

1

Special relativity

The essence of a physical theory expressed in mathematical form is the identification of mathematical concepts with certain physically measurable quantities. This must be our first concern . . .

Bernard Schutz, §7.1

Minkowski pointed out that it is very helpful to regard (t, x, y, z) as simply four coordinates in a four-dimensional space [that] we now call spacetime. This was the beginning of the geometrical point of view, which led directly to general relativity in 1914–16.

Bernard Schutz, §1.1

1.1 Exercises

- 1.1 Convert the following to [natural] units in which $c = 1$, expressing everything in terms of m and kg:
- (a) Worked example: 10 J.

Solution:

$$10 \text{ J} = 10 \text{ N m} = 10 \text{ kg m}^2 \text{ s}^{-2} = \frac{10 \text{ kg m}^2 \text{ s}^{-2}}{(3 \times 10^8 \text{ m s}^{-1})^2} = 1.11 \times 10^{-16} \text{ kg}.$$

- (c) Planck’s reduced constant, $\hbar = 1.05 \times 10^{-34} \text{ J s}$. (Note the definition of $\hbar$ in terms of Planck’s constant h : $\hbar \equiv h/2\pi$.)

Solution:

$$\hbar = 1.05 \times 10^{-34} \text{ J s} = \frac{1.05 \times 10^{-34} \text{ kg m}^2 \text{ s}^{-1}}{3 \times 10^8 \text{ m s}^{-1}} = 3.52 \times 10^{-43} \text{ kg m}.$$

- (e) Momentum of a car.

Solution:

$$p = \frac{30 \text{ ms}^{-1}}{3 \times 10^8 \text{ ms}^{-1}} \times 1000 \text{ kg} = 10^{-4} \text{ kg}.$$

(g) Water density, 10^3 kg m^{-3} .

Solution:

$$10^3 \text{ kg m}^{-3}.$$

We will learn in Chapter 8 how to express mass in terms of meters, see in particular eqn. (8.8).

1.2 Convert from natural units ($c = 1$) to SI units

(a) Velocity, $v = 10^{-2}$:

Solution:

$$v = 10^{-2} \times c[\text{m s}^{-1}] = 3 \times 10^6 [\text{m s}^{-1}].$$

(c) Time, $10^{18} [\text{m}]$:

Solution:

$$\frac{10^{18} [\text{m}]}{c[\text{m s}^{-1}]} = 3.3 \times 10^9 [\text{s}].$$

(e) Acceleration, $10 [\text{m}^{-1}]$:

Solution:

$$10 [\text{m}^{-1}] \times c^2[\text{m}^2 \text{ s}^{-2}] = 9 \times 10^{17} [\text{m s}^{-2}].$$

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

(c) The $\bar{t}$ and $\bar{x}$ axes of an observer $\bar{\mathcal{O}}$ who moves with velocity $v = 0.5$ in the positive x -direction relative to $\mathcal{O}$ and whose origin ($\bar{t} = \bar{x} = 0$) coincides with that of $\mathcal{O}$.

