Special relativity: The geometric formulation

3

- The constancy of light speed c allows us to interpret geometrically the relativistic invariant interval s as a length in the 4D manifold, called Minkowski spacetime, with a metric equal to diag(-1,1,1,1).
- The physics of SR, such as time dilation and length contraction, follows directly from the Lorentz transformation, which is a rotation in Minkowski spacetime.
- We introduce 4-vectors as the simplest example of tensors in Minkowski space, and construct their scalar products. Besides the 4-position vector x^μ, we also introduce the 4-velocity U^μ and 4-momentum p^μ, with components that transform into each other under a Lorentz transformation.
- The principal features of a spacetime diagram and its representation of Lorentz transformations are presented. In particular, the causal structure of SR is clarified in terms of lightcones in a spacetime diagram.

The new kinematics of special relativity discussed in the previous chapter can be expressed elegantly in the geometric formalism of a four-dimensional manifold, known as **spacetime**, as first formulated by Herman Minkowski. The following are the opening words of an address he delivered at the 1908 Assembly of German National Scientists and Physicians held in Cologne, Germany.

> The views of space and time which I wish to lay before you have sprung from the soil of experimental physics, and therein lies their strength. They are radical. Henceforth space by itself, and time by itself, are doomed to fade away into mere shadows, and only a kind of union of the two will preserve an independent reality.

In this geometric formulation, the stage on which physics takes place is Minkowski spacetime with the time coordinate being on an equal footing with the spatial coordinates. Here we introduce **flat spacetime**, which is the spacetime that is appropriate for the physics of special relativity. This prepares us for the study of the larger framework of **curved spacetime** in general relativity.

3.1 Minkowski spacetime

The fact that measurement results for time, as well as space, may be different for different observers means that time must be treated as a coordinate in much the same way as the spatial coordinates. The unification of space and time can be made explicit when space and time coordinates appear in the same position vector. A coordinate transformation may be regarded as a rotation in this 4D space, with the possibility of changing space and time coordinates (t, x, y, z) into each other—in much the same way the (x, y, z) coordinates change into each other under an ordinary 3D rotation. The 4D Minkowski spacetime has coordinates $\{x^\mu\}$. Henceforth, all Greek indices[1] such as μ will have the range 0, 1, 2, 3. Namely,

$$x^\mu = \left(x^0,\ x^1,\ x^2,\ x^3\right) = (ct,\ x,\ y,\ z)\,. \tag{3.1}$$

We have already shown in Section 2.4.2 that the interval between the coordinate origin and the position coordinate x^μ,

$$s^2 = -c^2t^2 + x^2 + y^2 + z^2, \tag{3.2}$$

is a relativistic invariant. Namely, if we were to make a coordinate transformation, with the new coordinates being $x'^\mu = (ct',\ x',\ y',\ z')$, the corresponding spacetime interval $s'^2 = -c^2t'^2 + x'^2 + y'^2 + z'^2$ would have the same value $s'^2 = s^2$. In Section 2.4.2 we demonstrated this invariance by a calculation using the explicit form of the Lorentz transformation. In Box 3.1, we present another derivation, showing that this result follows directly from the basic postulate that the speed of light, c, is the same in every reference frame: s is absolute because c is absolute.

This interval Δs has the physical significance of being directly related to the proper time: $\Delta s^2 = -c^2\Delta\tau^2$. Furthermore, we recall that this rest-frame time τ is related to the coordinate time t by the relativistic time-dilation formula of $t = \gamma\tau$. As discussed in Box 3.1, for a light ray, this interval vanishes; therefore, the concept of proper time is not applicable for a light ray. This is entirely consistent with the fact that there does not exist a coordinate frame in which the light velocity vanishes—that is, a frame in which light is at rest.

In 4D Euclidean space with Cartesian coordinates (w, x, y, z), the invariant length[2] is given as $s^2 = w^2 + x^2 + y^2 + z^2$. The minus sign in front of the c^2t^2 term in (3.2) means that if we regard ct as the fourth dimension, the relationship between coordinate and length measurements differs from that in Euclidean space. We say that Minkowski space is pseudo-Euclidean. In Section 3.1.1, we shall introduce generalized coordinates and distance measurements (via the metric) in Minkowski spacetime. In subsequent chapters, the same formalism will be shown to be applicable to coordinates in the warped spacetime of general relativity.

[1] Superscripts rather than subscripts are being used because this is the customary mathematical notation for coordinates in relativity, and it is used because relativity is expressed in tensor notation, which uses superscript indices for "contravariant components" of vectors and subscript indices for "covariant components" of vectors. Tensor notation will be explained in detail in Chapter 12. Until then, we shall always denote components of a 4D vector by using a superscript index. That is, we only need to work with contravariant components.

[2] We are familiar with the idea that in 3D Euclidean space the relation between coordinates and squared length is $l^2 = x^2 + y^2 + z^2$; hence, this straightforward generalization to 4D Euclidean space. However, in the following presentation, we shall refer to any such (quadratic in coordinates) invariant as a squared "length," regardless of whether we can actually visualize it as a quantity that can be measured by a yardstick or not.

Box 3.1 Δs is absolute because c is absolute

In the geometric formulation of special relativity, the spacetime interval

$$\Delta s^2 = \Delta x^2 + \Delta y^2 + \Delta z^2 - c^2\Delta t^2, \tag{3.3}$$

where $\Delta x = x_2 - x_1$, etc. plays a central role. Here we show that the invariance under a Lorentz transformation of this interval follows directly from the basic postulate of special relativity: the speed of light, c, is the same in every inertial reference frame. The interval Δs is absolute because c is absolute (Landau and Lifshitz, 1975).

First consider the special case in which the two events, $(\vec{x}_1, t_1)$ and $(\vec{x}_2, t_2)$, are connected by a light signal. The interval Δs^2 must vanish because in this case $|\Delta\vec{x}|/\Delta t = c$. When observed in another frame O', this interval also has a vanishing value $\Delta s'^2 = 0$, because the velocity of light remains the same in the new frame O'. From this, we infer that for *any* interval Δs connecting two events (not necessarily by a light signal), Δs and $\Delta s'$ must always be proportional to each other (because, if Δs^2 vanishes, so must $\Delta s'^2$):

$$\Delta s'^2 = F\,\Delta s^2. \tag{3.4}$$

The proportionality factor F can in principle depend on the coordinates and the relative velocity of these two frames: $F = F(\vec{x}, t, \vec{v})$. However the requirement that space and time be homogeneous (i.e. there is no privileged point in space and time) implies that there cannot be any dependence on $\vec{x}$ and t. That space is isotropic means that the proportionality factor cannot depend on the direction of the relative velocity $\vec{v}$ of the two frames. Thus, we can at most have it be dependent on the magnitude of the relative velocity, $F = F(v)$. We are now ready to show that, in fact, $F(v) = 1$.

Besides the system O', which is moving with velocity $\vec{v}$ with respect to system O, let us consider yet another inertial system O'', which is moving with a relative velocity of $-\vec{v}$ with respect to the O' system:

$$O \xrightarrow{\vec{v}} O' \xrightarrow{-\vec{v}} O''. \tag{3.5}$$

From the above consideration, and applying (3.4) to these frames:

$$\begin{aligned} \Delta s'^2 &= F(v)\,\Delta s^2, \\ \Delta s''^2 &= F(v)\;\Delta s'^2 = [F(v)]^2\,\Delta s^2. \end{aligned} \tag{3.6}$$

However, it is clear that the O'' system is in fact just the O system. This requires that $F(v)^2 = 1$. The solution $F(v) = -1$ being nonsensical, we conclude that this interval Δs is indeed an invariant: $\Delta s'' = \Delta s' = \Delta s$. Every inertial observer will see the same light velocity, and will therefore obtain the same value for this particular combination of space and time intervals.

3.1.1 Basis vectors, the metric and scalar product

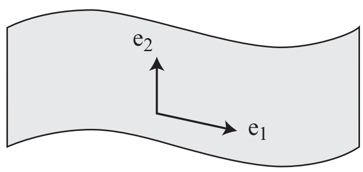

Fig. 3.1 Basis vectors for a 2D surface.

