v^2) + \frac{(-1/2)(-1/2 - 1)}{2!}(-v^2)^2 + \ldots \approx 1 + \frac{1}{2}v^2$$

$$dt = \gamma \, d\bar{t} \approx \left(1 + \frac{1}{2}v^2\right) d\bar{t}$$

(b) $dx \approx (1 - \frac{1}{2}v^2) \, d\bar{x}$

$$\gamma^{-1} = (1 - v^2)^{1/2} = \sum_{k=0}^{\infty} \binom{1/2}{k} x^k = 1 + (1/2)(-v^2) + \frac{(1/2)(1/2 - 1)}{2!}(-v^2)^2 + \ldots \approx 1 - \frac{1}{2}v^2$$

$$dx = \gamma^{-1} \, d\bar{x} \approx \left(1 - \frac{1}{2}v^2\right) d\bar{x}$$

(c) $W' \approx W + v - Wv(W + v)$ (with $|W| \ll 1$ as well)

$$W' = \frac{W + v}{1 + Wv} = (W + v)(1 + Wv)^{-1}$$

$$(1 + Wv)^{-1} = \sum_{k=0}^{\infty} \binom{-1}{k}(Wv)^k = 1 - Wv + \frac{1}{2} \cdot 1(1 + 1)(Wv)^2 - \frac{1}{6} \cdot 1(1 + 1)(1 + 2)(Wv)^3 + \ldots$$

$$\approx 1 - Wv + (Wv)^2$$

$$W' \approx (W + v)(1 - Wv + (Wv)^2) = W + v - Wv(W + v) + (Wv)^2(W + v)$$

$$\approx W + v - Wv(W + v)$$

What are the relative errors in these approximations when $|\mathbf{v}| = W = 0.1$?

TODO

15 Suppose that the velocity $\mathbf{v}$ of $\bar{\mathcal{O}}$ relative to $\mathcal{O}$ is nearly that of light, $|\mathbf{v}| = 1 - \varepsilon$, $0 < \varepsilon \ll 1$. Show that the same formulae of Exercise 14 become

(a) $dt \approx d\bar{t}/\sqrt{2\varepsilon}$

$$v = 1 - \varepsilon \implies v^2 = (1 - \varepsilon)^2 = 1 - 2\varepsilon + \varepsilon^2$$

$$\implies 1 - v^2 = 1 - (1 - 2\varepsilon + \varepsilon^2) = 2\varepsilon - \varepsilon^2 = 2\varepsilon\left(1 - \frac{\varepsilon}{2}\right)$$

$$\gamma = (1 - v^2)^{-1/2} = \left(2\varepsilon\left(1 - \frac{\varepsilon}{2}\right)\right)^{-1/2} = \frac{1}{\sqrt{2\varepsilon}}\left(1 - \frac{\varepsilon}{2}\right)^{-1/2} \approx \frac{1}{\sqrt{2\varepsilon}}$$

$$\mathrm{d}t = \gamma\,\mathrm{d}\bar{t} \approx \frac{\mathrm{d}\bar{t}}{\sqrt{2\varepsilon}}$$

(b) $\mathrm{d}x \approx \mathrm{d}\bar{x}\,\sqrt{2\varepsilon}$

$$v = 1 - \varepsilon \implies v^2 = (1 - \varepsilon)^2 = 1 - 2\varepsilon + \varepsilon^2$$

$$\implies 1 - v^2 = 1 - (1 - 2\varepsilon + \varepsilon^2) = 2\varepsilon - \varepsilon^2 = 2\varepsilon\left(1 - \frac{\varepsilon}{2}\right)$$

$$\gamma^{-1/2} = (1 - v^2)^{1/2} = \left(2\varepsilon\left(1 - \frac{\varepsilon}{2}\right)\right)^{1/2} = \sqrt{2\varepsilon}\left(1 - \frac{\varepsilon}{2}\right)^{1/2} \approx \sqrt{2\varepsilon}$$

$$\mathrm{d}x = \gamma^{-1}\,\mathrm{d}\bar{x} \approx \mathrm{d}\bar{t}\,\sqrt{2\varepsilon}$$

(c) $W' \approx 1 - \varepsilon(1 - W)/(1 + W)$

TODO

What are the relative errors on these approximations when $\varepsilon = 0.1$ and $W = 0.9$?

