Relativity made relatively easy

Andrew Steane

October 6, 2011

Contents

[1] This is loosely based on the discussion by Fayngold.

Chapter 1

Introduction

This book presents an extensive study of Special Relativity, aimed at an undergraduate level. It is not intended to be the first introduction to the subject for most students, although for a bright student it could function as that. Therefore basic ideas such as time dilation and space contraction are recalled but not discussed at length. However, I think it is also beneficial to have a thorough discussion of those concepts at as simple a level as possible, so I have provided one in another book called *The Wonderful World of Relativity.* The present book is self-contained and does not require knowledge of the first one, but a more basic text such as *The Wonderful World* or something similar is recommended as a preparation for this book.

The later chapters of the book go further than most undergraduate courses will want to go; they are intended to fill the gap between undergraduate and graduate study, and to offer general reading for the professional physicist.

Acknowledgements

I have, of course, learned relativity mostly from other people. All writers in this area have learned from the pioneers of the subject, especially Einstein, Lorentz, Maxwell, Minkowski and Poincaré. I am indebted also to tutors such N. Stone and W. S. C. Williams at Oxford University, and to authors who have preceded me, especially texts by (in alphabetical order) A. Einstein, R. Feynman, A. P. French, J.D. Jackson, H. Muirhead, W. Rindler, F. Rohrlich, R. Shankar, E. F. Taylor and J. A. Wheeler, and W.S.C. Williams.

Einstein emphasized the need to think of a reference frame in physical not abstract terms, as a physical entity made of rods and clocks. As a student I resisted this idea, feeling that a more abstract idea, liberated from mere matter, must be superior. I was wrong. The whole point of Relativity is to see that abstract notions of space and time are superfluous and misleading.

Feynman offered very useful guidance on how to approach things simply while retaining rigour. Readers familiar with his, Leighton and Sands' 'Lectures on Physics' will recognize that my treatment of Poynting's argument follows his quite closely, because I felt there was little room for improvement. I am happy to acknowledge this.

I learned a significant number of detailed points from Rindler's work; my contribution has been to clarify where possible. Appendix 2, for example, re-presents an argument I found in his book, with more comfort and explanation for the reader. The material in chapter 12 on radiation reaction and self-force is heavily indebted to Rohrlich. The presentation of General Relativity in chapter 14 owes much to J. Binney of Oxford University, as well as to W. Rindler once again.

Chapter 2

Basic ideas

The primary purpose of this chapter is to offer a way in for readers completely unfamiliar with Special Relativity, and to recall the main ideas for readers who have some preliminary knowledge of the subject. For the former category, Appendix 1 contains some of the basic arguments that will not be repeated in the main text (and that can be found in introductory texts such as *The Wonderful World of Relativity*). The right moment to turn to that appendix, if you need to, is after completing section 2.2.2 of this chapter.

In order to discuss space and time without being vague, it is extremely helpful to introduce the notion of a *reference body.* This is usually called an "*inertial frame of reference*", but that phrase is in some respects unfortunate. The phrase "frame of reference" is used in an abstract way in everyday language, but in physics we mean something more concrete: a large rigid physical object which could, in principle, exist in the vicinity of any system whose evolution we wish to discuss. Such a "reference body" clarifies what we mean when we talk of distance and time. By 'distance' we mean the number of particles or rods of the reference body between two places. The reference body keeps track of time as well, since the particles making it can be imagined to be tiny regular clocks (think of an atom with an internal vibration, for example). By 'time' at any given place we mean the number of repetitions of some such regularly repeating process ('clock') at that place.

"Frame of reference" and "reference body" are synonyms in physics. Most people like to think of a frame of reference as having the form of a scaffolding of ideally thin and rigid rods, with clocks attached. I sometimes like to think of it as a large brick (but one with the unusual property that others things can move through it unimpeded). It is a mistake to try to be too abstract here. Although the scaffolding or rigid body is not necessarily present, our reasoning about distance and time must be consistent with the fact that such a body might in principle be present in any region of spacetime.

An *observer* is a reasoning being who could in principle be situated at rest in some given frame