To set up a coordinate system for the 4D Minkowski space means to choose a set of four basis vectors $\{\mathbf{e}_\mu\}$, where $\mu = 0, 1, 2, 3$. Each $\mathbf{e}_\mu$, for a definite index value, is a 4D vector. (Figure 3.1 illustrates a case for a 2D space.) In contrast to the Cartesian coordinate system in Euclidean space (see Box 2.1),

978–0–19–957364–6 03-Cheng-c03 Cheng (Typeset by SPi, Chennai) 44 of 60 September 24, 2009 17:49

this in general is not an orthonormal set, $\mathbf{e}_\mu \cdot \mathbf{e}_\nu \neq \delta_{\mu\nu}$,[3] Nevertheless, we can represent such a collection of scalar products among the basis vectors as a symmetric matrix, called the **metric**, or the **metric tensor**:

$$\mathbf{e}_\mu \cdot \mathbf{e}_\nu \equiv g_{\mu\nu}. \tag{3.7}$$

We can display the metric as a 4×4 matrix with elements being dot products of basis vectors:

$$[\mathbf{g}] = \begin{pmatrix} g_{00} & g_{01} & .. \\ g_{10} & g_{11} & .. \end{pmatrix} = \begin{pmatrix} \mathbf{e}_0 \cdot \mathbf{e}_0 & \mathbf{e}_0 \cdot \mathbf{e}_1 & .. \\ \mathbf{e}_1 \cdot \mathbf{e}_0 & \mathbf{e}_1 \cdot \mathbf{e}_1 & .. \end{pmatrix}. \tag{3.8}$$

Thus, the diagonal elements are the (squared) lengths of the basis vectors, $|\mathbf{e}_0|^2$, $|\mathbf{e}_1|^2$, etc. while the off-diagonal elements represent their deviations from orthogonality. Any set of mutually perpendicular bases would be represented by a diagonal metric matrix, with diagonal entries of 1 if the basis vectors were of unit length (which is not required). For a Euclidean space with Cartesian coordinates, we have $g_{\mu\nu} = \delta_{\mu\nu}$.

We can expand any 4D vector[4] in terms of the basis vectors,

$$\mathbf{A} = \sum_\mu A^\mu \mathbf{e}_\mu \equiv A^\mu \mathbf{e}_\mu, \tag{3.9}$$

where the coefficients of expansion $\{A^\mu\}$ are labeled with superscript indices. From this point on, we shall adopt the **Einstein summation convention** of omitting the explicit display of the summation sign ($\sum$) whenever we have a pair of repeated indices[5] in one term of an expression, such as μ in (3.9). Consider the scalar product of two vectors and make their respective expansions, $\mathbf{A} \cdot \mathbf{B} = (A^\mu \mathbf{e}_\mu) \cdot (B^\nu \mathbf{e}_\nu) = (\mathbf{e}_\mu \cdot \mathbf{e}_\nu) A^\mu B^\nu$, which can be written as

$$\mathbf{A} \cdot \mathbf{B} = g_{\mu\nu} A^\mu B^\nu, \tag{3.10}$$

or displayed in matrix form as

$$\mathbf{A} \cdot \mathbf{B} = \begin{pmatrix} A^0 & A^1 & .. \end{pmatrix} \begin{pmatrix} g_{00} & g_{01} & .. \\ g_{10} & g_{11} & .. \end{pmatrix} \begin{pmatrix} B^0 \\ B^1 \end{pmatrix}. \tag{3.11}$$

The metric allows us to express the scalar product in terms of the vector components. A key feature is that all the indices are summed over (said to be "contracted") and the result has no free index left over. A scalar means that it is the same in all coordinates; it is unchanged under a transformation. In particular for the case $\mathbf{A} = \mathbf{B} = \mathbf{x}$, the position vector, then the invariant squared length $s^2 = \mathbf{x} \cdot \mathbf{x}$ can be written by (3.10) as

$$s^2 = g_{\mu\nu} x^\mu x^\nu. \tag{3.12}$$

We refer to this equation as the **metric equation**; it plays a central role in the geometric formulation of relativity. The coordinate-dependent metric turns the coordinate-dependent position components[6] $\{x^\mu\}$ into a coordinate-independent length s^2.

[3] $\delta_{\mu\nu}$ is the Kronecker delta:

$$\delta_{\mu\nu} = \begin{cases} 1; & \mu = \nu \\ 0, & \mu \neq \nu. \end{cases}$$

We can interpret its values as the elements of an identity matrix: $[\mathbf{1}]_{\mu\nu} = \delta_{\mu\nu}$.

[4] We denote 4D vectors with boldfaced letters, such as "**A**", and 3D vectors with an arrow on the top, such as "$\vec{A}$".

[5] Such repeated indices are called "dummy indices" and we are free to change their names, for example, $A^\mu \mathbf{e}_\mu = A^\nu \mathbf{e}_\nu = A^\lambda \mathbf{e}_\lambda$, etc. Also, note that although we apply this rule to any pair of repeated indices, strictly speaking it should always be a pair of repeated indices with one superscript index and the other subscript index. See Chapter 12 for further details.

[6] Here x^μ are understood to be the interval $\triangle x^\mu$ between the position point at x^μ and the origin of the coordinate system.

3.1.2 The Minkowski metric and Lorentz transformation

In Minkowski space we have the position 4-vector $x^\mu = (x^0,\ x^1,\ x^2,\ x^3) = (ct,\ x,\ \ y,\ z)$; the invariant length $s^2 = -c^2t^2 + x^2 + y^2 + z^2$ can be identified with the scalar product formula of (3.10) and (3.11):

$$s^2 = \mathbf{x}\cdot\mathbf{x} = \left(x^0\ x^1\ x^2\ x^3\right)\begin{pmatrix} -1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix}\begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \tag{3.13}$$

$$= \eta_{\mu\nu}x^\mu x^\nu. \tag{3.14}$$

Thus, the Minkowski space, with a pseudo-Cartesian coordinate system, has the metric

$$g_{\mu\nu} = \eta_{\mu\nu} = \mathrm{diag}\,(-1, 1, 1, 1)\,. \tag{3.15}$$

Because the metric $\eta_{\mu\nu}$ is constant (independent of position and time), Minkowski spacetime is a flat space,[7] as opposed to a curved space. It differs from the familiar Euclidean space by having a negative $\eta_{00} = -1$. We can think of x^0 as having an imaginary length. As we shall discuss in subsequent chapters, the spacetime manifold is warped in the presence of matter and energy. In Einstein's general theory of relativity, curved spacetime is the gravitational field and the metric $g_{\mu\nu}(x)$ for such a warped spacetime is necessarily position-dependent. The pseudo-Euclidean flat spacetime is obtained only in the **absence** of gravity. This is the limit of special relativity.

Lorentz transformations are the coordinate transformations between two frames moving with a constant velocity with respect to each other in Minkowski spacetime.[8] For example a boost with velocity v in the x^1 direction changes the vector components $x^\mu \to x'^\mu$ as shown in (2.11) of the previous chapter. We can write this transformation in matrix-component form as

$$\begin{pmatrix} x'^0 \\ x'^1 \\ x'^2 \\ x'^3 \end{pmatrix} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}, \tag{3.16}$$

where β and γ are defined in (2.12). We shall demonstrate in Box 3.2 that this explicit form of the Lorentz transformation follows directly from the requirement that the transformation leaves the length (in fact, any scalar product) and the metric of the space invariant. The Lorentz coordinate transformation of (3.16) may be written in matrix-component form as

$$x'^\mu = \mathbf{L}^\mu{}_\nu x^\nu. \tag{3.17}$$

L denotes the 4×4 Lorentz transformation matrix, with the left index (μ) being the row index and the right index (ν) the column index.[9] The components x^μ being those of a prototype 4-vector, any 4-vector **A** must have components A^μ that transform into components A'^μ of the same vector in another coordinate system, under a boost in the x^1 direction, in exactly the same

[7] In Chapter 5 we shall present a brief introduction to the geometric description of a curved space. We start with the familiar 2D space. Thus, say, a horizontal sheet of paper is "flat" and a spherical surface is "curved." This constant metric is a sufficient but not a necessary condition for a flat space. A flat plane can still have a position-dependent metric if we adopt a system such as the polar coordinate system.

[8] In our presentation, we shall restrict ourselves to coordinate transformations under which the coordinate origin $(t, x) = (0, 0)$ is fixed. These are Lorentz transformations. The combined transformations of a Lorentz and a coordinate translation ($x^\mu \to x^\mu + a^\mu$ with a^μ being a constant) is called a *Poincaré transformation*.