TODO

16 Use the Lorentz transformation, Equation 1.12, to derive (a) the time dilation, and (b) the Lorentz contraction formulae. Do this by identifying pairs of events where the separations (in time or space) are to be compared, and then using the Lorentz transformation to accomplish the algebra that the invariant hyperb b olae had been used for in the text.

(a) To derive the time dilation formula, we choose two events that occur at $x = c$, and times t_1 and t_2. Thus, from $\mathcal{O}$'s frame, the time elapsed between these two events is $\Delta t = t_2 - t_1$, and the distance between them is $\Delta x = 0$. Another observer, $\bar{\mathcal{O}}$, moves with some velocity v relative to $\mathcal{O}$. As it passes through the lines $t = t_1$ and $t = t_2$, its clock moves forward by a time $\Delta\tau = \bar{t}_2 - \bar{t}_1$. We now use the Lorentz transformation to write $\Delta\tau$ in terms of $\mathcal{O}$'s coordinates.

$$\Delta\tau = \bar{t}_2 - \bar{t}_1 = \gamma\left[(t_2 - vx_2) - (t_1 - vx_1)\right] = \gamma\left[(t_2 - t_1) + (vx_1 - vx_2)\right]$$

$$= \gamma[\Delta t + v\Delta x] = \gamma[\Delta t + v \cdot 0]$$

$$= \gamma\Delta t$$

and thus we have arrived at the formula for time dilation.

(b) To derive the Lorentz contraction formula, we take a slightly different approach. In the $\mathcal{O}$ frame, a stick

lies parallel to x, such that its length $\ell = x_2 - x_1$. In this frame, the world lines of the two ends of the stick form vertical lines. Another observer, $\bar{\mathcal{O}}$, moves with a velocity v, relative to $\mathcal{O}$. Two events, $\mathcal{A}$ and $\mathcal{B}$ occur on either end of the stick, such that $\bar{\mathcal{O}}$ observes the two events to be simultaneous. Thus, from the $\bar{\mathcal{O}}$ frame, the events are located a distance $\Delta\bar{x} = \bar{\ell}$ apart, and $\Delta\bar{t} = 0$. However, from the $\mathcal{O}$ frame, the events occur a distance $\Delta x = \ell$ apart, and a time separation $\Delta t \neq 0$.

$$\ell = x_2 - x_1 = \gamma\big[(\bar{x}_2 - v\bar{t}_2) - (\bar{x}_1 - v\bar{t}_1)\big] = \gamma\big[(\bar{x}_2 - \bar{x}_1) + v(\bar{t}_1 - \bar{t}_2)\big] = \gamma\bar{\ell}$$

$$\implies \bar{\ell} = \frac{\ell}{\gamma}$$

17 A lightweight pole, 20 m long, lies on the ground next to a barn 15 m long. An Olympic athlete picks up the pole, carries it far away, and runs with it toward the end of the barn at a speed 0.8. His friend remains at rest, standing by the door of the barn. Attempt all parts of this question, even if you can't answer some.

(a) How long does the friend measure the pole to be, as it approaches the barn?

We use the Lorentz contraction equation to find the length the friend measures.

$$\bar{\ell} = \ell/\gamma = \ell\sqrt{1 - v^2} = 20\,\mathrm{m}\sqrt{1 - 0.8^2} = 12\,\mathrm{m}$$

(b) The barn door is initially open and, immediately after the runner and pole are entirely inside the barn, the friend shuts the door. How long after the door is shut does the front of the pole hit the other end of the barn, as measured by the friend? Compute the interval between the events of shutting the door and hitting the wall. Is it spacelike, timelike, or null?

From the runner's point of view, we must consider the length contraction of the barn

(c) In the reference frame of the runner, what is the length of the barn and the pole?

(d) Does the runner believe that the pole is entirely inside the barn when its front hits the end of the barn? Can you explain why?

(e) After the collision, the pole and runner come to rest relative to the barn. From the friend's point of view, the 20 m pole is now inside a 15 m barn, since the barn door was shut before the pole stopped. How is this possible? Alternatively, from the runner's point of view, the collision should have stopped the pole *before* the door closed, so the door could not be closed at all. Was or was not the door closed with the pole inside?

(f) Draw a spacetime diagram from the friend's point of view and use it to illustrate and justify all your conclusions.