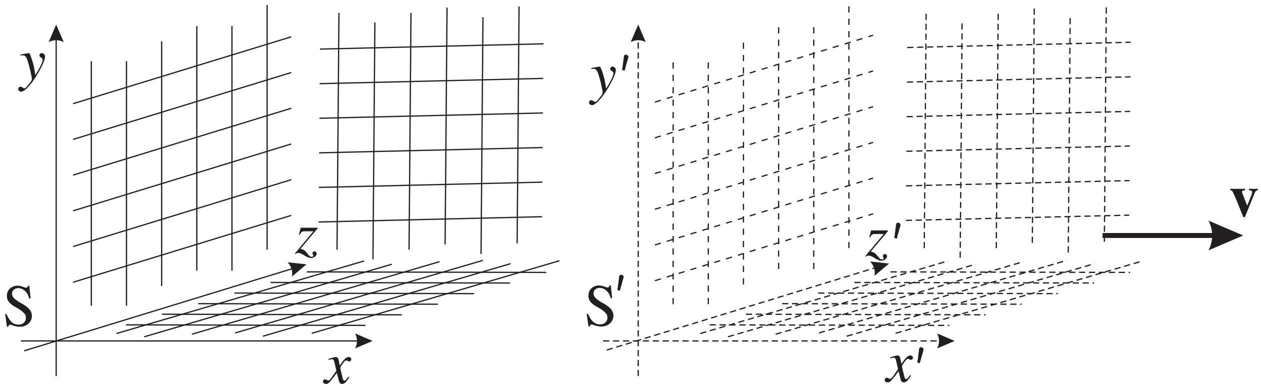

Figure 2.1: Two reference frames (=reference bodies) in standard configuration. S′ moves in the x direction relative to S, with its axes aligned with those of S. The picture shows the situation at the moment (defined in S) when the axes of S′ have just swept past those of S. The whole reference frame of S′ is in motion together at the same velocity $\mathbf{v}$ relative to S. Equally, the frame of S is in motion at velocity $-\mathbf{v}$ relative to S′.

of reference. We use the word 'observe' to mean *not* what the observer directly sees, but what he or she can deduce to be the case at each time and place in his/her reference frame. For example, suppose two explosions occur, and an observer is located closer to one than to the other in his own reference frame. If such an observer receives light flashes from the two explosions simultaneously, then he 'observes' (i.e. deduces) that the explosions were not simultaneous in his reference frame.

2.1 Classical physics

Let us briefly survey the connection between inertial reference frames according to classical physics, as developed by Galileo and Newton and others.

A crucial idea, first presented at length by Galileo, is the idea that the behaviour of physical systems is the same in any given inertial reference frame, irrespective of whether that frame may be in uniform motion with respect to others. For example, it is possible to play table tennis in a carriage of a moving railway train without noticing the motion of the train (as long as the rails are smooth and the train has constant velocity). There is no need to adjust one's calculations of the trajectory of the ball or the choice of force to apply using the bat: all the behaviour is the same as it would be in a motionless train. This idea, which we shall state more carefully in a moment, is called the Principle of Relativity; it is obeyed by both classical and relativistic physics.

When we analyze the motions of bodies, it is useful to introduce a coordinate system (in both space and time), which means we measure distances and times relative to a reference body (= inertial frame of reference). An *event* is a point in space and time. It is useful to know, for any given event, how the coordinates of the event relative to one reference body relate to the coordinates of the same event relative to another reference body. If reference frames F and F′ have all their axes aligned, but frame F′ moves along the positive x direction relative to F at speed v, then we say the reference frames are in *standard configuration* (figure 2.1). The coordinates of any given event, as determined in two reference frames in standard configuration, are related, *according to classical physics*, by

$$\begin{aligned} t' &= t, \\ x' &= x - vt, \\ y' &= y, \\ z' &= z. \end{aligned} \tag{2.1}$$

This set of equations is called the *Galilean transformation.* It can also be written in matrix notation as

$$\begin{pmatrix} t' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -v & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix}, \tag{2.2}$$

or

$$\begin{pmatrix} t' \\ \mathbf{r}' \end{pmatrix} = \mathcal{G} \begin{pmatrix} t \\ \mathbf{r} \end{pmatrix}, \tag{2.3}$$

where

$$\mathcal{G} \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ -v & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{2.4}$$

The inverse Galilean transformation is

$$\begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ v & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} t' \\ x' \\ y' \\ z' \end{pmatrix}. \tag{2.5}$$

which can also be written

$$\begin{pmatrix} t \\ \mathbf{r} \end{pmatrix} = \mathcal{G}^{-1} \begin{pmatrix} t' \\ \mathbf{r}' \end{pmatrix}. \tag{2.6}$$

The reader is invited to verify this, i.e. check that the matrix given in (2.5) is indeed the inverse of $\mathcal{G}$.

Matrix notation makes it easy to check things like the effect of transforming from one reference frame to another and then to a third. For example, the net effect of transforming to another frame and then back to the first is given by $\mathcal{G}^{-1}\mathcal{G}$ which is, of course, the identity matrix.

2.2 Special relativity

2.2.1 The postulates of Special Relativity

Turning now to Special Relativity, we shall find that the Principle of Relativity is still obeyed, but the Galilean transformation fails.

The Main Postulates of Special Relativity are

> **Postulate 1,** "Principle of Relativity": *The motions of bodies included in a given space are the same among themselves, whether that space is at rest or moves uniformly forward in a straight line.*
>
> **Postulate 2,** "Light speed postulate":
> Version A: *There is a finite maximum speed for signals.*
> Version B: *There is an inertial reference frame in which the speed of light in vacuum is independent of the motion of the source.*

The Principle of Relativity (Postulate 1) is obeyed by classical physics; the Light Speed Postulate is not. The Principle of Relativity can also be stated

> The laws of physics take the same mathematical form in all inertial frames of reference.

In Postulate 2, either version A or version B is sufficient on its own to allow Special Relativity to be developed. Version A does not mention light; this makes it clear that Special Relativity underlies all theories in physics, not just electromagnetism. For this reason version A is preferred. However we will preserve the practice of calling this postulate the "Light speed postulate" because in vacuum, far from material objects, light waves move at the maximum speed for signals. With this piece of information about light, one can use either version to derive the other.