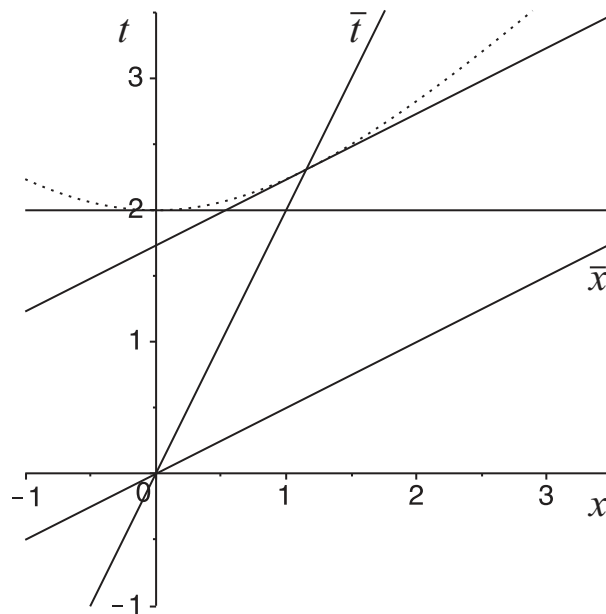


Figure 1.1

The $\bar{x}$ and $\bar{t}$ axes are the solution to Exercise 1.3(c). The dotted line is the invariant hyperbola with $\Delta s^2 = -4$. The solution to 1.3(h) is the horizontal line. The solution to 1.3(i) is the sloping line, parallel to the $\bar{x}$ -axis. It is tangent to the invariant hyperbola at the $\bar{t}$ -axis. These plots were made using the Maple™ worksheet that accompanies this book.

Solution: Recall from Schutz §1.5 that the $\bar{t}$ -axis follows from simple kinematics; it is just the line $t = x/v$, so here $t = 2x$. Recall also from §1.5 (see Schutz Fig. 1.5) that the $\bar{x}$ -axis was a straight line with slope equal to the inverse of that of the $\bar{t}$ -axis, $x = t/v$. (In SP1.3 you will prove this.) Here $t = x/2$. The solution was plotted in fig. 1.1.

- (h) The locus of events, all of which occur at the time $t = 2$ m (simultaneous as seen by $\mathcal{O}$).

Solution: See fig. 1.1.

- (i) The locus of events, all of which occur at the time $\bar{t} = 2$ m (simultaneous as seen by $\bar{\mathcal{O}}$).

Solution: The locus of events, all of which occur at the time $\bar{t} = 2$ m, have arbitrary $\bar{x}$, and so the solution is a straight line parallel to the $\bar{x}$ -axis. The coordinates in the

$\mathcal{O}$ frame are easily found with the Lorentz transformation. (See SP1.13 for a different approach.) From Schutz Eq. (1.12) we have

$$\bar{t} = 2 = \frac{t - vx}{\sqrt{1 - v^2}} \Rightarrow t = vx + 2\sqrt{1 - v^2} = x/2 + \sqrt{3}.$$

The solution was plotted in fig. 1.1.

- 1.5 (c) A second observer $\bar{\mathcal{O}}$ moves with speed $v = 0.75$ in the negative 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 the particle detectors sent out their signals simultaneously? If not, which signal was sent first?

Hint: See Schutz Fig. 1.5(b) for how the time and space axes look for a reference frame moving in the negative x -direction. Think carefully about what the $\bar{t}$ and $\bar{x}$ mean.

- (d) Compute the interval Δs^2 between the events at which the detectors emitted their signals, using both the coordinates of $\mathcal{O}$ and those of $\bar{\mathcal{O}}$.

Hint: Use the Lorentz transformation for a velocity boost to obtain the coordinates of the events in $\bar{\mathcal{O}}$.

- 1.6 Show that the interval

$$\Delta \bar{s}^2 = \sum_{\alpha=0}^3 \sum_{\beta=0}^3 M_{\alpha\beta} (\Delta x^\alpha) (\Delta x^\beta), \quad \text{Schutz Eq. (1.2)} \quad (1.1)$$

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.

Solution: Pick a pair of indices, $\alpha = \alpha^*$ and $\beta = \beta^*$ say, with $\alpha^* \neq \beta^*$, and where α^* and β^* are fixed integers in the set $\{0, 1, 2, 3\}$. 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. The RHS of eqn. (1.1) is called a *quadratic form*, and thus the interval of SR can be written as a symmetric quadratic form.

1.8 (a) Derive,

$$\Delta \bar{s}^2 = M_{00}(\Delta r)^2 + 2M_{0i}\Delta x^i \Delta r + M_{ij}\Delta x^i \Delta x^j, \quad \text{Schutz Eq. (1.3)} \quad (1.2)$$

where $\Delta r = \sqrt{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2}$, from eqn. (1.1) for general $M_{\alpha\beta}$. [You can assume $\Delta s^2 = 0$ and $\Delta t > 0$.]

