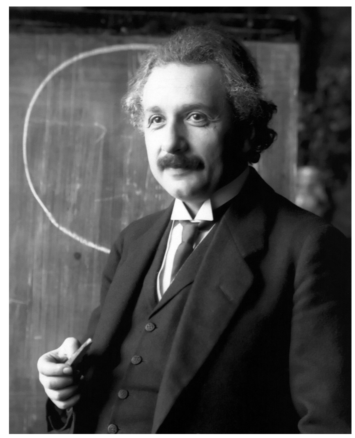
BEYOND EQUATIONS

Albert Einstein (1879 - 1955)



Einstein was born and raised in southern Germany. The family moved to Italy when he was a teenager, and he attended school there and in Switzerland. He did his advanced studies in physics at the Zurich Polytechnic, after which he was able to land a job at the Swiss patent office in Berne.

While working as a patent office clerk, in 1905 Einstein published five seminal papers: Two of them presented special relativity, one showed that light consists of particle-like "quanta", and two were on the existence and size of molecules. The revolutionary ideas presented in this "annus mirabilis" were slow to be generally accepted, but his paper on light quanta led to the Nobel Prize in 1921.

Finally becoming a professor, Einstein worked in Zurich and Prague, and later in Berlin, where in 1915 he published the general theory of relativity. When his prediction that light bends around the Sun was observationally confirmed in 1919, he because famous worldwide. Throughout all this time he displayed a penetrating physical intuition, seeming to see directly into the heart of nature.

Leaving Germany for good in 1933, Einstein became a founding professor at the Institute for Advanced Study in Princeton, New Jersey, where he worked until the end of his life. He spent his time working on a "unified field theory" and in trying to show that the theory of quantum mechanics is incomplete. He was unsuccessful at both of these enterprises; it was the enormous success of quantum mechanics, which (ironically) he had helped to invent in his early work on light quanta but could never fully accept, that made much of his later work fruitless. Nevertheless, he is universally recognized as the greatest physicist of the twentieth century.

Chapter 2

Relativity

In this chapter, we extend our review of mechanics to include Einstein's **special theory of relativity.** We will see that our previous Newtonian framework is a useful description of the mechanical world only when speeds are much less than that of light. We also use this chapter to introduce index notation and general technical tools that will help us throughout the rest of the book. Then in the following chapter we will show how relativity provides insights for an entirely different formulation of mechanics — the so-called **variational principle**.

2.1 Foundations

2.1.1 The Postulates

The most beautiful concepts in physics are often the simplest ones as well. In fact, the beautiful, revolutionary insights of special relativity are based on just two simple postulates:

1. **The principle of relativity**: *The fundamental laws of physics are valid in all inertial frames of reference.*

   We already introduced this principle in Chapter 1: it applies equally well to both newtonian and relativistic physics. There is also a second rather frugal, clean postulate inspired by electromagnetism: combined with the first, it leads to astounding conclusions that stretch one's imagination and intuition to the limit.

2. **The universal speed of light**: *The speed of light is the same in all inertial frames.*

The second postulate follows from the assertion that Maxwell's equations of electromagnetism are valid in all inertial frames. We know that Maxwell's equations lead to the wave equation

$$\frac{\partial^2\phi}{\partial x^2}+\frac{\partial^2\phi}{\partial y^2}+\frac{\partial^2\phi}{\partial z^2}-\frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2}=0 \tag{2.1}$$

for the electric potential $\phi(t,x,y,z)$ in vacuum, where c is the speed of light; this is in fact the wave equation for *light*. A convenient solution to this equation is given by the plane wave

$$\phi=\phi_0\cos(k\,(x-c\,t)) \tag{2.2}$$

where ϕ_0 is the amplitude and k is the wave number. Since (2.1) follows directly from a fundamental law of physics, c is a fundamental physical scale in Nature. Let us track the position of a particular wavefront in this plane wave. Take one of the crests $\phi=\phi_0$ with $x-c\,t=0$; the wavefront then evolves according to $x(t)=c\,t$ from the point of view of an observer at rest in frame $\mathcal{O}$. Now consider a different inertial frame $\mathcal{O}'$, moving in the positive x direction according to observers in $\mathcal{O}$, as shown in Figure 2.1, the exact same reference frames we introduced in Chapter 1. According to the second postulate, the same wavefront would be seen by $\mathcal{O}'$ as moving with the *same* speed c along x': $x'(t')=c\,t'$.

Panic ensues when we put these statements together with the Galilean transformation $x=x'+Vt$ of equation (1.1); this gives

$$c\,t=c\,t'+Vt=c\,t+Vt \tag{2.3}$$

since $t=t'$ in (1.1). This can be true only if the relative frame velocity V is zero!

To focus on the problem at hand, let us rephrase things in a slightly more general context. Say observer $\mathcal{O}$ is tracking a particle along a general trajectory $x(t)$. The same particle is seen by $\mathcal{O}'$ to evolve along $x'(t')$. A Galilean transformation tells us that $x(t)=x'(t')+Vt'$. Taking the time derivative of both sides of this equation, we get the usual velocity addition

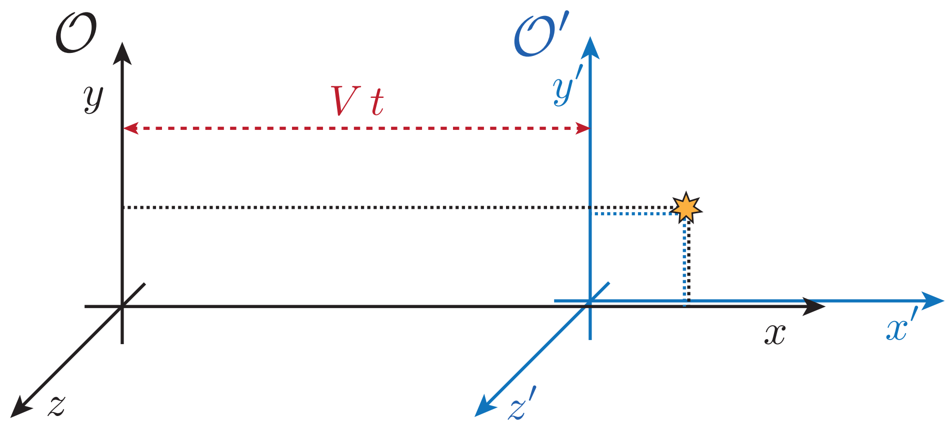


FIG 2.1 : Inertial frames $\mathcal{O}$ and $\mathcal{O}'$

rule (1.3)

$$\frac{d}{dt} = \frac{d}{dt'} : [x(t) = x'(t') + Vt'] \quad \Rightarrow \quad \frac{dx(t)}{dt} = \frac{dx'(t')}{dt'} + V$$
$$\Rightarrow \quad v_x = v'_x + V \ , \tag{2.4}$$

where $v_x = dx/dt$, $v'_x = dx'/dt'$, and we used $t = t'$ from equation (1.1). So if $v'_x = c$, then $v_x = c + V \neq c$ for $V \neq 0$, which contradicts the postulate, and we have a problem: *Galilean transformations are incompatible with the universal speed of light.* The second postulate can then be seen as a condition on the transformation rules relating the coordinate systems of inertial observers.

The overpriced million dollar question is then: what are the *correct* transformation equations relating the coordinates of $\mathcal{O}'$ and $\mathcal{O}$ to replace the Galilean transformation? Since the Galilean transformation arises intuitively from our basic sense of the world around us, it better be the case that it can be viewed at worse as a decent approximation to the correct transformation, which we now set out to find.

2.1.2 The Lorentz transformation

As shown aready in Figure 2.1, frames $\mathcal{O}$ and $\mathcal{O}'$ are assigned coordinate labels (t, x, y, z) and (t', x', y', z') respectively, such that $\mathcal{O}'$ moves with velocity $\boldsymbol{V}$ in the positive x direction as seen by observers at rest in frame $\mathcal{O}$, with the x and x' axes aligned and the y' axis parallel to the y axis, and the z' axis parallel to the z axis. According to the **Galilean** transformation, the coordinates in $\mathcal{O}$ are related to those in $\mathcal{O}'$ by (1.1)

$$x = x' + Vt', \quad y = y', \quad z = z', \quad t = t', \tag{2.5}$$

while the coordinates in $\mathcal{O}'$ are related to those in $\mathcal{O}$ by the inverse transformation

$$x' = x - Vt, \quad y' = y, \quad z' = z\ , \quad t' = t \tag{2.6}$$

which can be obtained from the first set simply by interchanging primed and unprimed coordinates and letting $V \to -V$.

Now the task at hand is to derive a replacement for the Galilean transformation, one that is consistent with the universal speed of light. We begin by assuming that the new transformation for x, y, z and inverse transformation for x', y', z' have the somewhat more general, but still linear, form

$$x = \gamma x' + \zeta t', \quad x' = \gamma' x + \zeta' t, \quad y = y', \quad z = z' \tag{2.7}$$

where γ, γ', ζ, and ζ' are constants, independent of position or time. That is, we have assumed for simplicity that the y and z transformations are the same as in the Galilean transformation and that the equations for $x(x', t')$ and for $x'(x, t)$ are still linear. We will have to see if these assumptions are consistent with the speed of light postulate; if not, we will have to try something more complicated. We will explicitly *not* assume that t is necessarily equal to t'.

Our goal now is to evaluate the constants γ, γ', ζ, and ζ'. We have four constants to determine, hence we need four physical conditions. We can determine three of the constants in terms of the fourth without ever invoking the second postulate.

First of all, from the meaning of the relative frame velocity V, the origin of the primed frame, $(x', y', z') = (0, 0, 0)$ must move with velocity V in the positive x direction as measured in the unprimed frame; *i.e.*, , if $x' = 0$, then $x = Vt$ (**condition 1**). This forces $\zeta' = -V\gamma'$ in the second of equation (2.7). We also want the origin of the unprimed frame to move in the *opposite* direction with speed V as measured in the primed frame; *i.e.*, if $x = 0$ then $x' = -Vt'$ (**condition 2**). This gives $\zeta = V\gamma$ in the first of equations (2.7). Therefore we can write

$$x = \gamma(x' + Vt'), \quad x' = \gamma'(x - Vt), \quad y = y', \quad z = z'. \tag{2.8}$$

Now the first postulate asserts that there is no preferred inertial frame of reference, so from the symmetry this implies we must have $\gamma' = \gamma$. Why is that?

Consider a clock A' at rest at the origin of $\mathcal{O}'$; it reads time t' and it always sits at $x' = 0$. When it reads $t' = 1$ s, its distance from the origin of $\mathcal{O}$, according to $\mathcal{O}$ observers, is $x = \gamma V(1\text{ s})$, from the first equation above. Consider another clock A at rest at the origin of $\mathcal{O}$; it reads time t and it always sits at $x = 0$. When it reads $t = 1$ s, its distance from the origin of $\mathcal{O}'$, from the point of view of $\mathcal{O}'$, is $x' = -\gamma' V(1\text{ s})$, from the second equation above: The minus sign simply reflects the fact that $\mathcal{O}$ moves in the *negative* x' direction from the point of view of $\mathcal{O}'$. However, except for this minus sign, which is related to the direction of travel, the distances moved by A and B when each reads 1 s *should be exactly the same*, according to the egalitarian first postulate. If they were different, it would allow us to say that one frame (say the frame in which the distance moved was greater) was fundamentally "better" that the other frame. This forces us to the conclusion that $\gamma' = \gamma$ (**condition 3**).