[9] We follow this convention regardless of whether the index is a superscript or subscript.

way as given in (3.16):

$$A'^{\mu} = \mathbf{L}^{\mu}_{\nu} A^{\nu}. \tag{3.18}$$

Box 3.2 Lorentz transformation as a rotation in 4D spacetime

The key feature of Minkowski spacetime is that it has the pseudo-Euclidean metric of (3.15). A (squared) length in this space is given by $s^2 = -c^2t^2 + x^2 + y^2 + z^2$. We can make this identification even more obvious by working with an imaginary coordinate $w = ict$ so that $s^2 = w^2 + x^2 + y^2 + z^2$. The coordinate transformation $\mathbf{L}$ can be thought of as a "rotation" of the 4D spacetime coordinates.[10] Any rotational transformation (by definition) preserves the length of vectors. This condition that a rotational transformation be length preserving is enough to fix the explicit form of the Lorentz transformation.[11] Consider the relation between two inertial frames connected by a boost (with velocity v) in the $+x$ direction. Since the (y, z) coordinates are not affected, we have effectively a two-dimensional problem. The rotation relations are just like (2.3):

$$\begin{aligned} w' &= \cos\phi\, w + \sin\phi\, x, \\ x' &= -\sin\phi\, w + \cos\phi\, x. \end{aligned} \tag{3.19}$$

Plugging in $w = ict$, we have

$$\begin{aligned} ct' &= \cos\phi\, ct - i\sin\phi\, x, \\ x' &= -i\sin\phi\, ct + \cos\phi\, x. \end{aligned} \tag{3.20}$$

Reparametrizing the rotation angle as $\phi = i\psi$ and using $-i\sin(i\psi) = \sinh\psi$ and $\cos(i\psi) = \cosh\psi$,[12] we get

$$\begin{aligned} ct' &= \cosh\psi\, ct + \sinh\psi\, x, \\ x' &= \sinh\psi\, ct + \cosh\psi\, x. \end{aligned} \tag{3.21}$$

Thus, in the (ct, x) space a Lorentz boost transformation has the matrix form of

$$[\mathbf{L}(\psi)] = \begin{pmatrix} \cosh\psi & \sinh\psi \\ \sinh\psi & \cosh\psi \end{pmatrix}. \tag{3.22}$$

To relate the parameter ψ, called the **rapidity** parameter,[13] to the boost velocity v, we concentrate on the coordinate origin $x' = 0$ of the O' system. Plugging $x' = 0$ into (3.21):

$$x' = 0 = ct\sinh\psi + x\cosh\psi \quad \text{or} \quad \frac{x}{ct} = -\frac{\sinh\psi}{\cosh\psi}. \tag{3.23}$$

The coordinate origin $x' = 0$ moves with velocity $v = x/t$ along the x axis of the O system:

$$v = \frac{x}{t} = -c\frac{\sinh\psi}{\cosh\psi} \quad \text{or} \quad \frac{\sinh\psi}{\cosh\psi} = -\frac{v}{c} = -\beta. \tag{3.24}$$

[10] For an introductory discussion of rotation, see Box 2.1.

[11] A more formal proof can be found in Chapter 12. The components of a vector A^{μ} change under a coordinate transformation as in (3.18) with its length being an invariant, $g'_{\mu\upsilon} A'^{\mu} A'^{\upsilon} = g_{\mu\upsilon} A^{\mu} A^{\upsilon}$. If $\mathbf{L}$ does not change the metric $\mathbf{g}' = \mathbf{g}$, the corresponding symmetry is called an **isometry**. Such a transformation must satisfy the "generalized orthogonality condition," $\mathbf{L}\mathbf{g}\mathbf{L}^{\mathsf{T}} = \mathbf{L}^{\mathsf{T}}\mathbf{g}\mathbf{L} = \mathbf{g}$. In Euclidean space with $\mathbf{g} = \mathbf{1}$, this reduces to the familiar statement (Problem 3.3) that the transformation matrix must be an orthogonal matrix (i.e. the transposed matrix is the inverse matrix: $\mathbf{L}^{\mathsf{T}} = \mathbf{L}^{-1}$). Further details will be provided in Chapter 12. In Problems 3.4 and Problem 12.7, the reader will be asked to show that the explicit form of $\mathbf{L}$ can be determined from such conditions.

[12] Recall the identities $\sinh\psi = (e^{\psi} - e^{-\psi})/2$ and $\sin\phi = (e^{i\phi} - e^{-i\phi})/2i$; also, $\cosh\psi = (e^{\psi} + e^{-\psi})/2$ and $\cos\phi = (e^{i\phi} + e^{-i\phi})/2$.

[13] The rapidity parameter ψ has the property that, while the addition of relative velocities is complicated as in (2.22), relative rapidity is additive because the hyperbolic tangent, which is the velocity, satisfies the trigonometry identity

$$\tanh(\psi_1 \pm \psi_2) = \frac{\tanh\psi_1 \pm \tanh\psi_2}{1 \pm \tanh\psi_1 \tanh\psi_2}.$$

From the identity $\cosh^2\psi - \sinh^2\psi = 1$, which may be written as $\cosh\psi\sqrt{1-(\sinh^2\psi/\cosh^2\psi)} = 1$, we find

$$\cosh\psi = \gamma \quad \text{and} \quad \sinh\psi = -\beta\cosh\psi = -\beta\gamma, \tag{3.25}$$

where $\beta = v/c$ and $\gamma = (1-\beta^2)^{-1/2}$. The coordinate transformation in matrix form (3.22) is found to be

$$\begin{pmatrix} ct' \\ x' \end{pmatrix} = \gamma \begin{pmatrix} 1 & -\beta \\ -\beta & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix}, \tag{3.26}$$

which is just the Lorentz transformation stated in (3.16).

3.2 Four-vectors for particle dynamics

As a further application of Minkowski 4-vectors, we consider some of the basic quantities involved in the description of particle dynamic: velocity, energy and momentum, as well as acceleration. We shall see that many interesting relativistic features can already be deduced by the proper construction of quantities (vectors, scalars, etc.) that have definite transformation properties under 4D rotation (Lorentz transformation) in Minkowski spacetime.

3.2.1 The velocity 4-vector

We have already shown in Chapter 2 [see Eqs. (2.19)–(2.22)] that the velocity components have rather complicated Lorentz transformation properties. This is because ordinary velocity dx^μ/dt is not a proper 4-vector; namely

$$\frac{dx'^\mu}{dt'} \neq \mathbf{L}^\mu_\nu \frac{dx^\nu}{dt} \quad \text{as} \quad t' \neq t. \tag{3.27}$$

While dx^μ is a 4-vector $\left(dx'^\mu = \mathbf{L}^\mu_\nu dx^\nu\right)$, the ordinary time coordinate t is not a Lorentz scalar—it is a component of a 4-vector: $x^\mu = \left(ct, x^1, x^2, x^3\right)$. Consequently, the quotient dx^μ/dt cannot be a 4-vector. This suggests that, in order to construct a velocity 4-vector, we should differentiate the displacement with respect to the proper time τ, which is a Lorentz scalar (recall that $s^2 = -c^2\tau^2$ is invariant under a Lorentz transformation):

$$U^\mu = \frac{dx^\mu}{d\tau}, \tag{3.28}$$

and we have, as in (3.18),

$$U'^\mu = \mathbf{L}^\mu_\nu U^\nu. \tag{3.29}$$

The relation between the 4-velocity U^μ and dx^μ/dt can be readily deduced, as coordinate time and proper time are related by the time dilation relation of

$t = \gamma\tau$, with

$$\gamma = \left(1 - \frac{v^2}{c^2}\right)^{-\frac{1}{2}} \quad \text{and} \quad v = \left|v^i\right| \quad \text{with} \quad v^i = \frac{dx^i}{dt}. \tag{3.30}$$

We have the components of 4-velocity related to the ordinary velocity components v^i with $(i = 1, 2, 3)$ as

$$U^\mu = \frac{dx^\mu}{d\tau} = \gamma\frac{dx^\mu}{dt} = \gamma\left(c,\, v^1,\, v^2,\, v^3\right). \tag{3.31}$$

As an instructive exercise (Problem 3.7), the reader is invited to deduce the velocity transformation rule (2.22) from the fact that U^μ is a 4-vector as in (3.29). It is easy to check that the "4-velocity length squared" $|\mathbf{U}|^2 \equiv \eta_{\mu\nu}U^\mu U^\nu$ is a Lorentz scalar, see Eq. (3.10), with the same value in every coordinate frame:

$$|\mathbf{U}|^2 = \gamma^2(-c^2 + v^2) = -c^2, \tag{3.32}$$

where we have used the definition of γ as given in (3.30). For any material particle $(v < c)$, we have $|\mathbf{U}|^2 = -c^2$. For photons (and any other particles with zero rest mass[14]), which can only travel at $v = c$, Eq. (3.32) would have a vanishing RHS:

[14] For massless particles in general, see the discussion below as well as that at the end of Section 3.2.2.

$$|\mathbf{U}|^2 = \gamma^2(-c^2 + c^2) = 0. \tag{3.33}$$

Thus U^μ is a null 4-vector. We note that because there is no rest frame for a photon, the concept of "the proper time of the photon" does not exist. In that case, one must replace τ by some curve parameter (say, λ) of the photon's trajectory, $x^\mu(\lambda)$; and the 4-velocity invariant becomes

$$|\mathbf{U}|^2 = \eta_{\mu\nu}U^\mu U^\nu = \eta_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}, \tag{3.34}$$

which vanishes because the invariant spacetime separation for light is $\eta_{\mu\nu}dx^\mu dx^\nu = 0$. That is, the 4-velocity for a light ray has zero length in spacetime (a null 4-vector).