Einstein used version B of the Light Speed Postulate. It is often stated as "the speed of light is independent of the motion of the source." In this statement the fact that motion can only ever be relative motion is taken for granted, and it is a statement about what is observed in any reference frame. In our version B we chose to make a slightly more restricted statement (picking just one reference frame), merely because it is interesting to hone ones assumptions down to the smallest possible set. By combining this with Postulate 1 it immediately follows that all reference frames will have this property.

In order to make clear what is assumed and what is derived, it is useful to add two further postulates to the list:

> **Postulate 0,** "Euclidean geometry": *The rules of Euclidean geometry apply to all spatial measurements within any given inertial reference frame.*
>
> **Postulate 3,** "Conservation of momentum": *Internal interactions among the parts of an isolated system cannot change the system's total momentum, where momentum is a vector function of rest mass and velocity.*

Postulate 0 (Euclidean geometry) is obeyed by Special Relativity but not by General Relativity. Postulate 3 (conservation of momentum) allows the central elements of dynamics to be deduced, including the famous formula "$E = mc^2$" (that formula cannot be derived from the Main Postulates alone).

2.2.2 Central ideas about spacetime

Recall that a 'point in spacetime' is called an *event*. This is something happening at an instant of time at a point in space, with infinitesimal time duration and spatial extension. For an example, tap the tip of a pencil once on a table top, or click your fingers.

A *particle* is a physical object of infinitesimal spatial extent, which can exist for some some extended period of time. The line of events which gives the location of the particle as a function of time is called its *worldline*, see figure 2.2.

If two events have coordinates (t_1, x_1, y_1, z_1) and (t_2, x_2, y_2, z_2) in some reference frame, then the quantity

$$s^2 = -c^2(t_2 - t_1)^2 + (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2 \tag{2.7}$$

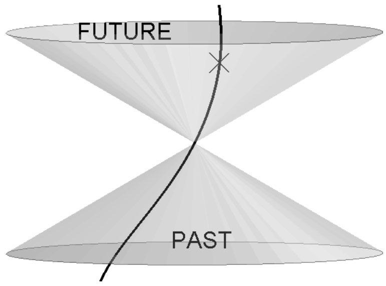

Figure 2.2: A spacetime diagram showing a worldline and a light cone (past and future branches). The cross ($\times$) marks an example event. The apex of the cone is another event.

is called the *squared spacetime interval* between them. Note the crucial minus sign in front of the first term. We emphasize it by writing (2.7) as

$$s^2 = -c^2 \Delta t^2 + \Delta x^2 + \Delta y^2 + \Delta z^2. \tag{2.8}$$

If $s^2 < 0$ then the time between the events is sufficiently long that a particle or other signal (moving at speeds less than c) could move from one event to the other. Such a pair of events is said to be separated by a *timelike interval.* If $s^2 > 0$ then the time between the events is too short for any physical influence to move between them. This is called a *spacelike interval.* If $s^2 = 0$ then we have a *null interval*, it means that a light pulse or other light-speed signal could move directly from one event to the other.

Although the parts t_i, x_i, y_i, z_i needed to calculate an interval will vary from one reference frame to another, we will find in chapter 3 that the net result, s^2, is **independent of reference frame**: all reference frames agree on the value of this quantity. This is similar to the fact that the length of a vector is unchanged by rotations of the vector. A quantity whose value is the same in all reference frames is called a *Lorentz invariant* (or Lorentz scalar). **Lorenz invariants play a central role Special Relativity.**

The set of events with a null spacetime interval from any given event lie on a cone called the *light cone.* The part (or 'branch') of this cone extending into the past is made of the worldlines of photons that form a spherical pulse of light collapsing onto the event, the part extending into the future is made of the worldlines of photons that form a spherical pulse of light emitted by the event. The cone is an abstraction: the incoming and outgoing light pulses don't have to be there. The past part of the light cone surface of any event A is called the *past light cone* of A, the future part of the surface is called the *future light cone* of A. The whole of the future cone (i.e. the body of the cone as well as the surface) is called the *absolute future* of A, it consists of all events which could possibly be influenced by A (in view of the Light Speed Postulate). The whole of the past cone is called the *absolute past* of A, it consists of all events which could possibly influence A. The rest of spacetime, outside either branch of the light cone, can neither

Einstein's train.
Suppose a fast-moving train is moving past a platform, and suppose that at the moment when the front of the train reaches the far end of the platform (where it is about to leave the station), a firecracker explodes there, leaving scorch marks on the train and platform. Similarly, when the back of the train arrives at the start of the platform (at the other end of the station), a firecracker explodes there, leaving scorch marks on the train and platform. We consider a train whose length is such that, in the reference frame of the platform, the lengths of train and platform are the same. In this case the two explosions are simultaneous in the reference frame of the platform. The flashes of light emitted by the explosions therefore arrive at the centre of the platform at the same time. However, the flashes of light do not arrive at the centre of the train together. An observer standing on the platform finds that when the flashes arrive at him, the train has moved on, so that the flash from the front of the train has already moved past the centre of the train, and the flash from the back has not yet arrived at the centre of the train. It follows that *the flash from the front of the train arrives at a passenger seated in the middle of the train before the flash from the back does.*
Observers in all reference frames must agree with this fact, i.e. one flash arrives at the passenger before the other, because the flashes could be arranged to trigger events at the passenger. Suppose, for example, that he carries a device which will smash a glass if the flashes arrive simultaneously (or if the rear flash arrives first). If the observer at rest on the platform finds that the glass is not smashed, then it is not, irrespective of which reference frame we adopt for the purpose of calculating time and space intervals. It follows that, in the reference frame of the train (i.e. that in which the train is at rest), the front flash arrives at the passenger before the rear flash does. Also, the two scorch marks at the front and back of the train are equidistant from the passenger in the middle of the train, and the light pulses have the same speed (by postulate 2). It follows that the firecracker explosion at the front of the train must have happened first, *before* the one at the back, and *not* simultaneous with it, in the reference frame of the train.
We infer that simultaneity is a relative concept: it depends on reference frame. It also follows that in the reference frame of the train, the train and the platform are not of the same length: the train must be longer than the platform.

influence nor be influenced by A. It consists of all events with a spacelike separation from A.