The transformation now becomes

$$x = \gamma(x' + Vt'), \quad x' = \gamma(x - Vt), \quad y = y', \quad z = z' \tag{2.9}$$

for some still undetermined value of γ. The Galilean transformation assumes $\gamma = 1$, but as we have seen, this choice is inconsistent with a universal speed of light.

We now finally require that if $x = ct$ then also $x' = ct'$, corresponding to a beam of light emitted from the mutual coordinate origins at the instant $t = t' = 0$ when the origins coincide (**condition 4**, *i.e.*, Postulate 2). The beam moves in the x directions of both frames with the same speed c, in which case the first two of equations (2.9) become

$$t = \gamma(1 + V/c)t' \quad \text{and} \quad t' = \gamma(1 - V/c)t. \tag{2.10}$$

We can eliminate t' by substituting the second equation into the first; this gives

$$t = \gamma^2(1 + V/c)(1 - V/c)t = \gamma^2(1 - V^2/c^2)t, \tag{2.11}$$

so we must choose $\gamma^2 = (1 - V^2/c^2)^{-1}$. Finally, we need the *positive* square root so that $\gamma \to 1$ as $V \to 0$, because as $V \to 0$ the two sets of axes coincide. Therefore the final results for $x(t', x')$ and $x'(t, x)$ are

$$x = \gamma(x' + Vt') \ \text{ and } \ x' = \gamma(x - Vt), \ \text{ where } \ \gamma = \frac{1}{\sqrt{1 - V^2/c^2}}. \tag{2.12}$$

We can now find the transformation equations for t and t'. Substitute $x' = \gamma(x - Vt)$ into the right-hand side of $x = \gamma(x' + Vt')$; the resulting equation can be solved for t' to give

$$t' = \gamma\left(t - \frac{Vx}{c^2}\right). \tag{2.13}$$

We can instead eliminate x between the two equations and then solve for t to give

$$t = \gamma\left(t' + \frac{Vx'}{c^2}\right), \tag{2.14}$$

which is the same as equation (2.13) if we interchange primed and unprimed coordinates and let $V \to -V$. Thus we have the amazing and profound result that *there is no longer an absolute time, the same in all frames.* Relativity shows that time and space have become closely intertwined.

The entire set of transformations from primed to unprimed coordinates can be written in the compact form

$$\begin{aligned} ct &= \gamma(ct' + \beta\, x'), \\ x &= \gamma(x' + \beta\, ct'), \\ y &= y', \\ z &= z', \end{aligned} \tag{2.15}$$

where

$$\beta \equiv V/c \quad \text{and} \quad \gamma = \frac{1}{\sqrt{1-\beta^2}}. \tag{2.16}$$

These equations are collectively called the **Lorentz transformation** or colloquially **Lorentz boost**. We have used the product ct in the equations, instead of just t by itself, so that the four coordinates ct, x, y, z all have the same dimension of length. The *inverse* Lorentz transformation, for (ct', x', y', z') in terms of (ct, x, y, z), is the same, with primed and unprimed coordinates interchanged and with $\beta \to -\beta$. We began with four constants $\gamma, \gamma', \zeta, \zeta'$ and found four conditions they must obey, which determined all four in terms of the relative frame velocity V.

Having found the transformation by invoking the speed of light only in the x direction, we can verify that the transformation works also for light moving in *any* direction. Let a flashbulb flash from the mutual origins of frames $\mathcal{O}$ and $\mathcal{O}'$ just as the origins pass by one another. In the unprimed frame, the square of the distance moved by the wavefront of light in time t is

$$x^2 + y^2 + z^2 = c^2 t^2. \tag{2.17}$$

That is, the light flash spreads out at speed c in all directions. Now, using the Lorentz transformation of equations (2.15), we can see how the flash moves in the primed frame. Rewriting (2.17) in terms of primed coordinates, we have

$$[\gamma(x' + \beta c t')]^2 + y'^2 + z'^2 = [\gamma(c t' + \beta x')]^2, \tag{2.18}$$

which, with a little algebra, yields

$$x'^2 + y'^2 + z'^2 = c^2 t'^2. \tag{2.19}$$

That is, the light flash also moves in all directions at speed c in frame $\mathcal{O}'$. Therefore the second postulate is obeyed for light moving in *any* direction, if we use the Lorentz transformation to transform coordinates.

Let us stare at the Lorentz transformation equation (2.15) for a while and observe some of its features:

- For $V \ll c$, *i.e.*, when the two observers are moving with respect to one another at a speed much less than that of light, we have $\beta \ll 1$ and $\gamma \sim 1$ to leading order in β, and the Lorentz transformation (2.15)) reduces to the Galilean transformation (1.1). That's a sanity check: our intuition led us to (1.1) because our daily experiences involve mechanics at speeds much less than that of light. Hence, we may still use Galilean transformations as long as we restrict ourselves to problems involving slow speeds and as long as we don't care about high-precision measurements. Obviously, Maxwell's equations involve light and so require the use of the full and correct form of the transformation of coordinates, the Lorentz transformations. This is why electromagnetism historically seeded the development of relativity.

- There are two novelties at work: the mixing of time and space, the coordinates t and x, t' and x'; and an interesting scale factor $\gamma \geq 1$. Figure 2.2 shows a plot of γ as a function of β. We can see that γ changes no more than 1% from unity for $0 \leq \beta \leq 0.1$, or speeds up to about 10% of light. A rough rule of thumb is then to require $v < 0.1c$

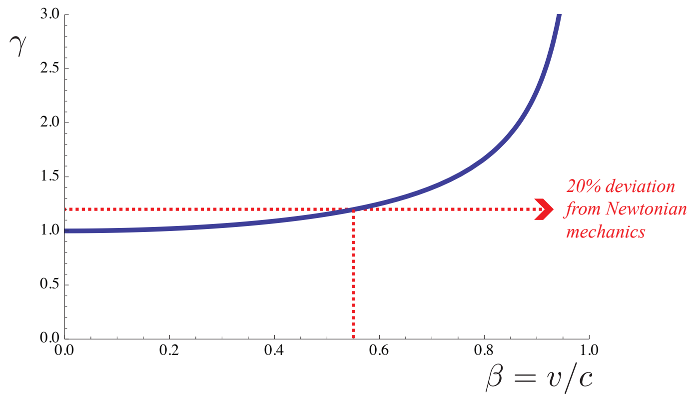


FIG 2.2 : Graph of the γ factor as a function of the relative velocity β. Note that $\gamma \cong 1$ for nonrelativistic particles, and $\gamma \to \infty$ as $\beta \to 1$.

for Newtonian mechanics. Note also the divergence as $V \to c$: the corresponding flip of the sign under the square root for $V > c$ implies an upper bound on speed $\beta < 1 \Rightarrow V < c$. Nature comes with a speed limit!

EXAMPLE 2-1: Rotation and rapidity

Consider two observers $\mathcal{O}$ and $\mathcal{O}'$, stationary with respect to one another and with identical origins, but with axes $\{x, y, z\}$ and $\{x', y', z'\}$ relatively rotated. Focus on a case where observer $\mathcal{O}$'s coordinate system is rotated with respect to $\mathcal{O}'$'s by a positive angle θ about the z axis

$$x = x' \cos\theta + y' \sin\theta, \quad y = -x' \sin\theta + y' \cos\theta, \quad z = z' \ . \tag{2.20}$$

It is often convenient to write this transformation in matrix notation:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} \ . \tag{2.21}$$

In general, a rotation can be written as

$$\boldsymbol{r} = \hat{\boldsymbol{\mathcal{R}}} \cdot \boldsymbol{r}' \ , \tag{2.22}$$

for $\boldsymbol{r} = (x, y, z)$, and $\boldsymbol{r'} = (x', y', z')$, with $\hat{\boldsymbol{\mathcal{R}}}$ a 3 by 3 rotation matrix satisfying the orthogonality condition $\hat{\boldsymbol{\mathcal{R}}}^t \cdot \hat{\boldsymbol{\mathcal{R}}} = \mathbf{1}$ as well as having the determinant $|\hat{\boldsymbol{\mathcal{R}}}| = 1$. Here $\hat{\boldsymbol{\mathcal{R}}}^t$ is the transpose matrix, the reflection of $\hat{\boldsymbol{\mathcal{R}}}$ about its principal diagonal.

Interestingly, we can present a Lorentz boost in analogy to rotations, making its structural form more elegant and transparent. To do so, we start by introducing a *four* component "position vector"

$$\boldsymbol{r} \equiv (c\,t, \boldsymbol{r}) = (c\,t, x, y, z)\,, \tag{2.23}$$

denoting an event in spacetime occurring at position (x, y, z) *and* time t. This is a natural notation, since Lorentz transformations mix space and time coordinates; again, we use $c\,t$ as the time component to give it the same dimension of length as the other components. We can now write the Lorentz boost in equation (2.15) as a matrix multiplication

$$\begin{pmatrix} c\,t \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \gamma & \gamma\,\beta & 0 & 0 \\ \gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} c\,t' \\ x' \\ y' \\ z' \end{pmatrix}. \tag{2.24}$$

Consider the parametrization

$$\gamma \equiv \cosh \xi \geq 1\ , \tag{2.25}$$

where ξ is called **rapidity**. Using the identity $\cosh^2 \xi - \sinh^2 \xi = 1$ one can easily show that

$$\gamma\,\beta = \sinh \xi\ , \tag{2.26}$$

so our Lorentz boost now becomes

$$\begin{pmatrix} c\,t \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \cosh \xi & \sinh \xi & 0 & 0 \\ \sinh \xi & \cosh \xi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} c\,t' \\ x' \\ y' \\ z' \end{pmatrix}, \tag{2.27}$$

much like a rotation but with *hyperbolic* trigonometric functions instead and a sign flip! We say that Lorentz transformations rotate time and space into one another.

We can write the most general Lorentz transformation in matrix notation as well,

$$\boldsymbol{r} = \hat{\boldsymbol{\Lambda}} \cdot \boldsymbol{r'}, \tag{2.28}$$

for $\boldsymbol{r} = (c\,t, x, y, z)$, and $\boldsymbol{r'} = (c\,t', x', y', z')$, with $\hat{\boldsymbol{\Lambda}}$ a *general* 4 by 4 matrix satisfying the condition

$$\hat{\boldsymbol{\Lambda}}^t \cdot \hat{\boldsymbol{\eta}} \cdot \hat{\boldsymbol{\Lambda}} = \hat{\boldsymbol{\eta}} \tag{2.29}$$

as well as

$$|\hat{\boldsymbol{\Lambda}}| = 1\ , \tag{2.30}$$

where $\hat{\boldsymbol{\eta}}$ is the 4 by 4 matrix

$$\hat{\boldsymbol{\eta}} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{2.31}$$

For a derivation of this general statement, see the Problems section of this chapter. Notice that

$$\hat{\boldsymbol{\Lambda}} = \begin{pmatrix} \cosh\xi & \sinh\xi & 0 & 0 \\ \sinh\xi & \cosh\xi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{2.32}$$

satisfies (2.29) and (2.30). Note also that $\hat{\boldsymbol{\eta}}$ is *almost* the identity matrix, but not quite, because of the minus sign in the first entry. It is known as the **metric** of flat spacetime. Correspondingly, $\hat{\boldsymbol{\Lambda}}$ satisfies an 'almost' orthogonality condition (2.29). We will revisit these observations in the upcoming sections as we develop our physical intuition for relativity.