3.2.2 Relativistic energy and momentum

For momentum, we naturally consider the product of the invariant mass m with the 4-velocity of (3.31):

$$p^\mu \equiv mU^\mu = \gamma\left(mc,\, mv^i\right), \tag{3.35}$$

with mv^i being the components of the nonrelativistic 3-momentum $\vec{p} = m\vec{v}$. The spatial components of the relativistic 4-momentum p^μ are the components of the relativistic 3-momentum, $p^i = \gamma mv^i$, which reduces to mv^i in the nonrelativistic limit of $\gamma = 1$. What then is the zeroth component of the 4-momentum? Let's take its nonrelativistic limit $(v \ll c)$:

$$p^0 = mc\gamma = mc\left(1 - \frac{v^2}{c^2}\right)^{-\frac{1}{2}}$$

$$\xrightarrow{\text{NR}} mc\left(1+\frac{1}{2}\frac{v^2}{c^2}+\cdots\right)=\frac{1}{c}\left(mc^2+\frac{1}{2}mv^2+\cdots\right). \quad (3.36)$$

The presence of the kinetic energy term $\frac{1}{2}mv^2$ in the nonrelativistic limit naturally suggests that we interpret cp^0 as the relativistic energy $E = cp^0 = \gamma mc^2$, which has a nonvanishing value mc^2 even when the particle is at rest

$$p^\mu = \left(E/c,\ p^i\right). \quad (3.37)$$

According to (3.32) and (3.35), the invariant square of the 4-momentum $\eta_{\mu\nu}p^\mu p^\nu$ must be $-(mc)^2$. Plugging in (3.37), we obtain the important relativistic energy–momentum relation:

$$E^2 = \left(mc^2\right)^2 + (\vec{p}c)^2 = m^2c^4 + \vec{p}^2c^2. \quad (3.38)$$

Once again, for a particle with mass,[15] we have the components of the relativistic 3-momentum and the relativistic energy as

$$p^i = \gamma mv^i \quad \text{and} \quad E = \gamma mc^2. \quad (3.39)$$

Thus the ratio of a particle's momentum to its energy can be expressed as that of velocity over c^2

$$\frac{p^i}{E} = \frac{v^i}{c^2}. \quad (3.40)$$

[15]The concept of a velocity-dependent mass $m^* \equiv \gamma m$ is sometimes used in the literature, so that $p^i = m^* v^i$ and $E = m^*c^2$. In our discussion we will avoid such usage and restrict ourselves only to the Lorentz scalar mass m, which equals m^* in the rest frame of the particle $\left(m^*|_{v=0} = m\right)$, hence, m is called the **rest mass**

Massless particles always travel at speed c

With $m = 0$, we can no longer define the 4-momentum as $p^\mu = mU^\mu$; nevertheless, since a massless particle has energy and momentum, we can still assign a 4-momentum to such a particle, with components just as in (3.37). When $m = 0$, the relation (3.38) with $p \equiv |\vec{p}|$ becomes

$$E = pc. \quad (3.41)$$

Plugging this into the ratio of (3.40), we obtain the well-known result that massless particles such as photons and gravitons[16] always travel at the speed of $v = c$. Another way of saying the same thing: with this energy momentum relation (3.40), the 4-momentum of a massless particle[17] must have components $p^\mu = \left(p,\ p^i\right)$ which, just like its 4-velocity, is manifestly a null 4-vector.

[16]Gravitons are the quanta of the gravitational field, just as photons are the quanta of the electromagnetic field.

[17]While we do not have $p^\mu = mU^\mu$, we still have the proportionality of the 4-momentum to its 4-velocity, $p^\mu \propto U^\mu$, with the 4-velocity defined as $U^\mu = dx^\mu/d\lambda$. In fact one can choose the curve parameter λ in such a way that $p^\mu = dx^\mu/d\lambda$.

Box 3.3 The wavevector

Here we discuss the wave 4-vector, which is closely related to the photon 4-momentum. Recall that for a dynamic quantity $A\,(\vec{x},\ t)$ to be a solution to the wave equation, its dependence on the space and time coordinates must be in the combination of $(\vec{x} - \vec{v}\,t)$, where $\vec{v}$ is the wave velocity. A harmonic electromagnetic wave is then proportional to $\exp i(\vec{k}\cdot\vec{x} - \omega t)$, with $k = |\vec{k}| = 2\pi/\lambda$ being the wavenumber, and $\omega = 2\pi/T$ being the

(*cont.*)

Box 3.3 (*Continued*)

angular frequency corresponding to a wave period of T, and their being related to the velocity of light by $\omega/k = c$. The phase factor $(\vec{k}\cdot\vec{x} - \omega t)$, which basically counts the number of peaks and troughs of the wave, must be a frame-independent quantity (i.e. a Lorentz scalar). To make this scalar nature explicit, we write this phase in terms of the 4-vector $x^\mu = (ct,\ \vec{x})$ as

$$\vec{k}\cdot\vec{x} - \omega t = \begin{pmatrix} ct, & \vec{x} \end{pmatrix} \begin{pmatrix} -1 & \\ & 1 \end{pmatrix} \begin{pmatrix} \omega/c \\ \vec{k} \end{pmatrix} \equiv \eta_{\mu\nu} x^\mu k^\nu.$$

From our knowledge that x^μ is a 4-vector and $\eta_{\mu\nu}x^\mu k^\nu$ a scalar, we conclude[18] that ω and $\vec{k}$ must also form a 4-vector, the "wavevector:" $k^\mu = \left(\omega/c,\ k^i\right)$. Under the Lorentz transformation, the components of the wavevector transform, according to (3.18), as

$$k^\mu \longrightarrow k'^\mu = [\mathbf{L}]^\mu_{\ \nu}\, k^\nu. \tag{3.42}$$

Specifically under a Lorentz boost in the $+x$ direction, we have from (3.16):

$$k'_x = \gamma\left(k_x - \beta\frac{\omega}{c}\right) \tag{3.43}$$

$$\omega' = \gamma\left(\omega - c\beta k_x\right) = \gamma\left(\omega - c\beta k\cos\theta\right) \tag{3.44}$$

where θ is the angle between the boost direction $\hat{x}$ and the direction of wave propagation $\hat{k}$. Since $ck = \omega$, we obtain in this way the **relativistic Doppler formula**,

$$\omega' = \frac{(1-\beta\cos\theta)}{\sqrt{1-\beta^2}}\omega \tag{3.45}$$

which is to be compared to the nonrelativistic Doppler relation $\omega' = (1-\beta\cos\theta)\,\omega$. We note that in the nonrelativistic limit there is no Doppler shift in the transverse direction, $\theta = \pi/2$, as compared to the relativistic "transverse Doppler effect" of $\omega' = \gamma\omega$. (One can trace the origin of this new effect back to SR time dilation.) In the longitudinal direction, $\theta = 0$, we have the familiar relation of

$$\frac{\omega'}{\omega} = \sqrt{\frac{1-\beta}{1+\beta}}, \tag{3.46}$$

which has the low-velocity (v) approximation of ($v \ll c$)

$$\frac{\omega'}{\omega} \approx (1-\beta) \quad \text{or} \quad \frac{\Delta\omega}{\omega} \approx -\frac{v}{c}. \tag{3.47}$$

Because of the $\omega = ck$ relation, the wave 4-vector $k^\mu = \left(k,\ k^i\right)$ has a null invariant length $\eta_{\mu\nu}k^\mu k^\nu = 0$, compatible[19] with the property of the 4-momentum for a photon as shown in (3.33).

[18] This follows from the quotient theorem, which is described in Problem 12.6.

[19] This is expected from the quantum mechanical relation of $p^\mu = \hbar k^\mu$, which has the zeroth component corresponding to $E = \hbar\omega$ and the spatial components, the de Broglie relation of $p = \hbar k = h/\lambda$.