The single most basic insight into spacetime that Einstein's theory introduces is the *relativity of simultaneity*: two events that are simultaneous in one reference frame are not necessarily simultaneous in another. In particular, if two events happen simultaneously at different spatial locations in reference frame F, then they will not be simultaneous in any reference frame moving relative to F with a non-zero component of velocity along the line between the events. An example is furnished by "Einstein's train," see box.

By careful argument from the postulates one can connect timing and spatial measurements in one inertial reference frame to those in any other inertial reference frame in a precise, quantitative way. In the next chapter we shall introduce the *Lorentz transformation* to do this in general. Arguments for some simple cases were presented in *The Wonderful World*; these are summarised in Appendix 1.

2.3 Matrix methods

By writing down the Galilean transformation using a matrix, we already assumed that the reader has some idea what a matrix is and how it is used. However, in case matrices are unfamiliar, we will here summarize the matrix mathematics we shall need. This will not substitute for a more lengthy course of mathematical training, but it may be a useful reminder.

A matrix is a table of numbers. We will only need to deal with real matrices (until chapter 17) so the numbers are real numbers. In an "$n \times m$" matrix the table has n rows and m columns. Here is a 2×3 matrix, for example:

$$\begin{pmatrix} 1.2 & -3.6 & 8 \\ 2 & 4.5 & 2 \end{pmatrix}. \tag{2.9}$$

If either n or m is 1 then we have a vector; if both are 1 than we have a scalar.

A vector of 1 row is called a row vector; a vector of 1 column is called a column vector:

$$\text{row vector: } (1, -3, 2), \qquad \text{column vector: } \begin{pmatrix} 1 \\ -3 \\ 2 \end{pmatrix}.$$

The sum of two matrices, written as $A+B$, is only defined (so it is only a legal operation) when A and B have the same shape, that is, the two matrices have the same number of rows n, and they also have the same number of columns m (but n does not have to equal m). $A+B$ is then defined to mean the matrix formed from the sums of the corresponding components of A and B. To be precise, if M_{ij} refers to the element of matrix M in the i'th row and j'th column, then the matrix sum is defined by

$$M = A + B \qquad \Leftrightarrow \qquad M_{ij} = A_{ij} + B_{ij}.$$

This rule applies to vectors and scalars too, since they are special cases of matrices, and it agrees with the familiar rule for summing vectors: add the components.

The product of two matrices, written as AB, is only defined (so it is only a legal operation) when A and B have appropriate shapes: the number of *columns* in the first matrix has to equal the number of *rows* in the second matrix. For example, a 2×3 matrix can multiply a 3×5 matrix, but it cannot multiply a 2×3 matrix. The product is defined by the mathematical rule

$$M = AB \qquad \Leftrightarrow \qquad M_{ij} = \sum_k A_{ik} B_{kj}. \tag{2.10}$$

It is important to note that this rule is **not commutative**: AB is not necessarily the same as BA. The rule is important in order to have a precise definition, but the use of subscripts and the sum can leave the operation obscure until one tries a few examples. It amounts to the following. You have to work your way through the elements of M one by one. To obtain the element of M on the i'th row and j'th column, take the i'th row of A and the j'th column of B. Regard these as two vectors and evaluate their scalar product: that is, 'dive' the row of A onto the column of B, multiply corresponding elements, and then sum. The result is the value of M_{ij}.

The only way to become familiar with matrix multiplication is by practice. By applying the rule, you will find that if a $k \times n$ matrix multiplies a $n \times m$ matrix then the result is a $k \times m$ matrix. This is a very useful check to keep track of what you are doing.

The whole point of matrix notation is that much of the time we can avoid actually carrying out the element-by-element multiplications and additions. Instead we manipulate the matrix symbols. For example, if $A + B = C$ and $A - B = D$ then we can deduce that $C + D = 2A$ without needing to carry out any element-by-element analysis. The following mathematical results apply to matrices (as the reader can show by applying the rules developed above):

$$\begin{aligned} A + B &= B + A \\ A + (B + C) &= (A + B) + C \\ (AB)C &= A(BC) \\ A(B + C) &= AB + AC. \end{aligned}$$

We shall mostly be concerned with square matrices and with vectors. The square matrices will be mostly 4×4, so they can be added and multiplied to give other 4×4 matrices. A square matrix can multiply a *column* vector, giving a result that is a column vector (since a 4×4 matrix multiplying a 4×1 matrix gives a 4×1 matrix). For example

$$\begin{pmatrix} 1.2 & -3.6 & 8 & 2 \\ 2 & 4.5 & 2 & 0.5 \\ -1 & 5 & 1 & -0.5 \\ -2 & 3.2 & 3 & -5 \end{pmatrix} \begin{pmatrix} 1 \\ -3 \\ 2 \\ -4 \end{pmatrix} = \begin{pmatrix} 20 \\ -9.5 \\ -12 \\ 14.4 \end{pmatrix}.$$

A square matrix can be multiplied from the left by a *row* vector, giving a result that is a row vector (since a 1×4 matrix multiplying a 4×4 matrix gives a 1×4 matrix).