2.2 Relativistic kinematics

Kinematics deals with how we describe motion, including the position, velocity, and acceleration of particles, for example, while stopping short of looking for physical underpinnings for that motion, which is the subject of *dynamics*. So we take up the essential topic of relativistic kinematics here, and then go on to relativistic dynamics in the following section.

2.2.1 Proper time

Consider a particle moving in the vicinity of an observer $\mathcal{O}$ who describes its trajectory by $x(t)$, $y(t)$, $z(t)$. The observer can describe the location of the particle in time and space using a position four-vector

$$\boldsymbol{r} = (c\,t, x, y, z) . \tag{2.33}$$

If dt, dx, dy, and dz represent infinitesimal steps in the evolution of the particle, we can also write the infinitesimal displacement four-vector as

$$d\boldsymbol{r} = (c\,dt, dx, dy, dz) . \tag{2.34}$$

Observer $\mathcal{O}$ may, for some yet mysterious reason, choose to compute the quantity

$$ds^2 = d\boldsymbol{r}^t \cdot \hat{\boldsymbol{\eta}} \cdot d\boldsymbol{r} = -c^2 dt^2 + dx^2 + dy^2 + dz^2. \tag{2.35}$$

The first part of this expression uses matrix notation: $\hat{\boldsymbol{\eta}}$ is the four by four matrix from (2.31)), and the 't' label denotes the transpose operation on the column vector that is $d\boldsymbol{r}$: that is, a row vector $d\boldsymbol{r}^t$ multiplies the matrix $\hat{\boldsymbol{\eta}}$ which then multiplies the column vector $d\boldsymbol{r}$. To see why this quantity is interesting to compute, consider the same quantity as computed by an observer $\mathcal{O}'$ at rest in the primed frame. Equation (2.28) prescribes that we must have

$$d\boldsymbol{r} = \boldsymbol{\Lambda} \cdot d\boldsymbol{r}' \ . \tag{2.36}$$

Substituting this into (2.35), we get

$$\begin{aligned} ds^2 &= d\boldsymbol{r}^t \cdot \hat{\boldsymbol{\eta}} \cdot d\boldsymbol{r} \\ &= d\boldsymbol{r}'^T \cdot \boldsymbol{\Lambda}^t \cdot \hat{\boldsymbol{\eta}} \cdot \boldsymbol{\Lambda} \cdot d\boldsymbol{r}' = d\boldsymbol{r}'^T \cdot \hat{\boldsymbol{\eta}} \cdot d\boldsymbol{r}' \\ &= -c^2 dt'^2 + dx'^2 + dy'^2 + dz'^2 \end{aligned} \tag{2.37}$$

where we used (2.29). Comparing (2.35) and (2.37), we now see that ds^2 is an *invariant* under Lorentz transformations! Observers $\mathcal{O}$ and $\mathcal{O}'$ use the same form of the expression in their respective coordinate systems and get the same value for ds^2. In general, quantities like ds^2 that remain the same under Lorentz transformations are said to be **scalar invariants** or **Lorentz invariants**.

There is a physical way to understand why ds^2 is the same in all inertial frames. Imagine that observer $\mathcal{O}'$ happens to be 'riding' with the particle at the given instant in time she measures the displacement four-vector $d\boldsymbol{r}'$. Observer $\mathcal{O}'$ would then see the particle momentarily at rest, with $dx' = dy' = dz' = 0$, since she is matching the particle's velocity at that instant: that is,

$$d\boldsymbol{r}' = (c\, dt', 0, 0, 0) \ . \tag{2.38}$$

Now $dt' \equiv d\tau$ is an advance in time on the watch of $\mathcal{O}'$, *i.e.*, a watch in the **rest frame** of the particle. We then have from (2.37)

$$ds^2 = -c^2 d\tau^2, \tag{2.39}$$

so the value of ds^2 measures the period of an infinitesimal clock tick *as measured in the rest frame of our particle*. No wonder it is an invariant quantity! The quantity τ is called the **proper time** of the particle. Equation (2.39) also helps us relate the proper time τ of the particle to the time t in the frame of reference of observer $\mathcal{O}$, since we know that

$$ds^2 = -c^2 d\tau^2 = -c^2 dt^2 + dx^2 + dy^2 + dz^2 \ . \tag{2.40}$$

Divide this equation by dt^2 to get

$$c^2 \frac{d\tau^2}{dt^2} = c^2 - \frac{dx^2}{dt^2} - \frac{dy^2}{dt^2} - \frac{dz^2}{dt^2} = c^2 - v^2 \tag{2.41}$$

where v is the speed of the particle, from which we find that

$$dt = \gamma \, d\tau \tag{2.42}$$

with $\gamma = 1/\sqrt{1-\beta^2}$ and $\beta \equiv v/c$. This implies that a time interval $d\tau$ in the rest frame of the particle is perceived by observer $\mathcal{O}$ as an interval $dt > d\tau$. The effect is known as **time dilation**: from the point of view of an inertial observer at rest in a frame in which some particle is moving with speed v, if the observer ages by (say) ten seconds, the particle may age by only one second in the observer's frame! We say the particle's time slows down as seen by observer $\mathcal{O}$. Note that this relation holds instantaneously even when the particle is accelerating. At every instant in time, its changing velocity results in a different amount for time dilation with respect to observer $\mathcal{O}$.

We have learned that time is not at all a universal observable: it is a 'malleable' quantity, with two observers in different reference frames disagreeing about its rate of advance. To talk about a notion of time that everyone agrees on in relativity, we need to refer to *proper* time — the time as measured in the rest frame of a particular reference observer or particle. For speeds small compared to that of light we have $\gamma \sim 1$ in (2.42), and we recover the approximate Galilean statement of universal time $dt = d\tau$.

At this point we will use our current discussion to introduce a notation that will come in handy for the rest of the book. We have already started to appreciate the elegance of lumping time and space together in a position four-vector, and we also demonstrated the use of matrix language in compactifying the notation. Putting these technologies together, let us label a position four-vector alternatively as

$$\boldsymbol{r} = (c\,t, x, y, z) \to (x_0, x_1, x_2, x_3) \ , \tag{2.43}$$

where $x_0 = ct$, $x_1 = x$, etc. Note in particular the superscript notation we adopt. We will reserve subscripts to distinguish between the coordinates of *different* particles. We can then neatly denote the components of the position four-vector as x_μ, where μ is an index that can be 0, 1, 2, or 3. We would then write the components of the displacement four-vector $d\boldsymbol{r}$ as dx_μ. Let us now rewrite equation (2.35) in terms of this new "index notation"

$$ds^2 = \sum_{\mu=0}^{3}\sum_{\nu=0}^{3} dx_\mu \eta_{\mu\nu} dx_\nu, \tag{2.44}$$

where we are now representing the $\hat{\boldsymbol{\eta}}$ matrix by its components: $\eta_{\mu\nu}$ is the entry in the $\hat{\boldsymbol{\eta}}$ matrix (2.31) in the μth row and νth column[1]. The two sums in the expression simply implement the usual matrix multiplication rule of multiplying rows against columns. Note also this expression is now in the form of a sum over numbers: dx_μ, dx_ν, $\eta_{\mu\nu}$ are just commutative numbers. Therefore we can write

$$ds^2 = \sum_{\mu=0}^{3}\sum_{\nu=0}^{3} dx_\mu \eta_{\mu\nu} dx_\nu = \sum_{\mu=0}^{3}\sum_{\nu=0}^{3} dx_\mu dx_\nu \eta_{\mu\nu} = \sum_{\mu=0}^{3}\sum_{\nu=0}^{3} \eta_{\mu\nu} dx_\mu dx_\nu . \tag{2.45}$$

As you can see, rewriting the sum symbol time after time becomes tedious. In most cases, indices such as μ and ν will always be summed over 0, 1, 2, and 3. Hence, we also adopt the **Einstein summation convention**: all indices appearing exactly twice in an expression are assumed to be summed over unless explicitly stated otherwise. For example, we may now write simply

$$ds^2 = \eta_{\mu\nu} dx_\mu dx_\nu \tag{2.46}$$

with an implied sum over μ and over ν. To say that ds^2 is a Lorentz invariant, we would then write

$$ds^2 = \eta_{\mu\nu} dx_\mu dx_\nu = \eta_{\mu'\nu'} dx_{\mu'} dx_{\nu'}, \tag{2.47}$$

[1]A note for readers with more advanced background in differential geometry: In more conventional and advanced notation, a distinction is made between upper and lower Lorentz indices – corresponding to mathematical objects in the so-called tangent and co-tangent spaces of spacetime. In our notation, all quantities are given in the *tangent space*; correspondingly, all index contractions with the spacetime metric must be explicitly shown. To simplify things, we also write all Lorentz indices as subscripts. For example, a tangent space vector's components v^μ will be denoted as v_μ; and a co-tangent space co-vector's components w_μ will be denoted as $\eta_{\mu\nu} w_\nu$.

where primed indices refer to coordinates in the coordinate system of $\mathcal{O}'$. Note that the $\eta_{\mu'\nu'}$'s are the same as the corresponding $\eta_{\mu\nu}$'s (see equation (2.31)).

Now we can rewrite equation (2.28) in our new notation:

$$dx_\mu = \Lambda_{\mu\nu'} dx_{\nu'} \ . \tag{2.48}$$

The components of the Lorentz matrix $\hat{\mathbf{\Lambda}}$ are represented by $\Lambda_{\mu\nu'}$ at the μth row and ν'th column. The ν' index appears twice in the expression on the right-hand side, so it is summed over: The sum implements the matrix multiplication $\mathbf{\Lambda} \cdot d\boldsymbol{r}'$. There is also a *single* index μ; that index is *not* repeated within the same expression, since it appears only *once* on each side of the equal sign and so is not summed over. For every possible value of μ we have a different equation — for a total of four. These are the relations for the four components of $d\boldsymbol{r}$. If we encounter an expression with an index that occurs *more* than twice in the same term, a mistake has been made. An expression $\eta_{\mu\nu} x_\mu A_\nu B_\mu$ is undefined, for example.

Index notation takes some time to get used to, but it is worth it. It is powerful, and the physics of relativity lends itself very naturally to this notation and language. It is one of those things that requires practice to get the hang of, but once mastery is achieved, it is difficult to remember how one got by in the past without it. As we proceed with the discussion of relativity, we adopt this notation from the outset to provide the reader with as much practice and exposure as possible.