Box 3.4 Covariant force

Just as the ordinary velocity $\vec{v}$ has a complicated Lorentz property—and so we introduced the object of 4-velocity—it is also not easy to relate different components of the usual force vector $F^i = dp^i/dt$ in different moving frames. However, the notions of 4-velocity and 4-momentum naturally lead us to the definition of 4-force, or the covariant force, as

$$K^\mu \equiv \frac{dp^\mu}{d\tau} = m\frac{dU^\mu}{d\tau}, \tag{3.48}$$

which, using (3.37), has components[20]

$$K^\mu = \frac{dp^\mu}{d\tau} = \gamma \frac{d}{dt}\left(E/c,\ p^i\right) = \gamma\left(\dot{E}/c,\ F^i\right). \tag{3.49}$$

Next we show that the rate of energy change $\dot{E}$ is given, just as in nonrelativistic physics, by the dot product $\vec{F}\cdot\vec{v}$. Because $|\mathbf{U}|^2$ is a constant, its derivative vanishes.

$$0 = m\frac{d}{d\tau}\left(\eta_{\mu\nu}U^\mu U^\nu\right) = 2\eta_{\mu\nu} m \frac{dU^\mu}{d\tau}U^\nu = 2\eta_{\mu\nu}K^\mu U^\nu, \tag{3.50}$$

where we have used (3.48) to reach the last equality. Substituting in the components of K^μ and U^ν from (3.49) and (3.31), we have

$$0 = \eta_{\mu\nu}K^\mu U^\nu = \gamma^2(-\dot{E} + \vec{F}\cdot\vec{v}), \tag{3.51}$$

thus $\dot{E} = \vec{F}\cdot\vec{v}$. With this expression for $\dot{E}$, we can display the components of the covariant force (3.49) as

$$K^\mu = \gamma\left(\vec{F}\cdot\vec{v}/c,\ F^i\right). \tag{3.52}$$

[20] One must keep in mind that the 3-force vector $F^i = dp^i/dt$ here is a differential of the relativistic momenum; hence it is related to the nonrelativistic force by $F^i = \gamma F^i_{\mathrm{NR}}$.

3.3 The spacetime diagram

Space and time coordinates are labels of physical processes taking place, one "event" following another, in the world. Any two given events may or may not be causally connected. Relativity brings about a profound change in this causal structure of space and time, which can be nicely visualized in terms of the **spacetime diagram**.[21] For a given event P at a particular point in space and particular instant of time, all the events that could in principle be reached by a particle starting from P one collectively labeled as the **future of P**, while all the events from which a particle can arrive at P form the **past of point P**. In order to appreciate the nontrivial causal structure brought about by the new relativistic conception of space and time, let us first recall the corresponding structure that one had assumed in pre-relativity physics. Here the notion of simultaneous events is of key importance. Those events that are neither the future nor the past of the event P form a 3D set of events **simultaneous with P**. This notion of simultaneous events allows one to discuss, in pre-relativity

[21] To have the same length dimension for all coordinates, the temporal axis is represented by $x^0 = ct$.

physics, all of space at a given instant of time, and as a corollary, allows one to study space and time separately. In relativistic physics, the events that fail to be causally connected to event P are much larger than a 3D space. As we shall see, all events outside the future and past lightcones are causally disconnected from the event P, which lies at the tip of the lightcones in the spacetime diagram.

3.3.1 Basic features and invariant regions

An event with coordinates (t, x, y, z) is represented by a **worldpoint** in the spacetime diagram. The history of events becomes a line of worldpoints, called a **worldline**. In Fig. 3.2, the 3D space is represented by a 1D x axis. In particular, a light signal $\Delta s^2 = 0$ passing through the origin is represented by a straight worldline at a 45° angle with respect to the axes: $\Delta x^2 - c^2\Delta t^2 = 0$, thus $c\Delta t = \pm\Delta x$. Any line with constant velocity $v = |\Delta x/\Delta t|$ would be a straightline passing through the origin. We can clearly see that those worldlines with $v < c$, corresponding to $\Delta s^2 < 0$, would make an angle greater than 45° with respect to the spatial axis (i.e. above the worldlines for a light ray). According to relativity, no worldline can have $v > c$. If there had been such a line, it would correspond to $\Delta s^2 > 0$, and would make an angle less than 45° (i.e. below the light worldline). Since $\Delta s^2 = \Delta x^2 + \Delta y^2 + \Delta z^2 - c^2\Delta t^2$ is invariant, it is meaningful to divide the spacetime diagram into regions, as in Fig. 3.3, corresponding to

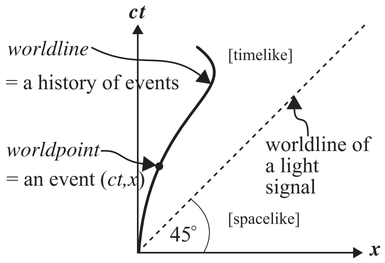

Fig. 3.2 Basic elements of a spacetime diagram, with two spatial coordinates (y, z) suppressed.

$\Delta s^2 < 0$	timelike
$\Delta s^2 = 0$	lightlike
$\Delta s^2 > 0$	spacelike

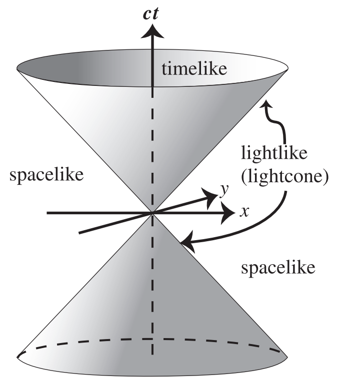

Fig. 3.3 Invariant regions in the spacetime diagram, with two of the spatial coordinates suppressed.

where the names of the region are listed on the right-hand column. The coordinate intervals being $c\Delta t = ct_2 - ct_1$, $\Delta x = x_2 - x_1$, etc. consider the separation of two events: one at the origin $(ct_1, \vec{x}_1) = (0, \vec{0})$, the other at a point in one of the regions $(ct_2, \vec{x}_2) = (ct, \vec{x})$:

- The light-like region has all the events which are connected to the origin with a separation of $\Delta s^2 = 0$. This corresponds to events that are connected by light signals. The 45° incline in Fig. 3.3, in which two spatial dimensions are displayed, forms a **lightcone**. It has a slope of unity, which reflects the fact that the speed of light is c. A vector that connects an event in this region to the origin, called **a light-like vector, is** a non-zero 4-vector having zero length, a **null vector**. The lightcone surface is a null 3-surface.
- The space-like region has all the events which are connected to the origin with a separation of $\Delta s^2 > 0$. (The 4-vector from the origin in this region is a **space-like vector**, having a positive squared length.) In the spacelike region, it takes a signal traveling at a speed greater than c in order to connect an event to the origin. Thus, an event taking place at any point in this region cannot be influenced causally (in the sense of cause-and-effect) by an event at the origin. We can alternatively explain it by going

978–0–19–957364–6 03-Cheng-c03 Cheng (Typeset by SPi, Chennai) 53 of 60 September 24, 2009 17:49

to another frame O' resulting in different spatial and time intervals $\Delta x' \neq \Delta x$ and $\Delta t' \neq \Delta t$. However the spacetime interval is unchanged, $\Delta s'^2 = \Delta s^2 > 0$. The form of (3.3), with the spatial terms being positive and the time term negative, suggests that we can always find an O' frame such that this event would be viewed as taking place at the same time $\Delta t' = 0$ as the event at the origin but at different locations $\Delta x' \neq 0$. This makes it clear that such a worldpoint (an event) **cannot be causally connected** to an event at the origin because the two events would have to be connected by an instantaneous signal, which is not possible, as no signal can travel faster than c. Thus the causally disconnected 3D space (represented by a horizontal plane) in pre-relativity physics is now enlarged to a much large region—all of the 4D subspace outside the lightcone.

- The time-like region has all the events which are connected to the origin with a separation of $\Delta s^2 < 0$. (The 4-vector from the origin in this region is a **time-like vector**, having a negative squared length.) One can always find a frame O' such that such an event takes place at the same locations, $x' = 0$, but at different time, $t' \neq 0$. This makes it clear that events in this region **can be causally connected** with the origin. In fact, all the worldlines passing through the origin will be confined to this region, inside the lightcone.[22] In Fig. 3.3, we have displayed the lightcone structure with respect to the origin of the spacetime coordinates ($t = 0$, $\vec{x} = 0$). It should be emphasized that each point in a spacetime diagram has a lightcone. The time-like regions with respect to several worldpoints are represented by the lightcones shown in Fig. 3.4. If we consider a series of lightcones having their vertices located along a given worldline, each subsequent segment must lie within the lightcone of that point (at the beginning of that segment). It is clear from Fig. 3.4 that any particle can only proceed in the direction of ever-increasing time. We cannot stop our biological clocks!

[22] The worldline of an inertial observer (i.e. moving with constant velocity) must be a straight line inside the lightcone. This straight line is just the time axis of the coordinate system in which the inertial observer is at rest.

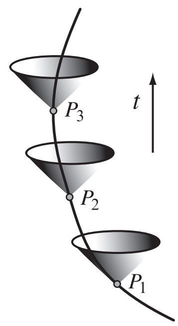

Fig. 3.4 Lightcones with respect to different worldpoints, P_1, P_2, ..., etc. along a time-like worldline, which can only proceed in the direction of ever-increasing time as each segment emanating from a given worldpoint must be contained within the lightcone with that point as its vertex.