Matrix inverse

Many, but not all, square matrices have an *inverse*. This is written M^{-1} and is defined by

$$MM^{-1} = M^{-1}M = I \tag{2.11}$$

where I is the *identity matrix*, consisting of 1's down the diagonal and zeros everywhere else. For example, in the 4×4 case it is

$$I = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The identity matrix has no effect when it multiplies another matrix: $IM = MI = M$ for all M. Inverses of non-square matrices can also be defined, but we shall not need them.

There is no definition of a 'division' operation for matrices (in the sense of one matrix 'divided by' another), but often multiplication by the inverse achieves what might be regarded as a form of division. For example, if $AB = C$ and A has an inverse, then by premultiplying both sides by A^{-1} we obtain $A^{-1}AB = A^{-1}C$, and therefore $B = A^{-1}C$ (by using the fact that $A^{-1}A = I$ and $IB = B$).

The inverse of a 2×2 matrix is easy to find:

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \Leftrightarrow \quad M^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Here the inverse exists when $ad - bc \neq 0$ and you can check that it satisfies (2.11).

There is also a general rule on how to find the inverse of a matrix of any size—you should consult a mathematics textbook when you need it.

The inverse of a product is the product of the inverses, but you have to reverse the order:

$$(AB)^{-1} = B^{-1}A^{-1}.$$

Proof: $(AB)(B^{-1}A^{-1}) = A(B^{-1}B)A^{-1} = AA^{-1} = I$ and you are invited to show by a similar method that $(B^{-1}A^{-1})(AB) = I$.

Transpose and scalar product

The *transpose* of a matrix, written M^T, is the matrix obtained by swapping the rows and columns. To be precise:

$$A = M^T \qquad \text{means} \qquad A_{ij} = M_{ji}.$$

For example, the transpose of the matrix displayed in (2.9) is

$$\begin{pmatrix} 1.2 & 2 \\ -3.6 & 4.5 \\ 8 & 2 \end{pmatrix}.$$

The transpose of a row vector is a column vector, and the transpose of a column vector is a row vector.

The following results are useful:

$$(A+B)^T = A^T + B^T \tag{2.12}$$
$$(AB)^T = B^T A^T \tag{2.13}$$
$$(A^T)^{-1} = (A^{-1})^T \tag{2.14}$$

Note the order reversal in (2.13). You can easily prove this result using (2.10). Then (2.14) follows since if M is the inverse of A^T then we must have $A^T M = I$, taking the transpose of both sides gives $M^T A = I^T = I$ and hence $M^T = A^{-1}$ and the result follows.

The product of a row vector and a column vector of the same length is often useful because it is simple: it is a 1×1 matrix, in other words a scalar. If we start with a pair of column vectors $\mathbf{u}$ and $\mathbf{v}$ of the same size, then we can obtain such a scalar by

$$\mathbf{u}^T \mathbf{v}. \tag{2.15}$$

This comes up often, so it is given a name: it is called the *scalar product* or *inner product* of the vectors. (The inner product of a pair of row vectors would be $\mathbf{u}\mathbf{v}^T$.) You can calculate it by multiplying corresponding components and summing. For example if $\mathbf{u}$ has components u_1, u_2, u_3 and $\mathbf{v}$ has components v_1, v_2, v_3 then

$$\mathbf{u}^T \mathbf{v} = u_1 v_1 + u_2 v_2 + u_3 v_3. \tag{2.16}$$

Most science or mathematics students will meet the scalar product first in the context of vector analysis in space, where one is typically dealing with three-component vectors representing things like displacement, velocity and force. In this context it can be convenient not to be too concerned whether the vectors are row or column vectors, and so the dot notation is introduced: the scalar product is written $\mathbf{u} \cdot \mathbf{v}$. In relativity we will be dealing with 4-component vectors in time and space, and for them we will introduce a special meaning for the dot notation and for the phrase 'scalar product'. In chapter 9 we shall also introduce the 'outer product' which enables a square matrix to be obtained from a pair of vectors.

Chapter 22

Appendix 1. Some basic arguments

22.1 Simultaneity and radar coordinates

If two events happen at the same place in some reference frame, then it is easy to define the time and distance between them. The distance in this case is zero, and the time is determined by the number of ticks of a clock situated at the location of both events. (In practice we would also need to agree some sort of standard of time—this is currently done by observing the oscillations of the nucleus of a caesium atom in vacuum, but the details are not necessary in order to study Relativity, we just need to agree that some such standard can be defined. The Principle of Relativity ensures that the definition applies in all reference frames equally.)