2.2.2 Four-velocity

Calculus is the natural language of motion: Newton invented differential calculus to make the discussion of motion more natural and precise.[2] Similarly, four-vector notation is the natural language of relativity, because relativistic physics inherently mixes time and space. One could proceed without the use of this mathematical language of four-vectors, but that would come at the expense of unnecessarily convoluting the discussion of the physics. It is important, however, to appreciate where the physics starts and where it ends. The tool of four-vectors we will use in this section is just that, a mathematical tool. It comes with layers of logic that make the symmetries underlying

[2] Ironically, his masterpiece, the *Principia*, uses no calculus at all, because few of his readers would have understood it: the *Principia* presents instead an exposition of mechanics in a rather awkward mathematical language that often obscures the physics at hand, particularly to present-day observers familiar with calculus!

relativity more transparent and hence guides us to the next natural steps in the discussion. Throughout, we still need to rely on the independent statements of physics, *i.e.*, the postulates of relativity, including the universality of the speed of light.

We start by looking for an observable quantity that relates to good old *velocity*, but which also fits more naturally into our new language of relativity. We want a 'four-velocity', a quantity with four components that maps onto the usual velocity at small speeds. For this new quantity to be natural in relativity, it should transform under Lorentz transformations in a simple way. Let's call the component of four-velocity v_μ, with $\mu = 0, 1, 2, 3$ as usual; we then require that

$$v_\mu = \Lambda_{\mu\nu'} v_{\nu'} \ . \tag{2.49}$$

Whatever v_μ may be for observer $\mathcal{O}$, it relates to $v_{\nu'}$ as seen by observer $\mathcal{O}'$ by this simple Lorentz transformation. It also needs to be related to our usual notion of velocity — the rate of change of position per unit time. However, time is not a universally invariant notion. The closest we can get to this concept is *proper time*, so the obvious candidate for four-velocity is

$$v_\mu \equiv \frac{dx_\mu}{d\tau} \ . \tag{2.50}$$

In this expression dx_μ is the four-displacement of a particle observed by $\mathcal{O}$, and $d\tau$ is the advance in proper time of the particle in question — which both observers agree upon. Since $dx_\mu = \Lambda_{\mu\nu'} dx_{\mu'}$ and $d\tau$ is invariant, we see that this definition does lead to (2.49) as required. But how does it relate to good old velocity? To see this, we need to write v_μ explicitly in terms of the coordinates of a fixed observer, say $\mathcal{O}$:

$$\begin{aligned} v_\mu \ &\rightarrow \ (\frac{dx_0}{d\tau}, \frac{dx_1}{d\tau}, \frac{dx_2}{d\tau}, \frac{dx_3}{d\tau}) = (c\,\frac{dt}{d\tau}, \frac{dx}{d\tau}, \frac{dy}{d\tau}, \frac{dz}{d\tau}) \\ &= \ (\gamma c, \gamma\frac{dx}{dt}, \gamma\frac{dy}{dt}, \gamma\frac{dz}{dt}), \end{aligned} \tag{2.51}$$

where we used the time dilation relation obtained previously,

$$dt = \gamma\, d\tau \ . \tag{2.52}$$

Note that $\gamma = 1/\sqrt{1 - v^2/c^2}$, where v is the speed of the particle as seen by $\mathcal{O}$.

We now recognize the last three components of our four-vector as γ times the ordinary velocity of the particle! We may write more compactly

$$v_\mu \to (\gamma c, \gamma \boldsymbol{v}), \tag{2.53}$$

lumping the last three entries together. For a slow-moving particle, we have $\gamma \sim 1$ and $v_\mu \sim (c, \boldsymbol{v})$, so we have achieved our goal of embedding velocity into the four-vector language.

What have we gained from this exercise except mild levels of enjoyment? Well, we now know how the ordinary velocity transforms under Lorentz transformations, as we shall see!

EXAMPLE 2-2: The transformation of ordinary velocity

We can now relate the ordinary velocity of the particle $\boldsymbol{v}$ as measured by observer $\mathcal{O}$ to the velocity $\boldsymbol{v}'$ as measured by $\mathcal{O}'$. To see this, we go back to equation (2.49) and expand it in the explicit coordinates of $\mathcal{O}$ and $\mathcal{O}'$.

Let us set up the problem. A particle is moving around in spacetime, and observers $\mathcal{O}$ and $\mathcal{O}'$ are measuring its trajectory. Frame $\mathcal{O}'$ is moving with respect to $\mathcal{O}$ with speed V in the positive x direction and their spatial axes are otherwise aligned; we then have the Lorentz transformation matrix from (2.24)

$$\boldsymbol{\Lambda} = \begin{pmatrix} \gamma_V & \gamma_V \beta_V & 0 & 0 \\ \gamma_V \beta_V & \gamma_V & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{2.54}$$

where $\beta_V = V/c$ and $\gamma_V = 1/\sqrt{1-\beta_V^2}$. That is, in the context of this problem we have added the subscript V to the β and γ that describe the transformation between primed and unprimed frames with relative velocity $\boldsymbol{V}$, to distinguish it from the β and γ involving the velocity $\boldsymbol{v}$ of a particle in frame $\mathcal{O}$, and the β' and γ' involving the velocity $\boldsymbol{v'}$ of the particle in frame $\mathcal{O}'$.

We can now write equation (2.49) in matrix notation

$$\begin{pmatrix} \gamma c \\ \gamma v_x \\ \gamma v_y \\ \gamma v_z \end{pmatrix} = \begin{pmatrix} \gamma_V & \gamma_V \beta_V & 0 & 0 \\ \gamma_V \beta_V & \gamma_V & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \gamma' c \\ \gamma' v'_x \\ \gamma' v'_y \\ \gamma' v'_z \end{pmatrix} . \tag{2.55}$$

All that is left is a simple matrix multiplication to obtain the relativistic velocity addition law. From the first row, we find

$$\gamma = \gamma_V \gamma' \left(1 + \beta_V \frac{v'_x}{c}\right), \tag{2.56}$$

which we can then use in the other three rows to get the velocity transformation

$$\begin{aligned} v_x &= \frac{v'_x + V}{1 + V v'_x/c^2} \\ v_y &= \frac{v'_y}{\gamma_V \left(1 + V v'_x/c^2\right)} \\ v_z &= \frac{v'_z}{\gamma_V \left(1 + V v'_x/c^2\right)} . \end{aligned} \tag{2.57}$$

Let us analyze the physical content of these equations.

- As a sanity check, we should first take the small-speed limit, for which $V v_{x'}/c^2 \ll 1$ and $\gamma_V \sim 1$; then

  $$\begin{aligned} v_x &= v'_x + V \\ v_y &= v'_y \\ v_z &= v'_z \ . \end{aligned} \tag{2.58}$$

  These are the familiar Galilean addition of velocity rules we know and love. So far, so good.

- The real deal happens when speeds in the problem compete with that of light. Let us say the particle is seen by observer $\mathcal{O}'$ to be traveling at the speed of light in the x' direction, $v'_x = c$, $v'_y = v'_z = 0$. We then have

  $$\begin{aligned} v_x &= \frac{c + V}{1 + V/c} = c \\ v_y &= \frac{0}{\gamma \left(1 + V/c\right)} = 0 \\ v_z &= \frac{0}{\gamma \left(1 + V/c\right)} = 0 \ . \end{aligned} \tag{2.59}$$

  At this point, we are justified in getting slightly emotional about the matter: everything works as it is supposed to according to our original postulates!

- How about intermediate speeds, which bridge the gap between low speeds and the speed of light? The simplest way to see the implication is to plot v_x as a function of v'_x for fixed V. Figure 2.3 shows such a plot. We now see explicitly how relativity caps speeds to be below that of light!

- It is also interesting to note that the relativistic velocity addition law has non-trivial structure in the y and z directions, transverse to the relative motion of the two observers. This comes about from the ratio γ'/γ; *i.e.*, it is due to the fact that the tick-rates of the two observers' clocks are different. Even though transverse *distances* are not affected by a change of reference frame, velocity depends also on the duration of clock ticks of each observer.

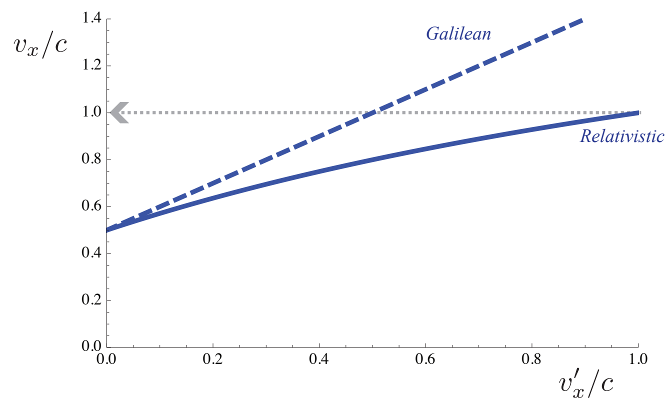


FIG 2.3 : The velocity v_x as a function of v'_x for fixed relative frame velocity $V = 0.5c$.

Note that without the use of four-vector language the derivation of these rather involved expressions for velocities would have been more painful. The formalism helps us embed velocity into a structure that transforms in a simple way under Lorentz transformations — given by (2.49). Yet, in explicit form, this rather simple expression metamorphoses to the beast that is (2.57). The strategy now becomes obvious: try to embed any physical quantity of interest into the language of four-vectors so that we get its Lorentz transformation for free; then decompose the transformation law into its components to see the physical implications.

EXAMPLE 2-3: Four-velocity invariant

Before we proceed to a similar treatment of momentum, let us introduce a simple technical exercise. We want to compute the quantity $v_\mu v_\nu \eta_{\mu\nu}$, in which the indices μ and ν are repeated and hence are to be summed over. It is an interesting quantity, since, writing it in matrix notation,

$$\boldsymbol{v}^t \cdot \hat{\boldsymbol{\eta}} \cdot \boldsymbol{v} = \boldsymbol{v'}^t \cdot \boldsymbol{\Lambda}^t \cdot \hat{\boldsymbol{\eta}} \cdot \boldsymbol{\Lambda} \cdot \boldsymbol{v'} = \boldsymbol{v'}^t \cdot \hat{\boldsymbol{\eta}} \cdot \boldsymbol{v'} \tag{2.60}$$

we get a Lorentz invariant, much like proper time. In index notation, we would write

$$v_\mu v_\nu \eta_{\mu\nu} = v_{\mu'} v_{\nu'} \eta_{\mu'\nu'} \ . \tag{2.61}$$

To compute this quantity, we write it in explicit form in terms of good old velocity,

$$v_\mu v_\nu \eta_{\mu\nu} = -\gamma^2 c^2 + \gamma^2 \left(v_x^2 + v_y^2 + v_z^2\right) = -\gamma^2 c^2 \left(1 - \frac{v^2}{c^2}\right) = -c^2 \ , \tag{2.62}$$

which is obviously an invariant quantity!

Now let us compute this same expression using a different technique. Since $v_\mu v_\nu \eta_{\mu\nu}$ is an invariant, we can evaluate it in *any* inertial frame. In particular, we can choose a frame $\mathcal{O}'$ that happens to be instantaneously at rest with respect to the particle whose velocity is represented in the expression. In that special frame we have $v'_x = v'_y = v'_z = 0$, and so we immediately get $v_{\mu'} v_{\nu'} \eta_{\mu'\nu'} = -c^2$. Therefore if we had been slightly more astute about things, we need not have done all our previous work in (2.62): by simply observing that we have an invariant, we would jump to a more convenient reference frame — the rest frame of the particle — and perform the computation there mentally. The moral: it pays to know your invariants!