3.3.2 Lorentz transformation in the spacetime diagram

The nontrivial parts of the Lorentz transformation (3.16) of intervals (taken, for example, with respect to the origin) are

$$\Delta x' = \gamma\left(\Delta x - \beta c \Delta t\right), \quad c\Delta t' = \gamma\left(c\Delta t - \beta \Delta x\right). \tag{3.53}$$

We can represent these transformed axes in the spacetime diagram:

- The x' axis corresponds to the $c\Delta t' = 0$ line. This corresponds, according the second equation above, to a line satisfying the relationship $c\Delta t = \beta \Delta x$. Hence, the x' axis is a straight line in the (x, ct) plane with a slope of $c\Delta t/\Delta x = \beta$.
- The ct' axis corresponds to the $\Delta x' = 0$ line. This corresponds, according the first equation above, to a line satisfying the relationship $\Delta x = \beta c \Delta t$. Hence, the ct' axis is a straight line with a slope of $c\Delta t/\Delta x = 1/\beta$.

Depending on whether β is positive or negative, the new axes either "close in" or "open up" from the original perpendicular axes. Thus we have the **opposite-angle rule**: the two axes make opposite-signed rotations of $\pm\theta$ (Fig. 3.5). The x axis rotates by $+\theta$ relative to the x' axis; the ct axis, by $-\theta$ relative to the ct' axis. The physical basis for this rule is the need to maintain the same slope ($= 1$; i.e. **equal angles** with respect to the two axes) for the lightcone in every inertial frame so that light speed is the same in every frame. Another important feature of the diagrammatic representation of the Lorentz transformation is that the new axes will have a scale **different** from the original one. Namely, the unit-lengths along the axes of the two systems are different. Let us illustrate this by an example. Consider the separation (from the origin O) of an event A on the ct' axis, which has O' system coordinates $\left(ct' = 1,\ x' = 0\right)$, see Fig. 3.6. What O system coordinates $(ct,\ x)$ does the worldpoint have?

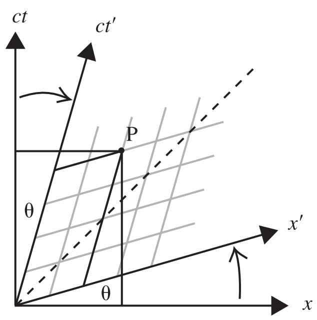

Fig. 3.5 Lorentz rotation in the spacetime diagram. The space and time axes rotate by the same amount but in opposite directions so that the lightcone (the dashed line) remains unchanged. The shaded grid represents lines of fixed x' and t'.

$$x' = \gamma\,(x - \beta ct) = 0 \Rightarrow \quad x = \beta ct,$$
$$ct' = \gamma\,(ct - \beta x) = ct\gamma\left(1 - \beta^2\right) = ct/\gamma = 1.$$

Hence this event has $(ct = \gamma, x = \gamma\beta)$ coordinates in the O system. Evidently, as $\gamma > 1$, a unit vector along the ct' direction has "projection" on the ct axis that is longer than unit length. This is possible only if there is a scale change when transforming from on reference system to another.

Consider another separation of an event B on the x axis, which has O coordinates $(ct = 0, x = 1)$. It is straightforward to check that it has O' system coordinates $\left(ct' = -\gamma\beta,\ x' = \gamma\right)$, again showing a difference in scales of the two systems.

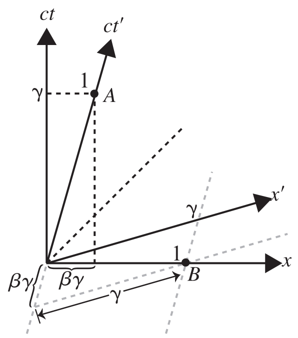

Fig. 3.6 Scale change in a Lorentz rotation. A unit length on the ct' axis has a *longer* projection, γ, onto the ct axis. The event A $\left(ct' = 1, x' = 0\right)$ in the O' frame is seen by an observer in the O frame to have coordinates $(ct = \gamma, x = \beta\gamma)$. Similarly, the event B, with coordinates $(ct = 0, x = 1)$ in the O frame, has coordinates $\left(ct' = \gamma\beta, x' = \gamma\right)$ in the O' frame. The two sets of dotted lines passing through worldpoints A and B are parallel lines to the axes of (ct, x) and $\left(ct', x'\right)$, respectively.

Box 3.5 Time dilation and length contraction in the spacetime diagram

The physics behind the scale changes discussed above is **time dilation** and **length contraction**. While the algebra involved in deriving these results from the Lorentz transformation (3.53) is simple, in order to obtain the correct result, one has to be very clear as to exactly what is being measured (spatial length or time), and which frame is being chosen as the rest frame for the appropriate object (in the derivation of time dilation, the object is a clock; in the derivation of length contraction, the object is any object whose spatial length is being measured). Knowing which frame has been chosen as the rest frame for the appropriate object, one must then have a mathematical way to express the fact that the clock is not moving or the object is not moving; one must have a mathematical condition or input to apply to the Lorentz transformation equations. The conventional way of doing this is to set $\Delta x' = 0$ for time dilation, and $\Delta t = 0$ for length contraction. For time dilation, we pick the O' frame to be the rest frame of the clock; since the clock is at rest in this frame, $\Delta x' = 0$. This is the crucial input into the Lorentz transformation equations in the derivation of time dilation. For length contraction, we also consider the O' frame to be the rest frame, but now it is the rest frame for an object whose length is being measured. The crucial input for this derivation is $\Delta t = 0$, reflecting the fact that the observer in the O frame must measure the ends of the object simultaneously.

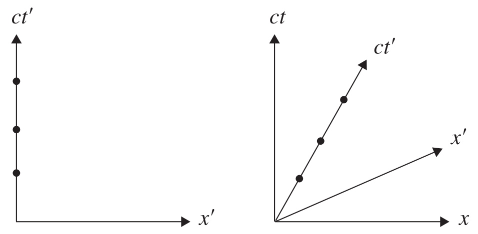

Fig. 3.7 Worldline of a clock, ticking at equal intervals: viewed in the rest frame of the clock, the O' system, and viewed in the moving frame, the O coordinate system.

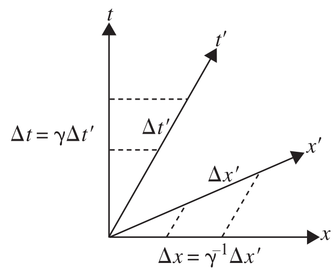

Fig. 3.8 Scale changes associated with the Lorentz rotation, reflecting the physics phenomena of time dilation and length contraction. The clock and object are moving with respect to the O system, but are at rest with respect to the O' system.

Time dilation A clock, ticking away in its own rest frame O' (also called the comoving frame), is represented by a series of worldpoints (the ticks of the clock) equally spaced on a vertical worldline ($\Delta x' = 0$) in the $\left(ct', x'\right)$ spacetime diagram. These same worldpoints when viewed in another inertial frame O, in which the O' system moves with $+v$ along the x axis, will appear as lying on an inclined worldline (Fig. 3.7). From the Lorentz transformation (2.40), as well as our previous discussion of the scale change under Lorentz rotation (also see Fig. 3.8), it is clear that the relationship between the time intervals in the two frames is

$$\Delta t = \gamma \Delta t' \quad \gamma > 1. \tag{3.54}$$

Thus we say that a moving clock (i.e. moving with respect to the O system) appears (Δt) to run slow. NB: keep in mind $\Delta x' = 0$; that is, there is no spatial displacement in the clock's rest frame, the comoving frame. However, there is a spatial displacement in the other, moving frame: $\Delta x \neq 0$.

Length contraction To obtain a length $\Delta x = x_1 - x_2$ in the O system of an object at rest in the O' system (and, therefore, moving with respect to the O system), we need to measure two events (t_1, x_1) and (t_2, x_2) simultaneously $\Delta t = t_1 - t_2 = 0$. (If you want to measure the length of a moving car, you certainly would not want to measure its front and back locations at different times!) The same two events,[23] when viewed in the rest frame of the object $\Delta x' = x_1' - x_2'$, will be measured according to (3.53) to have a greater separation (cf. Fig. 3.8):

$$\Delta x' = \gamma \Delta x > \Delta x. \tag{3.55}$$

[23]While we have simultaneous measurements in the moving frame, $\Delta t = 0$, these two events would be viewed as taking place at different times in the rest frame, $\Delta t' = \gamma\left(\Delta t - \frac{v}{c^2}\Delta x\right) \neq 0$. Of course, in the rest frame of the object, there is no need to perform the measurements simultaneously: in order to measure the front and back ends of a parked car, it is perfectly all right to make one measurement, take a lunch break, and then come back to measure the other end.