Solution: Start with eqn. (1.1), and partially expand the summations

$$\begin{aligned} \Delta \bar{s}^2 &= M_{00}(\Delta t)^2 + \sum_{i=1}^3 M_{0i} \Delta t \Delta x^i + \sum_{i=1}^3 M_{i0} \Delta x^i \Delta t + \sum_{i=1}^3 \sum_{j=1}^3 M_{ij} \Delta x^i \Delta x^j \\ &= M_{00}(\Delta t)^2 + 2 \sum_{i=1}^3 M_{0i} \Delta t \Delta x^i + \sum_{i=1}^3 \sum_{j=1}^3 M_{ij} \Delta x^i \Delta x^j. \quad \text{used } M_{i0} = M_{0i} \end{aligned}$$

Consider the case $\Delta s^2 = 0$, so from Schutz Eq. (1.1), $\Delta t = \pm \Delta r = \pm \sqrt{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2}$. Then, when $\Delta t > 0$,

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

which is eqn. (1.2).

(b) Since $\Delta \bar{s}^2 = 0$ in eqn. (1.2) for any $\{\Delta x^i\}$, replace Δx^i by $-\Delta x^i$ in eqn. (1.2) and subtract the resulting equations from eqn. (1.2) to establish that $M_{0i} = 0$ for $i = 1, 2, 3$.

Solution: Let us first recall why $\Delta\bar{s}^2 = 0$ in eqn. (1.2) for any $\{\Delta x^i\}$. We have set $\Delta s^2 = 0$ (because we were considering the path of a light ray) and it followed, based upon the universality of the speed of light, that we required also $\Delta\bar{s}^2 = 0$. Now why does $\Delta s^2 = 0$ for any Δx^i ? Because we have imposed that we are considering the path of a light ray, and regardless of the spatial point x^i on the light ray path we choose, it always has $(\Delta t)^2 = (\Delta r)^2$, so $\Delta s^2 = -(\Delta t)^2 + (\Delta r)^2 = 0$.

Now note that changing Δx^i to $-\Delta x^i$ does not change

$$\Delta r = \sqrt{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2}.$$

Thus the only term in eqn. (1.2) to change sign when changing Δx^i to $-\Delta x^i$ is the middle term, the sum over $2M_{0i}\Delta x^i\Delta r$. The final term does not because changing Δx^i to $-\Delta x^i$ also changes Δx^j to $-\Delta x^j$; the i and j are just dummy indices. So when we subtract $\Delta\bar{s}^2(\Delta t, \Delta x^i) - \Delta\bar{s}^2(\Delta t, -\Delta x^i)$ as instructed, using eqn. (1.2), we find:

$$\begin{aligned} 0 &= 0 - 0 = \Delta\bar{s}^2(\Delta t, \Delta x^i) - \Delta\bar{s}^2(\Delta t, -\Delta x^i) \\ &= M_{00}(\Delta r)^2 + 2\Delta r \sum_{i=1}^3 M_{0i}\Delta x^i + \sum_{i=1}^3 \sum_{j=1}^3 M_{ij}\Delta x^i\Delta x^j \\ &\quad - \left(M_{00}(\Delta r)^2 + 2\Delta r \sum_{i=1}^3 M_{0i}(-\Delta x^i) + \sum_{i=1}^3 \sum_{j=1}^3 M_{ij}(-\Delta x^i)(-\Delta x^j) \right) \\ &= 4\Delta r \sum_{i=1}^3 M_{0i}\Delta x^i. \end{aligned} \quad (1.3)$$

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

(c) Derive

$$M_{ij} = -M_{00}\delta_{ij}, \quad (i, j = 1, 2, 3) \quad \text{Schutz Eq. (1.4b)} \quad (1.4)$$

using eqn. (1.2) with $\Delta\bar{s}^2 = 0$. Hint: Δx , Δy , and Δz are arbitrary.