2.3 Relativistic dynamics

We are now prepared to investigate relativistic *dynamics*, including the causes of changes in motion and such concepts as relativistic momentum and energy.

2.3.1 Four-momentum

In classical mechanics, the momentum of a particle is a three-vector, generally with components in all three spatial directions. Can we find a four-vector related to this good old classical momentum $\boldsymbol{p} = m\boldsymbol{v}$? Constructing it will help us learn how momentum changes when we shift perspective from one inertial observer to another in a fully relativistic context. This project turns out to be *easy*, because we have already found an expression for the four-velocity: The natural choice for the four-momentum is

$$p_\mu = m\, v_\mu \ . \tag{2.63}$$

However, we need to be slightly careful. To get the required simple transformation rule

$$p_\mu = \Lambda_{\mu\nu'} p_{\nu'} \tag{2.64}$$

from (2.49), we need the mass parameter $m = m'$ to be an invariant. We do not want to bias ourselves towards a physical assumption that has yet to come out of the postulates of relativity. Hence, we need to justify this statement. Fortunately, we have already seen a similar situation when we encountered the transformation of time. There, we found that the notion of invariant time required a definition of proper time: time in the rest frame of the observed particle. We can then safely adopt the same physical principle: we introduce the notion of **rest mass**, mass of a particle measured in its rest frame. That is obviously a quantity all observers would agree upon. Let's denote it by m and refer to it as simply **mass** from here on; it is the only mass the particle has, the same in all reference frames. The four-momentum is then $p_\mu = mv_\mu$, as given already in equation (2.63).

Let us look at the components of this new quantity and understand their physical significance. Recalling that the four-vector velocity has components $(\gamma c, \gamma \boldsymbol{v})$, it follows that for a particle of mass m moving with velocity $\boldsymbol{v}$ with $\beta = v/c$, the four-momentum is

$$p_\mu \rightarrow (\gamma mc, \gamma m\boldsymbol{v}) \tag{2.65}$$

where we have collected the last three terms together into a traditional three-vector. At low speeds this has the familiar form $\boldsymbol{p} \sim (mc, m\boldsymbol{v})$ to linear order in v/c, with the addition of the zeroth component mc. The nonrelativistic momentum $\boldsymbol{p} = m\boldsymbol{v}$ is seen to be just an approximation to the true momentum of a particle, $\boldsymbol{p} = \gamma m\boldsymbol{v}$.

Note that even though the velocity $\boldsymbol{v}$ of a particle is restricted to be $v < c$, because of the γ factor there is no upper limit to the momentum of a particle. As the speed of the particle gets ever closer to the speed of light, the momentum grows without bound. So far, things look promising.

What is the meaning of the quantity γmc, the *zeroth* component of the momentum four-vector? The first clue to its meaning is the fact that in Newtonian mechanics, the momentum of a particle is conserved if there are no forces on it, and that is true in all inertial frames. By the principle of relativity we want to retain this property for relativistic particles as well, which means that the last three components (called the spatial components) of the momentum four-vector should be conserved in the absence of forces. However, when we Lorentz-transform the spatial components of a four vector in one frame, they become a *mixture* of space and time components in another inertial frame. Therefore to ensure conservation of the spatial components in *all* frames means that the zeroth component (also called the time component) must be conserved as well! So the zeroth component of the momentum four-vector must also be some conserved quantity. What quantity could it be?

A second clue to the meaning of γmc comes from evaluating it for non-relativistic velocities. Using the binomial series

$$(1+x)^n = 1 + nx + \frac{n(n-1)}{2!}x^2 + \frac{n(n-1)(n-2)}{3!}x^3 + ..., \tag{2.66}$$

valid for $|x| < 1$, it follows that for nonrelativistic velocities $v/c \ll 1$,

$$\gamma mc = mc\left(1 - \frac{v^2}{c^2}\right)^{-1/2} \cong \frac{1}{c}\left(mc^2 + \frac{1}{2}mv^2\right) \tag{2.67}$$

keeping the first two terms in the binomial series. This quantity is indeed conserved for a free nonrelativistic particle. We recognize the second term as the nonrelativistic **kinetic energy** of the particle, which of course is conserved in the absence of forces, while the first term is an invariant quantity proportional to the particle's mass.

Therefore we identify the zeroth component of the momentum four vector as E/c, where E is the **energy** of the particle. In Newtonian mechanics we traditionally take the energy of a particle (subject to no forces or potential energies) to be *zero* if it is at rest, but we now see that a particle at rest has the **mass energy**

$$E_0 = mc^2, \tag{2.68}$$

and if the particle is moving, it also has the kinetic energy

$$T = E - E_0 = (\gamma - 1)mc^2, \tag{2.69}$$

which is approximately $(1/2)mv^2$ in the nonrelativistic limit $v/c \ll 1$.

In summary, the momentum four-vector is actually an "energy-momentum" four-vector, with components

$$p_\mu = \left(\frac{E}{c}, \boldsymbol{p}\right) \tag{2.70}$$

where $E = \gamma mc^2 = mc^2 + (\gamma - 1)mc^2$, in which the first term mc^2 is the particle's mass energy and the second term $(\gamma - 1)mc^2$ is its kinetic energy.

EXAMPLE 2-4: Relativistic dispersion relation

We start with a technical exercise with interesting physical implications. We want to compute the relativistic invariant

$$p_\mu p_\nu \eta_{\mu\nu} = p_{\mu'} p_{\nu'} \eta_{\mu'\nu'} \ , \tag{2.71}$$

which has a similar structure to $v_\mu v_\nu \eta_{\mu\nu} = -c^2$. In fact, since the four-momentum $p_\mu = mv_\mu$, we can immediately write

$$p_\mu p_\nu \eta_{\mu\nu} = -m^2c^2. \tag{2.72}$$

Let us expand this expression in components as seen by an observer $\mathcal{O}$. Using $p_\mu = (E/c, \boldsymbol{p})$, we easily get

$$-\frac{E^2}{c^2} + \boldsymbol{p}^2 = -m^2\, c^2 \ . \tag{2.73}$$

Alternatively, we write

$$E(p) = \sqrt{(m\,c^2)^2 + p^2 c^2} \tag{2.74}$$

where we have taken $E > 0$. This is the relativistic **dispersion relation** for a particle with mass m. The non-relativistic limit at low speeds corresponds to $p \ll mc$, which gives, after expanding to leading order in p,

$$E(p) \simeq mc^2 \left(1 + \frac{1}{2}\frac{p^2 c^2}{(m\,c^2)^2} + \cdots\right) = mc^2 + \frac{p^2}{2\,m} + \cdots \tag{2.75}$$

Once again, we see the contribution of the mass energy $m\,c^2$ as well as the Newtonian kinetic energy term $T = p^2/(2m) = (1/2)mv^2$. The full relativistic form of the dispersion relation (2.74) allows us to also consider the opposite limit $m \to 0$ or $p \gg mc$, the case of a light or 'massless' particle

$$E(p) \simeq p\,c \ . \tag{2.76}$$

In the strict limit $m \to 0$, this expression become exact. Hence, we have to entertain the possibility of a massless particle that carries energy by virtue of its momentum! Substituting $E = \gamma mc^2$ and $p = \gamma mv$ in this expression, we also get

$$\gamma mc = \gamma mv \to v = c \ . \tag{2.77}$$

Therefore we conclude that massless particles must travel at the speed of light. We can reverse the argument and say that a particle with $v = c \Rightarrow \gamma \to \infty$ must have zero mass if it is to have finite energy $E = \gamma mc^2$. Since light travels at the speed of light (hopefully), there is perhaps a sense in which we can think of light as a bunch of massless particles. Historically, this simple observation helped seed the foundations of **quantum mechanics**.

We may then think of energy as being a more fundamental physical quantity than mass. It exists irrespective of whether a particle has or does not have mass. We will see later on that energy is indeed more fundamental — through a discussion of symmetries and conservation laws. In due time, we will also see that massless particles gravitate as well, and that even gravity also cares about all sorts of energy, not just mass energy.

EXAMPLE 2-5: Decay into two particles

Many particles decay into two other particles, as illustrated in Figure 2.4. The initial particle of mass m_0 is shown in its rest frame; it has energy $m_0 c^2$ and momentum zero. It

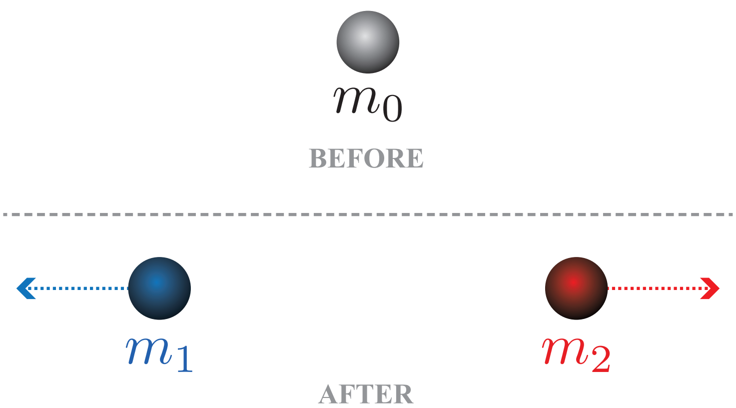


FIG 2.4 : A particle of mass m_0 decays into two particles with masses m_1 and m_2. Both energy and momentum are conserved in the decay, but mass is not conserved in relativistic physics. That is, $m_0 \neq m_1 + m_2$.

subsequently decays into two particles, with masses m_1 and m_2. These two final particles must move off in opposite directions to conserve momentum. As we will show, given the initial and final masses, conservation of energy and momentum are sufficient to determine the energies, momenta, and speeds of each final particle.

In relativity, just as in classical mechanics, we can assume that particles decay so quickly that any reasonable external forces have insufficient time to cause changes in momentum or energy during the very brief decay itself, so four-momentum is conserved. The initial four-momentum is entirely that of the particle of mass m_0,

$$p^i_\mu = \left(\frac{E_0}{c}, 0, 0, 0\right) \ . \tag{2.78}$$

The final four-momentum is the sum of two four-momenta

$$p^f_\mu = \left(\frac{E_1}{c}, \boldsymbol{p}_1\right) + \left(\frac{E_2}{c}, \boldsymbol{p}_2\right) \tag{2.79}$$

where we have aligned the x axis along the direction in which the two particles fly off.

Since we need

$$p^i_\mu = p^f_\mu \ , \tag{2.80}$$

we immediately see that

$$E_0 = E_1 + E_2 \tag{2.81}$$

and

$$0 = \boldsymbol{p}_1 + \boldsymbol{p}_2. \tag{2.82}$$

The mass m_0 is given; hence so is E_0, since

$$E_0 = m_0 c^2 \tag{2.83}$$

for $p_0 = 0$. The conservation law implies that the two particles emerge in opposite directions. Looking back at (2.81)) and (2.82), we actually have four unknowns that describe the problem: E_1, E_2, and the magnitudes of the momenta p_1 and p_2. Can we then unravel the kinematics of this problem with the only information we have, the masses of the particles? Equation (2.81) gives us one condition. Equation (2.82) leads to $p_1 = p_2$ – the magnitudes of the momenta must be the same – which is a second condition. We then need two additional conditions. These are the relations $E^2 - p^2c^2 = m^2c^4$ for each of the two emerging particles. Hence, we know the problem is solvable[3].