For events happening at different places, the time and distance between them has to be defined carefully. A convenient method is first to use 'radar coordinates,' and then derive times and distances from those. For any given reference frame F, we can pick a position to serve as the spatial origin O of a coordinate system. The particle located at such an origin will have a straight worldline. Now consider an arbitrary event R, not at the origin. We imagine a 'radar echo location' scenario. That is, at time t_1 the particle at the origin of F sends out an electromagnetic pulse, propagating at the speed of light c (think of it as a radio wave pulse or a flash of light, for example). We suppose that the pulse is reflected off some object present at event R (so R is the event of reflection), and then the pulse propagates back to the particle at the origin, arriving there at time t_2. The times (t_1, t_2) constitute the *radar coordinates* of event R in reference frame F. Together with the direction of travel of the pulses, they suffice to determine the position and time of R in frame F. Let's see how.

First, since the speed of light is independent of processes such as reflection, the outgoing and

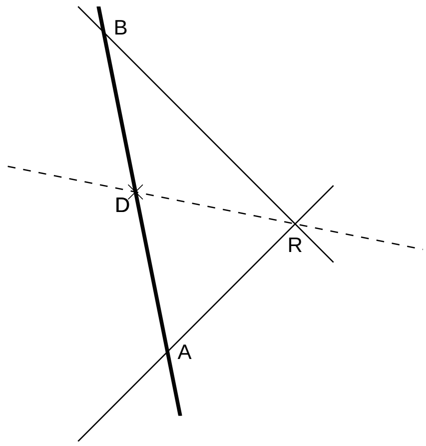

Figure 22.1: Identifying simultaneous events on a spacetime diagram. AB is a straight timelike worldline, so it can be the worldline of a particle in uniform motion, at the spatial origin of an inertial reference frame F. AR and RB are photon worldlines. D is half way between A and B. In the reference frame whose time axis is AB (i.e. frame F), R must be simultaneous with D, because the photon travel times of AR and RB must be equal. Applying the argument to further events, one concludes that all events along the dashed line are simultaneous in frame F.

incoming pulse must have the same speed c in F. It follows that the outgoing pulse takes the same time to get to R as the incoming one takes to come back, so R must occur at a time half way between t_1 and t_2, i.e.

$$t_R = \frac{t_2 + t_1}{2}$$

in frame F. Also, since the pulse traveled out and back at the speed c, the distance from O to R must be

$$x_R = \frac{c(t_2 - t_1)}{2}$$

in frame F. This pair of equations allows us to convert easily between radar coordinates and ordinary coordinates. It means that we only need one clock, at the origin, together with the Postulates of Relativity, to define a complete coordinate system for any given reference frame.

The worldline of a particle at the origin of F serves as the 'time axis' of F. The set of events simultaneous with R (according to reference frame F) consists of all those having the same value of the radar echo time $(t_2 + t_1)/2$. On a spacetime diagram these form a line that can be constructed as shown in figure 22.1. If we choose the scales such that photon worldlines have slope 45° on the diagram, then a line of simultaneity or 'distance axis' for frame F makes the same angle with a photon worldline as the time axis of F does. In other words, the distance axis and time axis of any reference frame are oriented such that the angle between them is bisected by a photon worldline. This is the graphical representation of the fact that the speed of light is the same in all reference frames.

The spacetime diagram construction allows us to see the relativity of simultaneity very easily: sets of events that are simultaneous in one reference frame are not simultaneous on another, because the lines of simultaneity associated with different frames cut through spacetime in different directions.

As a result of this, the temporal sequence of events with spacelike separation can depend on reference frame. The temporal sequence of events with timelike separation is, however, independent of reference frame, because lines of simultaneity never slope more than photon worldlines. This preserves the logic of cause and effect, so that Special Relativity is a physically and logically consistent theory.

22.2 Proper time and time dilation

For any pair of events having a timelike interval between them, the *proper time* is defined to be the time between such a pair of events, as observed in the reference frame in which the events happen at the same place. By definition, the proper time as related to the invariant interval by

$$\tau = \frac{s}{c} \tag{22.1}$$

(by using (2.7) with $\tau = t_2 - t_1$, and $x_2 = x_1,\ y_2 = y_1,\ z_2 = z_1$).

If a particle or system is in uniform motion at speed v relative to some reference frame F, then the time interval between a pair of events at the particle is related to the proper time by

$$T = \frac{\tau}{\sqrt{1 - v^2/c^2}} = \gamma\tau \tag{22.2}$$

where

$$\gamma(v) \equiv \frac{1}{\sqrt{1 - v^2/c^2}}. \tag{22.3}$$

This is called *time dilation*, because $T \geq \tau$. It means that all systems evolve more slowly if they are moving then if they are at rest. For example, a rabbit carried along in a rocket moving at $v = 0.999c$ would not experience the slightest difference in the laws of physics which determine its metabolism, and yet it would live for 150 years (i.e. 22 times longer than slow rabbits), when the time is measured by clocks relative to which it has this high speed. This effect has not been observed for rabbits, owing to experimental difficulties, but it has been observed for muons and other particles, and for atomic clocks.