18

(a) The Einstein velocity-addition law, Equation 1.13, has a simpler form if we introduce the concept of the *velocity parameter* u, defined by the equation

$$v = \tanh u.$$

Notice that for $-\infty < u < \infty$, the velocity is confined to the acceptable limits $-1 < v < 1$. Show that if

$$v = \tanh u$$

and

$$w = \tanh U,$$

then Equation 1.13 implies

$$w' = \tanh(u + U).$$

This means that velocity parameters add linearly.

There exists an identity:

$$\tanh(x + y) = \frac{\tanh(x) + \tanh(y)}{1 + \tanh(x)\tanh(y)}.$$

If we simply use $x = u$ and $y = U$, then we arrive at

$$\tanh(u + U) = \frac{\tanh(u) + \tanh(U)}{1 + \tanh(u)\tanh(U)} = w'$$

(b) Use this to solve the following problem. A star measures a second star to be moving away at speed $v = 0.9$. The second star measures a third to be receding in the same direction at 0.9. Similarly, the third measures a fourth, and so on, up to some large number N of stars. What is the velocity of the Nth star relative to the first? Give an exact answer and an approximation useful for large N.

Let w^N be the velocity of the Nth star relative to the original star, which we will call star 0. We will use an induction proof to find an expression for w^N. The base case is trivial, $w^0 = 0$, as the star does not move relative to itself. For the next case, $w^1 = v$, we still aren't really doing velocity addition, so we will skip to the w^2 case, where things get interesting, though we will later show that the general expression holds for w^0 and w^1.

For w^2, we simply use the Einstein velocity-addition law:

$$w^2 = \tanh(u + U) = \tanh\left(\tanh^{-1} v + \tanh^{-1} w^1\right) = \tanh\left(2\tanh^{-1} v\right).$$

Now I will prove that this is one instance of a general expression, that $w^N = \tanh\left(N\tanh^{-1} v\right)$.

$$w^N = \tanh\left(N\tanh^{-1} v\right)$$
$$\implies \tanh^{-1} w^N = N\tanh^{-1} v$$
$$\implies \tanh^{-1} w^N + \tanh^{-1} v = N\tanh^{-1} v + \tanh^{-1} v$$
$$\implies \tanh^{-1} w^N + \tanh^{-1} v = (N + 1)\tanh^{-1} v$$
$$\implies \tanh\left(\tanh^{-1} w^N + \tanh^{-1} v\right) = \tanh\left((N + 1)\tanh^{-1} v\right)$$
$$\implies w^{N+1} = \tanh\left((N + 1)\tanh^{-1} v\right).$$

If you can believe the last step, then this is proof that it works for all N. The last step is saying that, if we have a star N, moving away from star 0 at a speed w^N, and another star $N + 1$, moving away from star N at a speed v, then star $N + 1$ as observed from star 0 is given by the Einstein velocity-addition law, meaning we can rewrite that expression as w^{N+1}.

Now I'd like to go back and show that this works for $N = 0$ and $N = 1$. For $N = 1$, we get

$$w^1 = \tanh\left(1 \tanh^{-1} v\right) = v,$$

which is what we would expect, and for $N = 0$, we get

$$w^0 = \tanh\left(0 \tanh^{-1} v\right) = 0,$$

which we also expect. So the general expression,

$$w^N = \tanh\left(N \tanh^{-1} v\right),$$

holds true for all non-negative integers N. We can also write this more elegantly as

$$w^N = \tanh(Nu).$$

Now we want to consider the behaviour at large N. We first write tanh in its exponential form, as

$$w^N = \frac{1 - \exp(-2Nu)}{1 + \exp(-2Nu)}.$$

When N is very large, then the exponential in the bottom term goes to zero, allowing us to rewrite it as

$$w^N \approx 1 - \exp(-2Nu).$$

We can go a step further. Since $v = 0.9$, $u \approx 1.47$, which we can neglect for large N, and so we finally arrive at

$$w^N \approx 1 - \exp(-2N).$$

19

(a) Using the velocity parameter (u) introduced in Exercise 18, show that the Lorentz transformation equations, Equation 1.12, can be put in the form

$$\bar{t} = t \cosh u - x \sinh u \qquad\qquad \bar{y} = y$$

$$\bar{x} = -t \sinh u + x \cosh u \qquad\qquad \bar{z} = z$$

We start by putting γ in terms of u.