It follows that

$$E_2^2 - p_2^2c^2 = m_2^2c^4 = (m_0c^2 - E_1)^2 - p_1^2c^2 \tag{2.84}$$

using energy conservation for the first term and momentum conservation for the second term. Multiplying out the right-hand side, we find that

$$m_2^2c^4 = m_0^2c^4 - 2m_0c^2E_1 + E_1^2 - p_1^2c^2 = m_0^2c^4 - 2m_0c^2E_1 + m_1^2c^4. \tag{2.85}$$

We can then solve this last equation for E_1, giving

$$E_1 = \left(\frac{m_0^2 + m_1^2 - m_2^2}{2m_0}\right)c^2. \tag{2.86}$$

Thus we have solved the problem. Having found E_1 in terms of known quantities, we can also find E_2, both momenta, the particle velocities, and other quantities as well, using the conservation laws along with $E^2 - p^2c^2 = m^2c^4$.

In nuclear or particle physics, where two-particle decays are common, one usually uses *energy units* in calculations. In energy units the energy of a particle is given in MeV (10^6 electron volts), momenta in MeV$/c$, and masses in MeV$/c^2$. As a simple example, the π^0 meson, with mass 135 MeV$/c^2$, decays into two photons, each massless. Therefore the energy of photon 1 as seen from the rest frame of the meson is

$$E_1 = \left(\frac{m_0^2 + 0 - 0}{2m_0}\right)c^2 = \frac{m_0}{2}c^2 = 67.5\ \text{MeV} \tag{2.87}$$

and the magnitude of its momentum is $p_1 = E_1/c =$ 67.5 MeV$/c$.

[3] The reader may worry about one more unknown in the full problem: the angle at which the two particles emerge back to back. But this angle is undetermined because of the spherical symmetric attributes of the setup: any angle is consistent with energy and momentum conservation. Fixing the angle would require additional physical information about the natural laws underlying the decay process at hand.

2.3.2 Four-force

Finally, we seek a four-vector *force* that is responsible for *changes* in the four-momentum of a particle. A "four-force" would allow us to reformulate Newton's second law for relativistic mechanics, since non-conservation of momentum in Newtonian physics implies the presence of a force $\boldsymbol{F} = d\boldsymbol{p}/dt$. We then want to embed the notion of force in a four-vector as well. Let us label the four-vector force as f_μ and propose that

$$f_\mu = \frac{dp_\mu}{d\tau} \ . \tag{2.88}$$

We see that once again we have used proper time to measure rate of change. Therefore, observer $\mathcal{O}'$ would write

$$f_{\mu'} = \frac{dp_{\mu'}}{d\tau} \tag{2.89}$$

to describe the same physics – with the implicit Lorentz transformation of our new four-force $f_\mu = \Lambda_{\mu\mu'} f_{\mu'}$.

A force law is an independent statement of physics, so one then needs to check each component of f_μ – its detailed form in terms of the parameters of the particle and its environment – to see whether the Lorentz transformation changes it beyond the expected $f_\mu = \Lambda_{\mu\mu'} f_{\mu'}$. Since all inertial observers are to see the same physics, this should not happen! For now, let us assume that whatever forces appear on the left-hand side of (2.88)) are indeed consistent with the postulates of relativity. We want instead to focus on a much more urgent issue: what new physics does our reformulation of Newton's second law given by (2.88 add to the dynamics, on top of what we already know from the Newtonian realm?

To see the implications of (2.88), let us write it explicitly in component form,

$$(f_0, \boldsymbol{f}) = \gamma \frac{d}{dt} (\gamma mc, \gamma m\boldsymbol{v}) \tag{2.90}$$

where we use the time dilation relation (2.52) to write $d\tau$ in terms of observer $\mathcal{O}$'s time differential dt, and we collect the three spatial components of our four-vectors into the usual three-vector notation. We look at the easy part first: the spatial components are

$$\boldsymbol{f} = \gamma \frac{d}{dt} (\gamma m\boldsymbol{v}) \ . \tag{2.91}$$

Imagine that the particle is subject to no external forces, $\boldsymbol{f} = 0$. We then have momentum conservation

$$\frac{d}{dt}(\gamma m\boldsymbol{v}) = 0 \ , \tag{2.92}$$

where again the momentum is $\boldsymbol{p} = \gamma m\boldsymbol{v}$, and not the nonrelativistic approximation $\boldsymbol{p} = m\boldsymbol{v}$.

Defining force $\boldsymbol{F}$ as the rate of change of $\boldsymbol{p}$, we would need to write

$$\boldsymbol{F} \equiv \frac{d}{dt}(\gamma m\boldsymbol{v}) \ . \tag{2.93}$$

The quantity $\boldsymbol{F}$ corresponds to the force in Newtonian mechanics: it is the rate of change of momentum as seen by a given observer. Looking back at (2.91), we then interpret the lower case quantity $\boldsymbol{f}$ as

$$\boldsymbol{f} = \gamma\boldsymbol{F} \ . \tag{2.94}$$

Now what is the meaning of the *zeroth* component of (2.90)?

$$f_0 = \gamma\frac{d}{dt}(\gamma mc)? \tag{2.95}$$

Recall that the energy of the particle is $E = \gamma mc^2$, so

$$f_0 = \frac{d}{d\tau}(\gamma mc) = \frac{1}{c}\frac{dE}{d\tau}. \tag{2.96}$$

Earlier we showed that E and p obey $E^2 = m^2c^4 + p^2c^2$, so differentiating this equation with respect to τ gives

$$2E\frac{dE}{d\tau} = 2\left(p_x\frac{dp_x}{d\tau} + p_y\frac{dp_y}{d\tau} + p_z\frac{dp_z}{d\tau}\right)c^2 \equiv 2\left(\boldsymbol{p}\cdot\frac{d\boldsymbol{p}}{d\tau}\right)c^2, \tag{2.97}$$

and so, using $E = \gamma mc^2$ and $\boldsymbol{p} = \gamma m\boldsymbol{v}$,

$$\frac{dE}{d\tau} = \frac{1}{\gamma mc^2}(\gamma m\boldsymbol{v}\cdot\boldsymbol{f})\,c^2 = \boldsymbol{v}\cdot\boldsymbol{f} \Rightarrow \frac{dE}{dt} = \boldsymbol{F}\cdot\boldsymbol{v} \tag{2.98}$$

which we recognize as the rate at which the force does work on the particle, *i.e.*, the *power input* to the particle. So finally the four components of the force four-vector are

$$f_\mu \to \left(\gamma\frac{1}{c}\boldsymbol{v}\cdot\boldsymbol{F}, \gamma\boldsymbol{F}\right), \tag{2.99}$$

where the zeroth component of the four-force is the rate at which the energy flows in/out of the system. We now have also learned how force must transform under Lorentz transformations, since f_μ is a four-vector and we have $f_{\mu'} = \Lambda_{\mu' \mu} f_\mu$.

Summarizing the dynamical results so far, we are led by the postulates of relativity to a modification of the transformation rules that relate the perspectives of inertial observers. We then developed a mathematical language that naturally lends itself to Lorentz transformations, and we discussed four-vectors and Lorentz invariants. Next, we attempted to embed various physical quantities, such as velocity, momentum and force, into the language of four-vectors. In doing so, we wrote quantities that match the corresponding Newtonian ones at low speeds, but are packaged in a manner that easily determines how they change under Lorentz transformations. This led us to a revised velocity addition law, a new understanding of momentum and energy, including a realization that mass is a form of energy, and finally a revised concept of force and of Newton's second law of motion.

2.3.3 Dynamics in practice

At this point it is useful to step back and think about mechanics in light (no pun intended) of all the new revisions we have talked about. A nice organizing framework is to revisit the three laws of Newton and fit them into the postulates of relativity.

- **Unchanged first law**: There exists a class of observers — henceforth labeled **inertial observers** — for whom the laws of physics are the same. Given one inertial observer, another observer is inertial if their two frames have a constant relative velocity. In an inertial frame, in the absence of forces, a particle will move in a straight line at constant speed.

- **New fourth law**: The universal **speed of light** is a law of physics: light moves at the same speed with respect to all inertial observers. This implies that the inertial reference frames defined in the first law are connected to each other by **Lorentz transformations**.

- **Revised second law**: The rate of change of four-momentum is the four-force

$$f_\mu = \frac{d}{d\tau}\left(m v_\mu\right) = \frac{dp_\mu}{d\tau} \tag{2.100}$$

where $v_\mu = dx_\mu/d\tau$ and τ is proper time. In the absence of a four-force, energy and momentum are conserved.

- **Extended third law**: For every four-force there is an equal but opposite four-force. The spatial part of this statement ensures total momentum conservation for an isolated system: for an isolated system of particles, action-reaction pairs cancel so that the total force is zero and total momentum is conserved. We will see this in more detail in a later chapter on systems of particles. The temporal part of our new statement is about energy conservation for an isolated system: you can see this from the fact the the fourth component of the four-force measures rate of change of energy.

One can then use these statements to study mechanics with speeds all the way up to that of light. At low speeds we drop the new fourth law and recover Newton's three laws as approximate laws of physics, and Galilean transformations connect inertial reference frames. Beyond these four statements, what then remains is to complete the dynamical picture by incorporating forces consistent with the postulates of relativity.

Physical intuition is developed through explicit examples. Hence, we proceed now with a few case studies.

EXAMPLE 2-6: Uniformly accelerated motion

Consider a particle of mass m moving in one spatial direction, say along the x axis of an observer $\mathcal{O}$, and suppose that this particle is subjected to an external four-force

$$(f_0, \boldsymbol{f}) = \left(\gamma \frac{1}{c} \boldsymbol{v} \cdot \boldsymbol{F}, \gamma \boldsymbol{F}\right), \tag{2.101}$$

with $\boldsymbol{F}$ a constant three-vector pointing in the positive x direction. We want to see whether such a constant force can accelerate the particle past the speed of light! Writing the component of $f_\mu = dp_\mu/d\tau$ in the x direction, we get

$$\gamma F_x = \gamma \frac{d}{dt}\left(\gamma m v_x\right) \ . \tag{2.102}$$

Simplifying things, we have

$$F_x = m\frac{d}{dt}\left(\gamma v_x\right), \tag{2.103}$$

which is a differential equation we can solve for v_x. The left-hand side is a constant, and the velocity $v_x(t)$ appears both explicitly and also implicitly in the gamma-factor $\gamma = (1 - (v_x)^2/c^2)^{-1/2}$. Integrating equation (2.103) with $v_x(0) = 0$, we get

$$\frac{F_x}{m}t = \frac{v_x}{\sqrt{1 - v_x^2/c^2}} \tag{2.104}$$

which we can solve for $v_x(t)$,

$$\frac{v_x(t)}{c} = \frac{F_x t/mc}{\sqrt{1 + \left(F_x t/mc\right)^2}} \ . \tag{2.105}$$

We recognize $a = F_x/m$ as the Newtonian acceleration, which is a constant in this case. Therefore in terms of a,

$$\frac{v_x(t)}{c} = \frac{a\,t/c}{\sqrt{1 + \left(a\,t/c\right)^2}} \ . \tag{2.106}$$

The at factor looks very familiar, but the square root piece changes the ball game. At early times, when the particle has not yet acquired enough speed, we have $a\,t/c \ll 1$ and we recover the Newtonian expression $v_x(t) = a\,t$. At large times, however, relativity kicks in to ensure we do not violate the upper speed limit

$$\frac{v_x(t)}{c} \to 1 \tag{2.107}$$

as $t \to \infty$. Figure 2.5(a) shows a plot of $v_x(t)$ with the corresponding tapering-off feature at large speeds.