Solution: Recall from Exercise 1.8(b) that adding to eqn. (1.2) the following

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

gives

$$0 = M_{00}(\Delta r)^2 + \sum_{i=1}^3 \sum_{j=1}^3 M_{ij}\Delta x^i\Delta x^j. \quad (1.5)$$

Suppose, $\Delta x = \Delta r$, $\Delta y = \Delta z = 0$. Substituting into eqn. (1.5) 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 is also possible that $\Delta x = \Delta y = \Delta r/\sqrt{2}$ and $\Delta z = 0$. Substitution into eqn. (1.5) gives that

$$\begin{aligned} 0 &= (M_{12} + M_{21})(\Delta r)^2/2 + M_{11}(\Delta r)^2/2 + M_{22}(\Delta r)^2/2 + (\Delta r)^2 M_{00} \\ &= (M_{12} + M_{21})(\Delta r)^2/2 - M_{00}(\Delta r)^2/2 - M_{00}(\Delta r)^2/2 + (\Delta r)^2 M_{00} \\ &= (M_{12} + M_{21})(\Delta r)^2/2 = M_{21}(\Delta r)^2. \end{aligned} \tag{1.6}$$

The final step used $M_{\alpha\beta} = M_{\beta\alpha}$, as proved in Exercise 1.6. And since $(\Delta r)^2$ was arbitrary, we have $M_{21} = 0 = M_{12}$. Similarly, $M_{13} = M_{31} = 0 = M_{23} = M_{32}$. In summary,

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

which is eqn. (1.4).

- 1.9 Explain why the line $\mathcal{PQ}$ in Schutz Fig. 1.7 is drawn in the manner described in the text. [Note that in Schutz Fig. 1.7 the $\mathcal{F}$ should be a $\mathcal{Q}$ to be consistent with the text and with the corresponding figure in the first edition (Schutz, 1985, Fig. 1.7).]

Solution: The line $\mathcal{PQ}$ is described in the paragraph after Schutz Eq. (1.5) as perpendicular to the y -axis, parallel to the t - x plane, and parallel to the $\bar{t}$ -axis in Schutz Fig. 1.5(a). The line $\mathcal{PQ}$ represents the path of a clock that is stationary in the $\bar{\mathcal{O}}$ frame. Because the $\bar{\mathcal{O}}$ frame moves in the x -direction its path must be orthogonal to the y -axis. And furthermore it must be parallel to the t - x plane, as argued for a clock at the origin of the $\bar{\mathcal{O}}$ frame in Schutz §1.5. In fact the clock is simply displaced a fixed distance from $y = 0$ along the y - or $\bar{y}$ -axis and moves parallel to the $\bar{t}$ -axis.

- 1.11 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 .

Hint: Regardless of how large a and b are, consider the approximate behavior when $|x|$ and $|t|$ are much greater than $|a|$ and $|b|$.

- 1.12 (a) Use the fact that the tangent to the hyperbola $\mathcal{DB}$ in Schutz Fig. 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 $\mathcal{B}$.

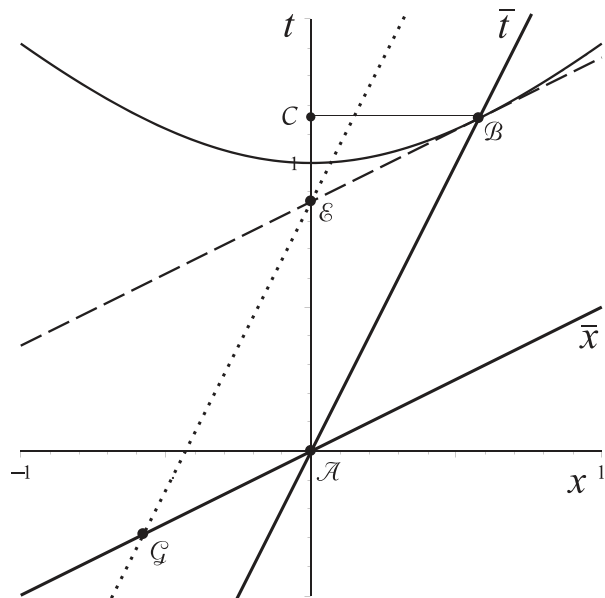


Figure 1.2 Similar to Schutz Fig. 1.14. The dotted line is the path of a second clock at rest in $\overline{\mathcal{O}}$ needed to infer that the moving clock along the t -axis runs slowly.