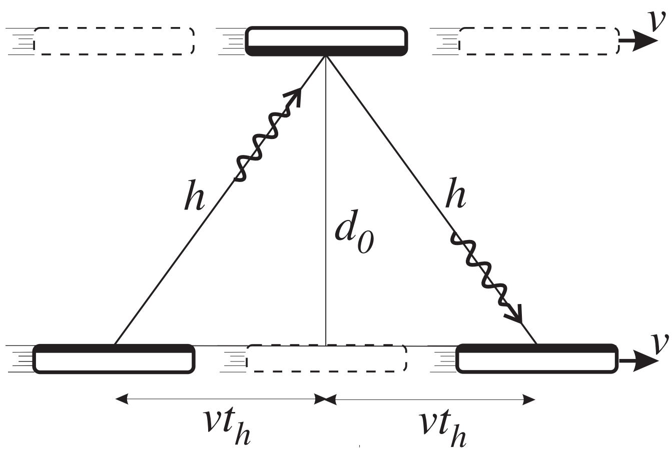

Figure 22.2: The photon clock. A pair of mirrors is attached to a rod of rest length d_0, such that light can bounce between them. In the rest frame, the time for a round trip of a light pulse between the mirrors is $\tau = 2d_0/c$. The figure shows the situation observed in a reference frame travelling to the left relative to this 'clock.' A such a reference frame, the mirrors have a speed v to the right. Let the round trip time in this reference frame be t. To complete a round trip the light pulse must travel a distance $2h$. Clearly since $h > d_0$ we must find $t > \tau$: time dilation. Pythagorus' theorem, $h^2 = d_0^2 + (vt_h)^2$ where $t_h = t/2$. Therefore $t = 2h/c = 2(d_0^2 + v^2t^2/4)^{1/2}/c$. Solving for t one finds $t = \gamma\tau$. The argument hinges on the fact that the events 'pulse leaves' and 'pulse returns' are just that: *events*, so t is the time between the *same two events* as those whose time separation is τ in the rest frame.

To deduce the time dilation factor, one can argue from the invariance of the interval (see exercises) but the conceptually most simple derivation is arguably the 'photon clock' argument, see figure 22.2.

22.3 Lorentz contraction

If a pair of events has a spacelike separation, then there exists a reference frame in which they are simultaneous. The distance between the events, as observed in such a reference frame, is called the *proper distance* L_0. Owing to the fact that we normally study the evolution of particles and systems, the concept of proper time is much more useful in practice than the concept of proper distance. However, distance is also needed in order to get a complete description when changing from one reference frame to another.

A physical object can be regarded as a set of worldlines (those of the particles of the object). If these wordlines are straight and parallel then the object has constant velocity and fixed size. The spatial size of an object is defined as the size of the region of space it occupies at any instant of time. Owing to the relativity of simultaneity, this concept is a relative one, i.e. it is well-defined only once a reference frame is specified, and the value obtained for the size can depend on which reference frame is adopted.

Suppose we choose some direction in space, and take an interest in sizes of physical objects along this direction. The length of an object in its own rest frame (i.e. the frame in which its velocity is zero), along the chosen direction, is called its *rest length.* In any reference frame moving with respect to the rest frame, the object has some non-zero velocity $\mathbf{v}$. In such a reference frame the length of the object, along the chosen direction, is given by

$$L = \frac{L_0}{\gamma(v_{\parallel})}, \qquad (22.4)$$

where $v_{\parallel}$ is the component of velocity along the chosen direction. Since $L \leq L_0$ this is called *Lorentz contraction* or *space contraction.* It means that an object in motion is contracted along the direction of motion compared to its size when at rest. For example, a rabbit carried along in a rocket moving at $v = 0.999c$ would have physical dimensions approximately 15 cm $\times$ 20 cm $\times$ 1 cm, when measured by rods relative to which it moves at this high speed.

To deduce this result one can use a photon clock once again, but now the clock is oriented along the direction of motion. In other words we are using a radar method again.

Consider a simple rod or stick. Suppose a pulse of light is emitted from one end of the stick and travels to the other end, where it is reflected and comes back to the first end. In the rest frame of the stick, the time between emission and final reception must be $2L_0/c$ where L_0 is the stick's rest length. The events of emission and final reception happen at the same place in

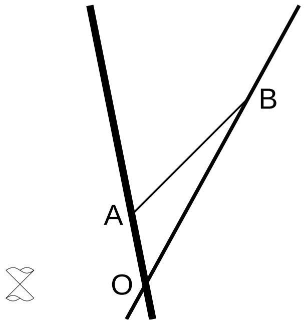

Figure 22.3: A simple set of events suited to reasoning about the Doppler effect. Two particles (with relative velocity less than c) pass one another at the origin O. At event A the first particle sends a light signal to the second particle, where it arrives at event B.

the rest frame of the stick, so their time separation in that frame is the proper time between them, $\tau = t_2 - t_1$. The time between these events in any other frame must therefore be

$$T = \gamma\tau = \gamma \frac{2L_0}{c}.$$

The stick singles out a direction in space by its own axis, and we now consider a reference frame in which it moves in that direction. In such a reference frame, the emitted light pulse moves at speed c and the far end of the stick moves at speed v. It follows that the time taken for the light pulse to reach the point of reflection is $L/(c-v)$ and the time taken to come back is $L/(c+v)$, where L is the length of the stick in the new reference frame. Hence

$$T = \frac{L}{c-v} + \frac{L}{c+v} = \frac{2cL}{c^2 - v^2} \quad \Rightarrow L = \frac{c^2 - v^2}{c^2}\gamma L_0 = \frac{L_0}{\gamma}.$$

This is eq. (22.4) in the case where $v_{\parallel} = v$. The general case is treated in chapter 3.1.