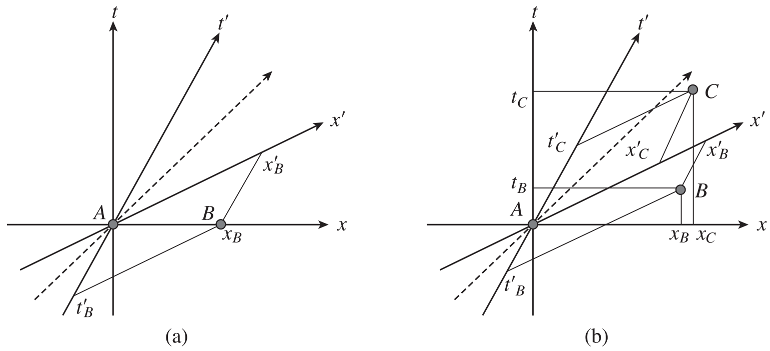

Fig. 3.9 (a) Relativity of simultaneity: $t_A = t_B$ but $t'_A > t'_B$. (b) Relativity of event order: $t_A < t_B$ but $t'_A > t'_B$. However, there is no change of event order with respect to A for all events located above the x' axis, such as event C. This certainly includes the situation in which C is located in the forward lightcone of A (above the dashed line).

Relativity of simultaneity, event-order and causality

It is instructive to use the spacetime diagram to demonstrate some of the physical phenomena we have discussed previously. In Fig. 3.9, we have two events A and B, with A being the origin of the coordinate system O and O': $\left(x_A = t_A = 0, x'_A = t'_A = 0\right)$. In Fig. 3.9(a), the events A and B are simultaneous, $t_A = t_B$, with respect to the O system. But in the O' system, we clearly have $t'_A > t'_B$. This shows the relativity of simultaneity. In Fig. 3.9(b), we have $t_A < t_B$ in the O frame, but we have $t'_A > t'_B$ in the O' frame. Thus, the temporal order of events can be changed by a change of reference frames. However, this change of event order can take place only if event B is located in the region below the x' axis.[24] This means that if we increase the relative speed between these two frames O and O' (with the x' axis moving ever closer to the lightcone) more and more events can have their temporal order (with respect to A at the origin) reversed as seen from the perspective of the moving observer. On the other hand, for all events above the x'axis, the temporal order is preserved. For example, with respect to event C, we have both $t_A < t_C$ and $t'_A < t'_C$. Now, of course, the region above this x' axis includes the forward lightcone of event A. This means that for two events that are causally connected (between A and any worldpoint in its forward lightcone), their temporal order cannot be changed by a Lorentz transformation. The principle of causality is safe under special relativity.

[24]The x' axis having a $1/\beta$ slope means that the region below it corresponds to $(\Delta x/\Delta t) > c/\beta$. This is clearly in agreement with the Lorentz transformation of $\Delta t' = \gamma\,(\Delta t - \beta\Delta x/c)$ to have opposite sign to Δt.

3.4 The geometric formulation of SR: A summary

Let us summarize the principal lessons we have learnt from this geometric formulation of special relativity:

- The stage on which physics takes place is Minkowski spacetime with the time coordinate being on an equal footing with spatial ones. “Space and time are treated symmetrically.” A spacetime diagram is often useful in clarifying ideas in relativity, especially its causal structure.
- Minkowski spacetime has a pseudo-Euclidean length (squared) of $\Delta s^2 = -c^2\Delta t^2 + \Delta x^2 + \Delta y^2 + \Delta z^2$. This relation between length and coordinate $\left[\Delta s^2 = g_{\mu\nu}\Delta x^\mu \Delta x^\nu\right]$ can be stated by saying that

Minkowski space has a flat-space metric of $g_{\mu\nu} = \text{diag}\,(-1, 1, 1, 1) \equiv \eta_{\mu\nu}$.

- Δs is invariant under transformations among inertial frames of reference. Such length-preserving transformations in a space with the metric $\eta_{\mu\nu}$ are just the Lorentz transformations, from which we can derive all the physical consequences of time dilation, length contraction, relativity of simultaneity, etc. Thus, in this geometric formulation, we can think of the metric as embodying all of special relativity.
- That one can understand special relativity as a theory of flat geometry in Minkowski spacetime is the crucial step in the progression towards general relativity. In GR, as we shall see, this geometric formulation is generalized into a warped spacetime. The corresponding metric must be position-dependent, $g_{\mu\nu}(x)$, and this metric acts as the generalized gravitational potential.
- In our historical introduction, SR seems to be all about light; the speed c actually plays a much broader role in relativity:
  - c is the universal maximal and absolute speed of signal transmission: massless particles (e.g. photons and gravitons) travel at the speed c, while all other ($m \neq 0$) particles move at a slower speed.
  - c is the universal **conversion factor** between space and time coordinates[25] that allows space and time to be treated symmetrically (i.e. on an equal footing) in relativity.
  - c is absolute and is just the speed so that $\Delta s^2 = -c^2\Delta t^2 + \Delta\vec{x}^2$ is an invariant interval under coordinate transformations. This allows Δs to be viewed as the length in spacetime. Thus, constancy of c underlies the entire geometric formulation of relativity.

[25] Since space and time have different units, we must have a conversion factor connecting the space and time coordinates. We mention two other fundamental physics conversion factors: Newton's constant G_N and Planck's constant h (see Section 6.3.2).

Review questions

1. (a) Give the definition of the metric tensor in terms of the basis vectors.
   (b) What is the invariant interval (the "length") Δs^2 between two neighboring events with coordinate separation $(c\Delta t, \Delta x, \Delta y, \Delta z)$ in Minkowski spacetime?
   (c) When the metric is displayed as a square matrix, what is the interpretation of its respective diagonal and off-diagonal elements?
   (d) What is the metric for an n-dimensional Euclidean space, in Cartesian coordinates, with a (squared) length given by $\Delta s^2 = \Delta x_1^2 + \Delta x_2^2 + \cdots + \Delta x_n^2$?
2. (a) What is the essential input needed for the proof that $\Delta s^2 = -c^2\Delta t^2 + \Delta\vec{x}^2$ has the same value in every inertial frame of reference?
   (b) What is the physical meaning of s that every observer can agree on its value?
3. What are the components of the position 4-vector in Minkowski spacetime?
4. If A^μ is a 4-vector (e.g. it is the 4-velocity or 4-momentum, etc.), how do these components transform under a coordinate transformation $A^\mu \to A'^\mu$ if the position 4-vector changes as $x^\mu \to x'^\mu = [\mathbf{L}]^\mu_\nu\, x^\nu$?
5. Under a coordinate change $O \to O'$, how is $\eta_{\mu\nu}A^\mu B^\nu$ related to $\eta_{\mu\nu}A'^\mu B'^\nu$?
6. From the condition $\Delta s'^2 = \Delta s^2$, derive the explicit form of the Lorentz transformation for a boost $\vec{v} = +v\hat{x}$.

978–0–19–957364–6 03-Cheng-c03 Cheng (Typeset by SPi, Chennai) 58 of 60 September 24, 2009 17:49

7. Why is dx^μ/dt not a 4-vector? How is it related to the velocity 4-vector U^μ? The squared length of the 4-velocity $\eta_{\mu\nu}U^\mu U^\nu$ should be a Lorentz invariant; what is this invariant?
8. What are the definitions (in terms of particle mass and velocity) of relativistic energy E and momentum $\vec{p}$ of a particle? Display their non-relativistic limits. What components of momentum 4-vector p^μ are E and $\vec{p}$? How are they related to each other?
9. In the spacetime diagram, display the time-like, space-like, and light-like regions. Also, draw in a worldline for some inertial observer.
10. The coordinate frame O' is moving at a constant velocity v in the $+x$ direction with respect to the coordinate frame O. Display the transformed axes (x', ct') in a two-dimensional spacetime diagram with axes (x, ct). You are not asked to solve the Lorentz transformation equations; but only to justify the directions of the new axes.
11. Two events A and B are simultaneous ($t_A = t_B$), but not equilocal ($x_A \neq x_B$), in coordinate frame O. Use a spacetime diagram to show that these same two events are viewed as taking place at different times, $t'_A \neq t'_B$, by an observer in the O' frame (in motion with respect to the O frame).
12. In a spacetime diagram, display two events with a temporal order of $t_A > t_B$ in the O frame such that they can possibly appear to have a reversed order $t'_A < t'_B$ in the O' frame. What is the condition that Δx, Δt and v (relative speed between the O and O' frames) must satisfy in order to have this reversal of temporal order? Explain why event A cannot possibly be caused by event B.
13. **Length contraction** means that the measured length interval of $\Delta x = x_1 - x_2$ is less than the corresponding rest-frame length $\Delta x' = x'_1 - x'_2$. What is the condition on the time coordinates of these two events, (t_1, t_2) and (t'_1, t'_2), and why is this a necessary condition? **Time dilation** means that the measured time interval of $\Delta t = t_1 - t_2$ is longer than the corresponding rest-frame interval $\Delta t' = t'_1 - t'_2$. What is the condition on the spatial coordinates, (x_1, x_2) and (x'_1, x'_2), of these two events? Why is this a necessary condition? Use these conditions and the Lorentz transformation to derive the result of length contraction and time dilation, respectively.

