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 $3.00 \times 10^8\,\mathrm{ms}^{-1}$ 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

$$\begin{aligned}\mathrm{d}\bar{s}^2 &= M_{00}(\mathrm{d}r)^2 \\ &\quad + 2\left(\sum_{i=1}^{3} M_{0i}\,\mathrm{d}x^i\right)\mathrm{d}r \\ &\quad + \sum_{i=1}^{3}\sum_{j=1}^{3} M_{ij}\,\mathrm{d}x^i\,\mathrm{d}x^j\,,\end{aligned} \tag{Schutz 1.3}$$

where

$$M_{0i} = 0 \tag{Schutz 1.4a}$$

and

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

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

1.14 Exercises

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

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

(a) 10 J

$$
\begin{aligned}
10\,\mathrm{J} = 10\,\mathrm{N\,m} = 10\,\mathrm{kg\,m^2\,s^{-2}} &\approx 10\,\mathrm{kg\,m^2\,s^{-2}} \cdot ((3\times10^8)^{-1}\mathrm{m^{-1}\,s})^2 \\
&\approx 10\,\mathrm{kg}(3\times10^8)^{-2} = 10\,\mathrm{kg}\left(\frac{1}{9}\times10^{-16}\right) \approx 1.11\times10^{-16}\,\mathrm{kg}
\end{aligned}
$$

(b) 100 W

$$
\begin{aligned}
100\,\mathrm{W} = 100\,\mathrm{kg\,m^2\,s^{-3}} &\approx 100\,\mathrm{kg\,m^2\,s^{-3}} \cdot ((3\times10^8)^{-1}\mathrm{m^{-1}\,s})^3 \\
&\approx 100\,\mathrm{kg\,m^{-1}}(3^{-3}\times10^{-24}) = \frac{100}{27}\times10^{-24}\mathrm{kg\,m^{-1}} \approx 3.7\times10^{-24}\,\mathrm{kg\,m^{-1}}
\end{aligned}
$$

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

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

$$
v = 10^{-2} = 10^{-2}c = 10^{-2}3\times10^8\,\mathrm{m\,s^{-1}} = 3\times10^6\,\mathrm{m\,s^{-1}}
$$

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

$$
\begin{aligned}
P = 10^{19}\mathrm{kg\,m^{-3}} &\approx 10^{19}\mathrm{kg\,m^{-3}}(3\times10^8\,\mathrm{m\,s^{-1}})^2 \\
&\approx 10^{19}\mathrm{kg\,m^{-3}}(9\times10^{16}\,\mathrm{m^2\,s^{-2}}) = 9\times10^{35}\,\mathrm{N\,m^2}
\end{aligned}
$$

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 $x = 1\,\mathrm{m}$.

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

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

5

TODO

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

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

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

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

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 $\mathcal{O}$:

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

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).

The simplest $M_{\alpha\beta}$ to find are those relating to $\bar{y}$ (M_{22}) and $\bar{z}$ (M_{33}). For $\bar{y}$ we have:

$$\mathrm{d}\bar{y}^2 = (a\,\mathrm{d}y)^2 = M_{22}(\mathrm{d}x^2)^2 \implies M_{22} = a^2,$$

and similarly for z we have $M_{33} = b^2$.

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 $\mathrm{d}t^2$ term, which corresponds to $\alpha = \beta = 0$, which gives us

$$\mathrm{d}\bar{s}^2 = M_{00}(\mathrm{d}t)^2 + \ldots,$$

now we use the equivalence of $\mathrm{d}t$ and $\mathrm{d}r$ to make the substitution

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

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

$$\begin{aligned} \mathrm{d}\bar{s}^2 &= M_{00}(\mathrm{d}r)^2 \\ &+ 2\left(\sum_{i=1}^{3} M_{0i}(\mathrm{d}x^i)(\mathrm{d}t)\right) \\ &+ \ldots. \end{aligned}$$

Now we use the equivalence of $\mathrm{d}t$ and $\mathrm{d}r$ once again, and pull the term out of the sum, giving us

$$\begin{aligned} \mathrm{d}\bar{s}^2 &= M_{00}(\mathrm{d}r)^2 \\ &+ 2\left(\sum_{i=1}^{3} M_{0i}\,\mathrm{d}x^i\right)\mathrm{d}r \\ &+ \ldots. \end{aligned}$$

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):

$$\begin{aligned} \mathrm{d}\bar{s}^2 &= M_{00}(\mathrm{d}r)^2 \\ &+ 2\left(\sum_{i=1}^{3} M_{0i}\,\mathrm{d}x^i\right)\mathrm{d}r \\ &+ \sum_{i=1}^{3}\sum_{j=1}^{3} M_{ij}\,\mathrm{d}x^i\,\mathrm{d}x^j\,. \end{aligned}$$

(b) Since $\mathrm{d}\bar{s}^2 = 0$ in Equation (Schutz 1.3), for *any* $\mathrm{d}x^i$, replace $\mathrm{d}x^i$ with $-\mathrm{d}x^i$, and subtract that result from the original equation. This will establish that $M_{0i} = 0$.

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

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

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

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

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

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

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

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

$$\begin{aligned}-M_{00}\Big((\mathrm{d}x^2) + (\mathrm{d}y^2) + (\mathrm{d}z^2)\Big) &= \sum_{i=1}^{3} M_{ii}(\mathrm{d}x^i)^2 \\ \implies -(M_{00})\delta_{ij} &= M_{ij},\end{aligned}$$

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.

The key here is that both $\mathcal{P}$ and $\mathcal{L}$ must emit or absorb a beam of light, which intersect the same reflection points at $\mathcal{A}$ and $\mathcal{B}$. This means that they occur at $t = \bar{t} = 0$. This also means that from $\bar{\mathcal{O}}$, it appears that the light reflected from

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 $t = \pm\sqrt{x^2 - c^2}$.

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.