We can also look at the particle's trajectory, shown in Figure 2.5(b) by integrating

$$\frac{dx}{dt} = \frac{at}{\sqrt{1 + (at/c)^2}} \tag{2.108}$$

with $x(0) = 0$ for a particle that starts at the origin. One gets

$$x^2 - c^2t^2 = \frac{c^4}{a^2} \ . \tag{2.109}$$

In the $c\,t$-x plane, this has the shape of a hyperbola. We will revisit this in the next section when we discuss Minkowski diagrams.

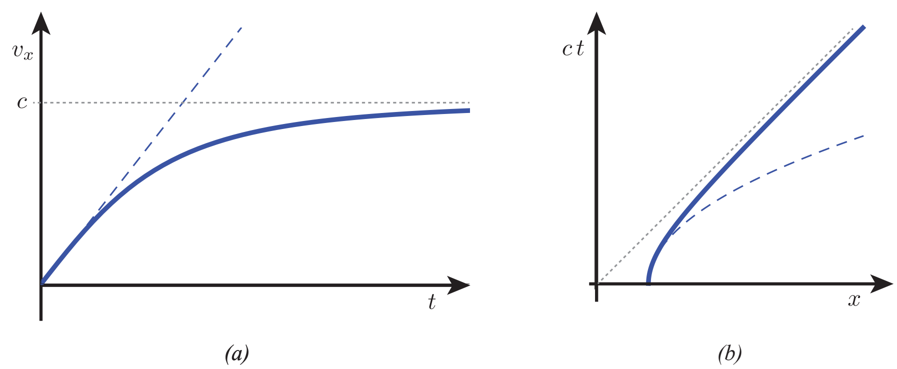


FIG 2.5 : Plots of relativistic constant-acceleration motion. (a) shows $v_x(t)$, demonstrating that $v_x(t) \to c$ as $t \to \infty$, *i.e.*, the speed of light is a speed limit in Nature. The dashed line shows the incorrect Newtonian prediction. (b) shows the hyperbolic trajectory of the particle on a $c\,t$-x graph. Once again the dashed trajectory is the Newtonian prediction.

EXAMPLE 2-7: The Doppler effect

The Doppler effect is the shifting of frequencies of sound or light between the perspectives of observers who are moving with respect to one another. We are most familiar with it in the context of sound (because the speed of sound is much less than the speed of light), when for example we notice a change in the pitch of the siren of an ambulance as it passes by. Sound also propagates in some *medium*, whether air, liquid, or solid, so that it has a particular fixed speed given by the properties of the particular medium *in the medium's rest frame*. Its speed is therefore *not* an invariant and will be subject to the velocity addition rule. Hence, the more interesting scenario for relativity involves the Doppler effect for light, because in that case there is no medium to provide a preferred frame of reference. We want to find how the frequency of light shifts as seen by different moving observers.

Consider our usual setup of observer $\mathcal{O}'$ moving with a constant speed V along the positive x direction *towards* another observer $\mathcal{O}$ as shown in Figure 2.6. Observer $\mathcal{O}'$ aims a laser beam of frequency ν' towards $\mathcal{O}$, and we want to find the frequency ν perceived by $\mathcal{O}$; that is, we seek the Lorentz transformation of **frequency**.

In another of his landmark papers of 1905, Einstein showed that light consists of particles now called **photons**, and that the energy E and momentum p of a photon are each proportional to frequency, $E = h\nu$ and $p = E/c = h\nu/c$, where h is Planck's constant. This means that the four-momentum of the laser beam is

$$p_\mu = \left(\frac{E}{c}, p, 0, 0\right) = \left(\frac{h\,\nu}{c}, \frac{h\,\nu}{c}, 0, 0\right) . \tag{2.110}$$

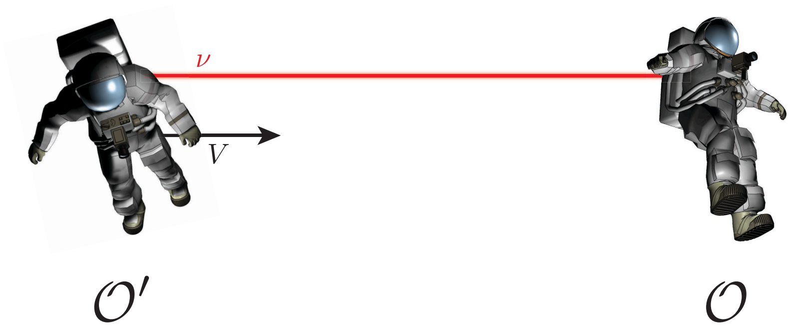


FIG 2.6 : Observer $\mathcal{O}'$ shooting a laser towards observer $\mathcal{O}$ while moving towards $\mathcal{O}$.

All that is left to do is write the Lorentz transformation $p_\mu = \Lambda_{\mu\nu'} p_{\nu'}$ in explicit component form. That is,

$$\begin{pmatrix} h\,\nu/c \\ h\,\nu/c \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \gamma_V & \gamma_V\,\beta_V & 0 & 0 \\ \gamma_V\,\beta_V & \gamma_V & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} h\,\nu'/c \\ h\,\nu'/c \\ 0 \\ 0 \end{pmatrix} \tag{2.111}$$

where $\beta_V = V/c$. This leads to

$$\nu = \gamma_V\,\nu' + \gamma_V\,\beta_V\,\nu' , \tag{2.112}$$

where Planck's constant has dropped out of the equation. A little algebra then shows that

$$\frac{\nu}{\nu'} = \sqrt{\frac{1+\beta}{1-\beta}} > 1. \qquad \textit{approaching} \text{ observers} \tag{2.113}$$

This applies to the scenario where the laser beam is aimed from $\mathcal{O}'$ *towards* $\mathcal{O}$ as $\mathcal{O}'$ moves in the positive x direction – implicit in the fact that the x component of p_μ in (2.110) is taken to be positive and it is assumed that the beam does arrive at $\mathcal{O}$. In short, this applies when the distance between the two observers is *shrinking*. The frequency received is greater than the frequency emitted, $\nu > \nu'$, known as a **blueshift**, for obvious reasons. To see the other possibility – *i.e.*, $\mathcal{O}$ and $\mathcal{O}'$ moving *away* from each other, we can just flip the sign of β in this expression

$$\frac{\nu}{\nu'} = \sqrt{\frac{1-\beta}{1+\beta}} < 1. \qquad \textit{receding} \text{ observers} \tag{2.114}$$

Now the distance between the two observers is increasing, and we find that the received frequency is *less* than the emitted one: we say there has been a *redshift*.

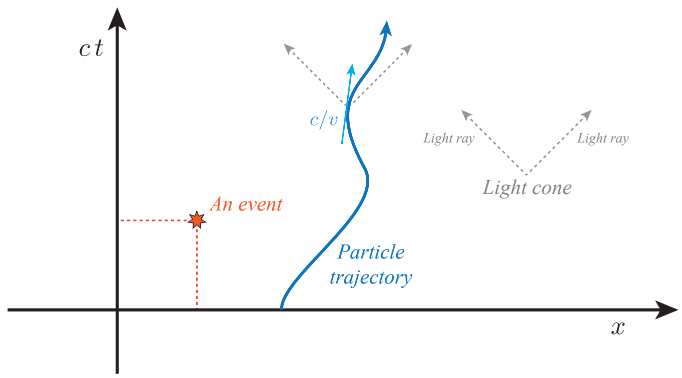


FIG 2.7 : A point on a Minkowski diagram represents an event. A particle's trajectory appears as a curve with a slope that exceeds unity everywhere.

Doppler shifts for light are an extremely useful tool in physics and technology, from determining the speeds of stars in our own galaxy and of distant galaxies, leading to Hubble's discovery of the expanding universe, to the use of frequency shifts in the Global Positioning System (GPS) for location and navigation. In fact, if special-relativistic time and Doppler-shift predictions were not included there would be large errors in position measurements using GPS. Interestingly, it turns out to be equally essential for GPS to include additional effects due to Earth's gravity, as contained in Einstein's **general theory of relativity**.

2.3.4 Minkowski diagrams

A particularly useful way to depict relativistic dynamics involves a visual tool called a **Minkowski diagram**. Simply put, it is a plot of the trajectory of a particle on a graph where the horizontal axis is one of the spatial directions and the vertical axis represents time, or actually the product ct. Figure 2.7 shows an example. The trajectory of the particle appears as a line moving upward, forward in time. It is sometimes referred to as the **worldline** of the particle. Light rays appear on such a diagram as straight lines at $45°$, as shown in the figure. A tangent to a trajectory corresponds to c/v, the inverse relative speed of the corresponding particle. Notice that the worldline of the particle in the figure has a slope greater than unity everywhere, since $v/c < 1$.

An isolated point on a Minkowski diagram corresponds to an event, since

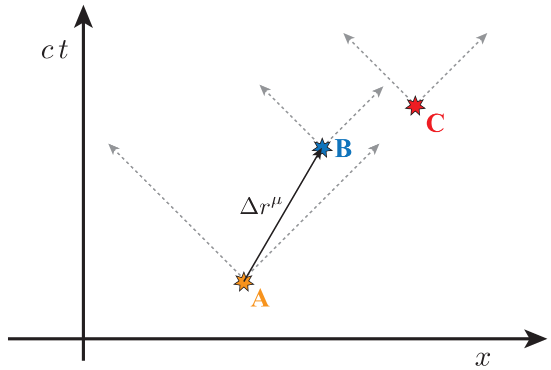


FIG 2.8 : Three events on a Minkowski diagram. Events A and B are timelike separated; A and C are lightlike separated; and B and C are spacelike separated.

it has a definite time and position. If two events can be connected by the physical trajectory of a particle (whose slope is everywhere greater than unity), the two events are said to be **timelike separated**. The physical implication is that earlier events can *talk* to the later event with physical signals involving matter or light. A quick way to determine whether two events are timelike separated is to draw a forward pointing **lightcone** wedge with its apex at the earlier event, as shown in Figure 2.8. The other event should then lie within the lightcone. We say that the two events are causally connected; *i.e.*, event A in the figure can cause event B. Event C lies outside the lightcone for B: reaching it requires signal propagation faster than light, *i.e.*, a curve that has at least some interval where its slope is less than unity. Such events are said to be causally disconnected; we also say events B and C are **spacelike** separated. Event C in the figure lies on the lightcone of A. This means that it can be reached from A with a light signal. A and C are then said to be **lightlike** separated.