Problems

3.1 **Inverse basis** We can introduce a set of "inverse basis vectors," denoted by an upper index, $\{\mathbf{e}^\mu\}$ which when multiplied with the basis vectors yield the identity: $\mathbf{e}^\mu \cdot \mathbf{e}_\nu = \delta^\mu_\nu$ and the products among the inverse bases form the "inverse metric" $g^{\mu\nu} = \mathbf{e}^\mu \cdot \mathbf{e}^\nu$.

(a) Show that one can use the inverse bases to project out the components A^μ from the vector $\mathbf{A}$.

(b) Prove that $g^{\mu\nu}$ is the inverse to the metric $g_{\mu\nu}$.

3.2 **Contraction and dummy indices** A scalar product $\mathbf{A} \cdot \mathbf{B} = g_{\mu\nu}A^\mu B^\nu$ can be written as a "contraction" of two vectors with upper and lower indices $\mathbf{A} \cdot \mathbf{B} = A^\mu B_\mu$ where $B_\mu = g_{\mu\nu}B^\nu$. We have used the Einstein summation convention and summed over repeated indices. As a practice of "contraction over repeated (dummy) indices," prove that the scalar product of symmetric and antisymmetric tensors will always vanish: $T_{\mu\nu}S^{\mu\nu} = 0$ where $T_{\mu\nu} = -T_{\nu\mu}$ and $S^{\mu\nu} = +S^{\nu\mu}$.

3.3 **Rotation matrix is orthogonal** Explicitly demonstrate that the rotation matrix in Eq. (2.4) satisfies the relation

$$\left[\mathbf{R}^{-1}(\theta)\right] = [\mathbf{R}(-\theta)] = \left[\mathbf{R}^{\mathsf{T}}(\theta)\right]$$

hence the orthogonality condition $\mathbf{R}\mathbf{R}^{\mathsf{T}} = \mathbf{1}$.

3.4 **Orthogonality fixes the rotation matrix** In Problem 3.3, you have been asked to show from the explicit form of a rotational matrix that it is an orthogonal matrix. Here you are asked to prove the converse: the orthogonality condition can fix the rotation matrix explicitly. Consider a rotation around the z axis:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix},$$

where the effective 2×2 rotation matrix $[\mathbf{R}]$ with real elements (a, b, c, d) must satisfy the orthogonal condition $[\mathbf{R}][\mathbf{R}]^{\mathsf{T}} = [\mathbf{1}]$ so that the length $x^2 + y^2$ is an invariant. Show that this condition fixes the explicit form of the rotation matrix to be

$$[\mathbf{R}] = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}.$$

3.5 **Group property of Lorentz transformations** Use simple trigonometry to show that the rotation and boost operators given in (2.4) and (3.22) satisfy the group property:

$$\mathbf{R}(\theta_1)\mathbf{R}(\theta_2) = \mathbf{R}(\theta_1 + \theta_2)$$
$$[\mathbf{L}(\psi_1)][\mathbf{L}(\psi_2)] = [\mathbf{L}(\psi_1 + \psi_2)]. \quad (3.57)$$

The expression of "group" used above is in the sense of a mathematical group, which is composed of elements satisfying multiplication rules such as the ones discussed here.

3.6 **Group multiplication leads to velocity addition rule** Use (3.57) of Problem 3.5 to prove the velocity addition rule (2.22).

3.7 **Lorentz transform and velocity addition rule** The velocity 4-vector U^μ being a 4-vector has the Lorentz boost transformation as in Eq. (3.16). From this, derive the velocity addition rule (2.22).

3.8 **Antiproton production threshold** Because of baryon number conservation, the simplest reaction to produce an antiproton $\bar{p}$ in proton–proton scattering is $pp \to ppp\bar{p}$. Knowing that the rest energy of a proton is $m_p c^2 = 0.94\,\text{GeV}$, use the invariant $\eta_{\mu\nu}p^\mu p^\nu$ to find the minimum kinetic energy a (projectile) proton must have in order to produce an antiproton after colliding with another (target) proton at rest.

3.9 **A more conventional derivation of the Doppler effect** A light signal of frequency ω sent from (x, t) is received at (x', t') with frequency ω'. The receiver is moving in the $+x$ direction with velocity v. One can derive the (longitudinal) Doppler formula (3.46) by the observation that the phase of a light wave (essentially the counting of peaks) remains the same for the sender and receiver: $d\phi = \omega d\tau = \omega' d\tau'$ (where τ and τ' are the proper times of the sender and receiver, respectively). From this, using the time dilation formula, one can relate the ratio ω'/ω to the coordinate time ratio dt'/dt, and finally to the relative velocity $\beta = v/c$.

3.10 **Twin paradox measurements and the Doppler effect** In presenting the twin paradox (Box 2.5) we used the traveling Al's observation of Bill's annual (birthday) fireworks to determine their respective ages. In the outward bound part of the journey ($\beta = 4/5$) Al sees the firework every three years, and in the inward bound journey, every four months. Discuss these time intervals from the viewpoint of the Doppler effect and show that the results are compatible with formula (3.46).

3.11 **Spacetime diagram for the twin paradox** Provide a spacetime diagram corresponding to the twin paradox discussed in Box 2.5. Let Bill's rest frame O having coordinates (x, t), the outbound-Al's rest frame O' with (x', t'), and inbound-Al's frame O'' with (x'', t''). Draw your diagram so that the perpendicular lines represent the (x, t) axes.

(a) Mark the event when Al departs in the spaceship by the worldpoint O; the event when Al returns and is reunited with Bill by the worldpoint Q; and the event corresponding to the event when Al turns around (from outward bound to inward bound) by the point P. Thus, the stay-at-home Bill has worldline OQ, and Al's worldline has two segments: OP for the outward bound and PQ for the inward bound parts of his journey.

(b) On the t axis, which should coincide with the worldline OQ, also mark the points M, P' and P'' which should be simultaneous with the turning point P as viewed in the coordinate frames of O, O' and O'', respectively.

(c) Indicate the time values of t_M, $t_{P'}$ and $t_{P''}$ (i.e. the elapsed times since Al's departure at point O in the (x, t) coordinate system. In particular, show how changing the inertial frame from O' and O'' brings about a time change of 32 years in the O frame.

3.12 **The twin paradox—the missing 32 years** In Problem 3.11, the event P' on Bill's worldline is viewed by the outward-bound Al to be simultaneous to the turning-point P just *before* Al turns around, and the event P'' simultaneous to P just *after*. They are viewed to have different time by the stay-at-home Bill, $t_{P'} \neq t_{P''}$. This just emphasizes again the point that time is just another coordinate label. When we change the frame of reference, all coordinates make their corresponding changes. Two different points P' and P'' in Bill's rest frame (the O system) are simultaneous to P when viewed from two different inertial frames, the O' and O'' systems, respectively. Thus a difference $t_{P'} \neq t_{P''}$ is brought about simply by a change of coordinate: $O' \longrightarrow O''$. Our discussion just below (2.41) and (2.45) suggests that $t_{P''} - t_{P'} = 32$ years. We have already calculated $t_{P'} = 9$ years by the relative motion between the O and O' frames. Let us verify the expected result of $t_{P''} = 41$ years in two ways.

(a) Calculate $t_{P''} = t_Q - t_{QP''}$, where $t_{QP''}$ is the time interval measured in Bill's rest frame of the second 15 year of Al's journey (for the worldline PQ).

(b) The turning point P is seen in the outward-bound O' frame to have the time $t'_{OP} = 15$ year; from the perspective of the inward-bound Al (O'' frame) how long an interval t''_{OP} does this first half of the journey appear to be? $t_{P''}$ can then be obtained by noting that $t_{P''}$ is the O frame measurement of this t''_{OP} time interval.

3.13 **Spacetime diagram for the pole-and-barn paradox** Draw a spacetime diagram for the pole-and-barn paradox as discussed in Box 2.4. Let (x, t) be the coordinates for the ground (barn) observer, with (x', t') the rest frame of the runner (pole). Show the worldlines for the front door (F), rear door (R) of the barn, and front end (A), back end (B) of the pole. Your diagram should display the order reversal phenomenon discussed in Box 2.4: $t_{AR} > t_{BF}$ and $t'_{AR} < t'_{BF}$.