22.4 Doppler effect, addition of velocities

Suppose two particles, moving along a line, pass one another at event O, then at event A the first particle sends a light signal to the second, where it arrives at event B; see figure 22.3.

We take O as the origin of position and time. If the relative velocity of the particles is $\mathbf{v}$ then, in the reference frame F of the emitter, the reception event takes place a distance $d = vt_B$ away, so the signal travel time is $d/c = vt_B/c$. It follows that $t_B = t_A + (v/c)t_B$, therefore

$$t_B = \frac{t_A}{1 - v/c}. \tag{22.5}$$

Events O and B take place at the same place in the reference frame F$'$ of the *receiver*, so their time separation t'_B is a proper time in that frame, so $t_B = \gamma t'_B$, hence

$$\frac{t'_B}{t_A} = \frac{1}{\gamma}\frac{1}{1 - v/c} = \sqrt{\frac{1 - v^2/c^2}{(1 - v/c)^2}} = \sqrt{\frac{1 + v/c}{1 - v/c}}. \tag{22.6}$$

We can use this result to deduce the Doppler effect for light waves. All we need to do is suppose that the emitter emits regularly space signals, once every time t_A in his reference frame, then the above argument applies to all the signals, and the receiver will receive them spaced in time by t'_B as given by (22.6). This set of signals could in fact be one continuous stream of light waves, with period t_A. Then the event A could be, for example, "the electric field of the light wave at the emitter is at a maximum", and the event B could be "the electric field of the light wave at the receiver is at a maximum". Since we define the period of a wave to be the time interval between successive maxima, it follows that (22.6) relates the periods observed at the emitter and receiver. By taking the inverse, we obtain the relationship between the frequencies. Hence if light waves of frequency ν_0 are emitted by a particle, then the frequency observed by any particle moving directly away from the emitter at speed v is given by

$$\nu = \sqrt{\frac{1 - v/c}{1 + v/c}}\,\nu_0. \tag{22.7}$$

This is called the *longitudinal Doppler effect* or just Doppler effect. It permits one to deduce, for example, the speed of a star relative to Earth, from the frequency of the received light, if one has independent evidence of what the emitted frequency was.

Now we shall use the Doppler effect to deduce a formula concerning relative velocities. Suppose F$'$ moves relative to F with speed u, and F$''$ moves relative to F$'$ with speed v (i.e. v is the speed of F$''$ *as observed in frame* F$'$), all motions being along the same direction, then we can calculate the speed w of F$''$ relative to F by using the Doppler effect formula three times, as follows:

$$\nu' = \sqrt{\frac{1 - u/c}{1 + u/c}}\,\nu, \quad \text{and} \quad \nu'' = \sqrt{\frac{1 - v/c}{1 + v/c}}\,\nu' = \sqrt{\frac{1 - w/c}{1 + w/c}}\,\nu.$$

It follows that

$$\left(\frac{1 - u/c}{1 + u/c}\right)\left(\frac{1 - v/c}{1 + v/c}\right) = \frac{1 - w/c}{1 + w/c}$$

and after a little algebra we obtain

$$w = \frac{u + v}{1 + uv/c^2}. \tag{22.8}$$

This is the formula for "relativistic addition of velocities." The generalization to velocities in any direction is presented in chapter 3. The formula predicts that w is never greater than c as long as u and v are both less than or equal to c; this is in agreement with the Light Speed Postulate. The result predicted by classical physics is $w = u + v$; the relativistic formula reproduces this in the limit $uv \ll c^2$.

Note that the velocities in the formula (22.8) are all what we call 'relative velocities' and they concern three different reference frames. There is another type of velocity that can be useful in calculations, which we shall refer to as 'closing velocity.' The concept of 'closing velocity' applies in a *single* reference frame, and it refers to *the rate of change of distance between* two objects, all distances and times being measured in a single reference frame. When both objects are moving relative to the reference frame, a closing velocity is *not* necessarily the velocity of any physical object or signal, and it can exceed the speed of light. For example, an observer standing at the collision point of a modern particle accelerator will observe a bunch of particles coming towards him from the right at a speed very close to c, and another bunch approaching from the left at a speed very close to c. The positions of the two particle bunches can be written $z_1 = d - vt$ and $z_2 = -d + vt$ in the rest frame of such an observer, where $v \simeq c$ and d is a constant (equal to half the distance between the bunches at $t = 0$). He finds the distance between the bunches to be $z_1 - z_2 = 2d - 2vt$. The rate of change of this distance is $2v$. The two particle bunches may then be said to have a 'closing velocity' of $2v$, and this can easily exceed the speed of light. Nevertheless, the *relative velocity* of the bunches is $w = 2v/(1 + v^2/c^2)$, which is less than c. The relative velocity is the velocity which an observer moving along with one bunch will find the other to have.

An example of closing velocity can be found in equation (22.5). We can obtain that equation by looking at the situation at time t_A. At this moment, the emitter considers that the receiver is a distance vt_A away, and he finds that the signal sent out at speed c has a closing velocity with the receiver of $c - v$, therefore it takes a time $vt_A/(c - v)$ to close with the receiver. Hence $t_B = t_A + vt_A/(c - v)$ which you can verify agrees with (22.5).