Solution: This example shows that time dilation is self-consistent. From the perspective of an observer in $\overline{\mathcal{O}}$, the time interval $\mathcal{AE} = \Delta\tau$ corresponds to the proper time of a moving clock, whose world line in Schutz Fig. 1.14 is the t -axis, see fig. 1.2. An observer at rest in $\overline{\mathcal{O}}$ needs two clocks to record the time interval $\Delta\bar{t} = \bar{t}_E - \bar{t}_A$ corresponding to the proper time interval $\Delta\tau$. The clock moving from $\mathcal{A}$ to $\mathcal{B}$ is one of those two clocks, recording $\bar{t}_A$. The other is drawn as a dotted line (fig. 1.2) that passes through $\mathcal{E}$, recording $\bar{t}_E$. The fact that the line of simultaneity in $\overline{\mathcal{O}}$ passes through $\mathcal{B}$ and $\mathcal{E}$ means that $\bar{t}_E = \bar{t}_B$, and hence $\Delta\bar{t} = \bar{t}_B - \bar{t}_A$. Recall the time dilation formula,

$$\Delta\tau = \Delta t \sqrt{1 - v^2}, \quad \text{Schutz Eq. (1.10)} \quad (1.7)$$

where Δt was the so-called improper time, an interval measured by two clocks. Here $\Delta\bar{t}$ plays the role of Δt (improper time measured by two clocks):

$$\Delta\tau = \Delta\bar{t} \sqrt{1 - v^2}, \quad (1.8)$$

implying $\Delta\tau < \Delta\bar{t}$ for $|v| > 0$.

Don't be thrown off by the Δt in eqn. (1.7) not having a bar above it, while it does in eqn. (1.8) above. *It is not the symbol that is important but the role played by the thing it depicts.* The roles of the $\mathcal{O}$ and $\overline{\mathcal{O}}$ frame have been reversed in this exercise, which was the point of discussion around Schutz Fig. 1.14.

1.12 (b) Calculate that¹

$$(\Delta s^2)_{\mathcal{AE}} = (1 - v^2) (\Delta s^2)_{\mathcal{AB}}. \quad (1.9)$$

Solution: Start with the LHS of eqn. (1.9):

$$\begin{aligned} (\Delta s^2)_{\mathcal{AE}} &\equiv -(t_{\mathcal{E}} - t_{\mathcal{A}})^2 + (x_{\mathcal{E}} - x_{\mathcal{A}})^2 && \text{definition of the interval} \\ &= -t_{\mathcal{E}}^2 + x_{\mathcal{E}}^2 && \mathcal{A} \text{ is the origin} \\ &= -t_{\mathcal{E}}^2. && \mathcal{E} \text{ on } t\text{-axis} \end{aligned} \quad (1.10)$$

From fig. 1.2 herein it is clear that

$$\begin{aligned} t_{\mathcal{E}} &= t_{\mathcal{B}} - x_{\mathcal{B}}v && \text{dashed line parallel to } \bar{x}\text{-axis, slope is } v \\ &= t_{\mathcal{B}} - (t_{\mathcal{B}}v)v \\ &= t_{\mathcal{B}}(1 - v^2). \end{aligned} \quad (1.11)$$

Now consider the RHS of eqn. (1.9),

$$(\Delta s^2)_{\mathcal{AB}} = -t_{\mathcal{B}}^2 + x_{\mathcal{B}}^2 = -t_{\mathcal{B}}^2 + (vt_{\mathcal{B}})^2 = -t_{\mathcal{B}}^2(1 - v^2). \quad (1.12)$$

Combining eqns. (1.10, 1.11, 1.12) one finds,

$$(\Delta s^2)_{\mathcal{AE}} = -t_{\mathcal{B}}^2(1 - v^2)^2 = (1 - v^2) (\Delta s^2)_{\mathcal{AB}}. \quad (1.13)$$

1.12 (c) Use (b) to show that $\overline{\mathcal{O}}$ regards $\mathcal{O}$'s clocks to be running slowly, at just the right rate.