$$\gamma = (1 - v^2)^{-1/2} = (1 - \tanh^2 u)^{-1/2} = \frac{1}{\text{sech } u} = \cosh u.$$

Now we can substitute this into the Lorentz transformation equations

$$\bar{t} = \gamma(t - vx) = \cosh u(t - x\tanh u) = t\cosh u - x\sinh u$$

$$\bar{x} = \gamma(x - vt) = \cosh u(x - t\tanh u) = x\cosh u - t\sinh u$$

(b) Use the identity $\cosh^2 u - \sinh^2 u = 1$ to demonstrate the invariance of the interval from these equations.

$$ds^2 = -dt^2 + dx^2 + dy^2 + dz^2$$

$$d\bar{s}^2 = -(dt\cosh u - dx\sinh u)^2 + (dx\cosh u - dt\sinh u)^2 + dy^2 + dz^2$$

$$= -\left(dt^2\cosh^2 u - \cancel{dx\,dt\sinh u\cosh u} + dx^2\sinh^2 u\right)$$

$$+ \left(dx^2\cosh^2 u - \cancel{dt\,dx\sinh u\cosh u} + dt^2\sinh^2 u\right) + dy^2 + dz^2$$

$$= -dt^2\left(\cancel{\cosh^2 u - \sinh^2 u}\right) + dx^2\left(\cancel{\cosh^2 u - \sinh^2 u}\right) + dy^2 + dz^2$$

$$= ds^2$$

(c) Draw as many parallels as you can between the geometry of spacetime and ordinary two-dimensional Euclidean geometry, where the coordinate transformation analogous to the Lorentz transformation is

$$\bar{x} = +x\cos\theta + y\sin\theta,$$

$$\bar{y} = -x\sin\theta + y\cos\theta.$$

What is the analog of the interval? Of the invariant hyperbolae?
The analog of the interval would be

$$d\bar{r}^2 = d\bar{x}^2 + d\bar{y}^2 = (dx\cos\theta + dy\sin\theta)^2 + (dy\cos\theta - dx\sin\theta)^2 +$$

$$= dx^2\cos^2\theta + \cancel{2\,dx\,dy\sin\theta\cos\theta} + dy^2\sin^2\theta$$

$$+ dy^2\cos^2\theta - \cancel{2\,dx\,dy\sin\theta\cos\theta} + dx^2\sin^2\theta$$

$$= dx^2(\sin^2\theta + \cos^2\theta) + dy^2(\sin^2\theta + \cos^2\theta)$$

$$= dx^2 + dy^2$$

The analog of the invariant hyperbola would be the invariant circle, as $\bar{x}$ and $\bar{y}$ are both equations of a circle.

20 Write the Lorentz transformation equations in matrix form.

$$\bar{t} = \gamma(t - vx) \qquad\qquad\qquad \bar{t} = \gamma t - \gamma vx + 0y + 0z$$

$$\bar{x} = \gamma(x - vt) \qquad\qquad\qquad \bar{x} = -\gamma vt + \gamma x + 0y + 0z$$

$$\bar{y} = y \qquad\qquad\qquad\qquad\quad \bar{y} = y$$

$$\bar{z} = z \qquad\qquad\qquad\qquad\quad \bar{z} = z$$

$$\begin{pmatrix} \bar{t} \\ \bar{x} \\ \bar{y} \\ \bar{z} \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma v & 0 & 0 \\ -\gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix}$$

21

(a) Show that if the two events are timelike separated, there is a Lorentz frame in which they occur at the same point, i.e. at the same spatial coordinate values.

If the two events are timelike separated, then it must be possible to have an object with a worldline which crosses the two points, as it is inside the light cone. If such an object exists, then we can draw a Lorentz frame for it, so its time axis, $\bar{t}$ is that line, meaning $\bar{x} = 0$ for both events.

(b) Similarly, if the two events are spacelike separated, there is a Lorentz frame in which they occur simultaneously.

If the two events are spacelike separated, then it must be possible to draw a coordinate frame where $\bar{x}$ has slope v in $\mathcal{O}$'s frame. This means that $\bar{t} = 0$ for both events, and so they are simultaneous.