Solution: This corresponds to verifying eqn. (1.8) above; recall $\Delta\tau = t_{\mathcal{E}}$ and $\Delta\bar{t} = \bar{t}_{\mathcal{B}}$. To find $\bar{t}_{\mathcal{B}}$ use the fact that the interval is invariant between Lorentz frames,

$$\begin{aligned} (\Delta s^2)_{\mathcal{AB}} &= -t_{\mathcal{B}}^2 + x_{\mathcal{B}}^2 = -\bar{t}_{\mathcal{B}}^2 + \bar{x}_{\mathcal{B}}^2 \\ &= -\bar{t}_{\mathcal{B}}^2. && \mathcal{B} \text{ on } \bar{t}\text{-axis} \end{aligned} \quad (1.14)$$

Combining eqns. (1.10, 1.13, 1.14)

$$\begin{aligned} -t_{\mathcal{E}}^2 &= (\Delta s^2)_{\mathcal{AE}} = (1 - v^2) (\Delta s^2)_{\mathcal{AB}} = -(1 - v^2) \bar{t}_{\mathcal{B}}^2 \\ t_{\mathcal{E}} &= \bar{t}_{\mathcal{B}} \sqrt{1 - v^2}. && \text{took square root} \end{aligned} \quad (1.15)$$

1.13 The half-life of the elementary particle called the pi 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

¹ We had corrected a typo in the original question, replacing $\mathcal{AC}$ with $\mathcal{AE}$. SP1.15 explores the other possible interpretation.

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.

Hint: Study the solution to Exercise 1.12, and make the analogy with the situation here. Think of the pion as a clock of sorts; its birth is say at time zero and its decay is another tick of the clock. In making the analogy with Exercise 1.12, pay attention to which time intervals are measured by one clock (proper time intervals) and which involve two physically separated clocks.

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

(a) $\Delta t \approx \left(1 + \frac{1}{2}v^2\right) \Delta \bar{t},$ (1.16)

(b) $\Delta x \approx \left(1 - \frac{1}{2}v^2\right) \Delta \bar{x},$ (1.17)

(c) $w' \approx w + v - wv(w + v),$ (with $|w| \ll 1$ as well). (1.18)

What are the relative errors in these approximations when $v = w = 0.1$.

(a) Solution: Recall the time dilation formula was given in eqn. (1.7), with here $\Delta \tau = \Delta \bar{t}$. Solving for Δt , and expanding the RHS in a Taylor series in the small parameter v we obtain

$$\begin{aligned} \Delta t &= \Delta \bar{t} \frac{1}{\sqrt{1 - v^2}} = \Delta \bar{t} (1 - v^2)^{-1/2} \\ &= \Delta \bar{t} \left(1 + \frac{1}{2}v^2 + \frac{3}{8}v^4 + \dots\right) \quad \text{used eqn. (B.2)} \\ &\simeq \Delta \bar{t} \left(1 + \frac{1}{2}v^2\right). \end{aligned} \tag{1.19}$$

For the Taylor series we have used the binomial series, eqn. (B.2) of Appendix B, a result well worth remembering! The largest term we ignored was $\frac{3}{8}v^4$. You will often see this written as $O(v^4)$, read “of order v to the fourth.” This means that we are focusing attention on the important part, i.e. v^4 , and ignoring the irrelevant numerical factor $3/8$ that is close to unity. The higher order terms in the series were $O(v^6)$ and these are clearly much smaller since $v \ll 1$. The relative error is then

$$\frac{\frac{3}{8}v^4}{(1 - v^2)^{-1/2}} \approx \frac{3}{8}v^4 = 3.75 \times 10^{-5}.$$

In fact the relative error can be calculated exactly to be 3.76×10^{-5} , see accompanying Maple™ worksheet.