There is an algebraic way to determine whether two events are lightlike, spacelike, or timelike separated. If we look at the position four-vector $\Delta r_\mu = (c\,\Delta t, \mathbf{\Delta r})$ pointing from one event to the other (see Figure 2.8), if the slope of this four-vector on the corresponding Minkowski diagram is greater than unity, then the events are timelike separated and we have

$$\Delta r_\mu \Delta r_\nu \eta_{\mu\nu} < 0 \; . \tag{2.115}$$

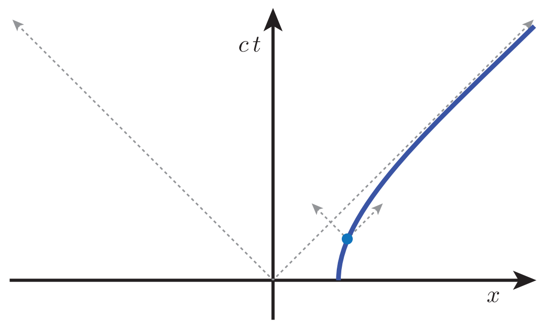


FIG 2.9 : The hyperbolic trajectory of a particle undergoing constant acceleration motion on a Minkowski diagram.

Similarly, we would get

$$\Delta r_\mu \Delta r_\nu \eta_{\mu\nu} > 0 \tag{2.116}$$

for spacelike separated events, and

$$\Delta r_\mu \Delta r_\nu \eta_{\mu\nu} = 0 \tag{2.117}$$

for lightlike separated ones. It is also useful to define this concept for any four-vector, such as the velocity, momentum, and force four-vectors. For any such four-vector, denoted by A_μ in general, we can write

$$\begin{aligned} A_\mu A_\nu \eta_{\mu\nu} &> 0 \quad \text{spacelike} \\ A_\mu A_\nu \eta_{\mu\nu} &< 0 \quad \text{timelike} \\ A_\mu A_\nu \eta_{\mu\nu} &= 0 \quad \text{lightlike} \end{aligned} \tag{2.118}$$

Note in particular that the momentum and velocity four-vectors are timelike, while the force four-vector is spacelike.

As an exercise in Minkowski diagrams, consider the trajectory of a particle under the influence of a constant four-force, as encountered in a previous example. From (2.109), we can now plot the profile of the worldline in Figure 2.9. We see that the particle starts at rest with infinite slope (*i.e.*, zero speed), then speeds up and asymptotically reaches the speed of light at

$45°$ slope in the figure. We note that the slope is everywhere greater than unity, as expected.

Another use of Minkowski diagrams is to picture the relation between the coordinate systems of two observers. The same set of events on a Minkowski diagram can get labeled via different coordinates by different inertial observers. Figure 2.10 shows the grid lines of observer $\mathcal{O}'$, who happens to be moving with speed V along the x axis of $\mathcal{O}$. The ct' axis is the worldline of observer $\mathcal{O}'$ as seen by $\mathcal{O}$, since it is obtained by setting $x' = 0$: After all, the ct axis of $\mathcal{O}$ is nothing but the trajectory of its origin on the Minkowski diagram at $x = 0$. From (2.15)), we then see that the $c\,t'$ axis is a straight line with slope c/V. The x' axis is given by $ct' = 0$ in the same spirit (as is the x axis of $\mathcal{O}$ given by the $ct = 0$ condition); from (2.15) we can see that it is a straight line with slope V/c. The ct' and x' axes are reflected images of each other across the lightcone at the origin. The figure shows a geometric realization of how an event gets labeled by the two observers: each projects the event along her time and space axes, along ct and x for $\mathcal{O}$, and ct' and x' for $\mathcal{O}'$. The reader is however cautioned in using concepts from Euclidean geometry on the diagram for measuring distances. The vertical axes here represent time! To measure the spacetime "distance" between two events separated by say Δt and Δx, you want to use $-c^2\Delta t^2 + \Delta x^2$, *not* $c^2\Delta t^2 + \Delta x^2$. That is, you want to use the Minkowski metric (2.31). Let us look at some examples using Minkowski diagrams to develop our visual intuition of relativity.

EXAMPLE 2-8: Time dilation

Consider our usual setup of two observers $\mathcal{O}$ and $\mathcal{O}'$. The Minkowski diagrams are shown in Figure 2.11 corresponding to a relative velocity $V = (3/5)c$, *i.e.*, observer $\mathcal{O}'$ is moving in the positive x direction with $(3/5)c$ relative to $\mathcal{O}$. In Figure 2.11(a), we show two events corresponding to two ticks of a clock carried by $\mathcal{O}$. In Figure 2.11(b), we show two events corresponding to two ticks of a clock carried by $\mathcal{O}'$ instead. Let us focus on Figure 2.11(a). $\mathcal{O}$'s clock ticks are separated by Δt. Using (2.15) with $\Delta x = 0$, we have

$$c\,\Delta t' = \gamma(c\,\Delta t - \beta\Delta x) \Rightarrow \Delta t' = \gamma\Delta t\ . \tag{2.119}$$

The corresponding time interval $\Delta t'$ observed in the primed frame is then *greater* than Δt. To $\mathcal{O}'$, this clock is moving in the negative x' direction and runs slow: this is the phenomenon of time dilation. Putting numbers in with $V = (3/5)c$, we have $\Delta t' = \gamma\Delta t = \Delta t/\sqrt{1 - V^2/c^2} = (5/4)\Delta t$. Figure 2.11(a) shows the same conclusion graphically.

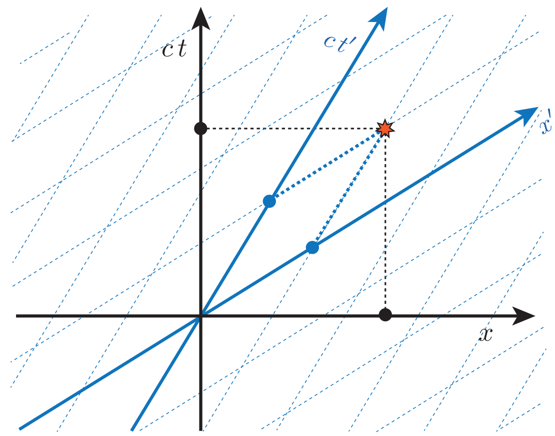


FIG 2.10 : The grid lines of two observers labeling the same event on a spacetime Minkowski diagram.

What if we reverse our perspective? Consider a clock carried by $\mathcal{O}'$ instead, which ticks with interval $\Delta t'$? Figure 2.11(b) depicts the corresponding scenario. Algebraically, the tick events of $\mathcal{O}'$'s clock have $\Delta x' = 0$. Using (2.15) once again, we now get

$$c\,\Delta t = \gamma(c\,\Delta t' + \beta \Delta x') \Rightarrow \Delta t = \gamma \Delta t' \; . \tag{2.120}$$

Observer $\mathcal{O}$ will then perceive this clock-tick separation as $\Delta t = (5/4)\Delta t' > \Delta t'$. To $\mathcal{O}$, this clock is moving in the positive x direction, and once again runs slow. In summary, from the standpoint of any inertial observer, a moving clock runs slow by a factor of γ.

EXAMPLE 2-9: Length contraction

Minkowski diagrams are shown in Figure 2.12 for a primed frame $\mathcal{O}'$ and unprimed frame $\mathcal{O}$ corresponding to a relative velocity $V = (3/5)c$. Figure 2.12(a) depicts a scenario where observer $\mathcal{O}'$ carries a meter stick along with her. The dashed lines are the trajectories of the endpoints of the meter stick. If $\mathcal{O}'$ wants to measure the length of the stick, she must measure the locations of both ends *at the same time* t'. The corresponding measurement is shown in Figure 2.12(a) through two events occurring at $t' = 0$ at the endpoints. We then have $\Delta t' = 0$ and $\Delta x' = L_0$ where L_0 is the **rest length** of the stick. If observer $\mathcal{O}$ is to measure the length of the same stick, he must use two events at the endpoints of the stick simultaneously in *his* reference frame, *i.e.*, two events with $\Delta t = 0$ and some value of Δx. Using (2.15) with $\Delta t' = 0$ and $\Delta x' = L_0$, one gets, after some straightforward algebra, $\Delta x = L_0/\gamma = (4/5)L_0 < L_0$! The moving stick is therefore shorter to $\mathcal{O}$. This is the phenomenon of **length contraction** or **Lorentz contraction**. If we consider a stick carried by $\mathcal{O}$ instead, the scenario is shown

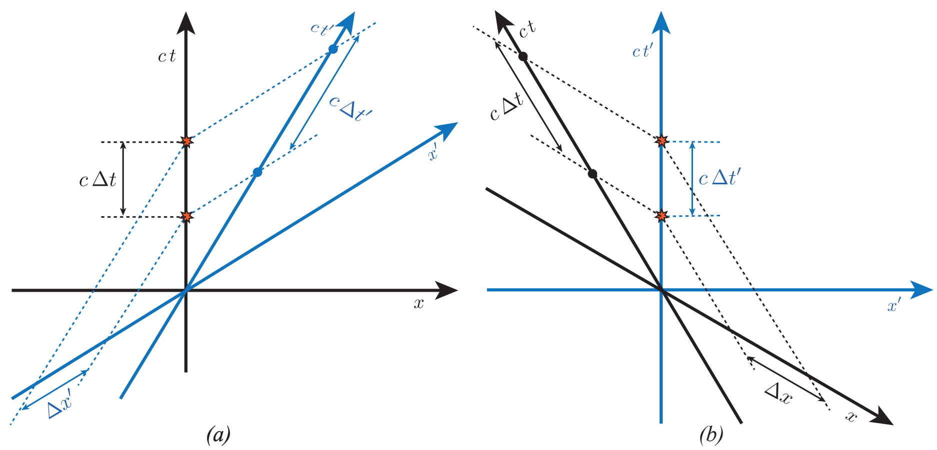


FIG 2.11 : The time dilation phenomenon. (a) Shows the scenario of a clock carried by observer $\mathcal{O}$. (b) shows the case of a clock carried by $\mathcal{O}'$.

in Figure 2.12(b). This time the rest length of the stick is given by $\Delta x = L_0$; and it is $\mathcal{O}'$ who perceives the stick moving, now in the negative x' direction. Once again, we can check that $\mathcal{O}'$ measures a length $\Delta x' = (4/5)L_0 < L_0$. Moving objects are contracted by a factor of γ along the direction of motion. In fact, relativity introduces more elaborate geometric aberrations of moving objects, including a pseudo-rotation effect and preservation of circular shapes. We leave some of the discussion to problems at the end of the chapter.

A crucial ingredient in this analysis is the realization that two events that are *simultaneous* in one reference are not necessarily so in another: this is known as **the relativity of simultaneity**. In the case at hand, the measurement of the locations of the two endpoints of the stick are simultaneous to one observer, and hence constitute a read out of the length of the stick. However, these same two measurements, as shown in the figure, are *not* simultaneous to the other observer, and hence cannot constitute a reading of the length of the stick as measured by this other observer.

EXAMPLE 2-10: The twin paradox

Relativity abounds with so-called "paradoxes" – thought experiments that appear at first to lead to conceptual contradictions. However, they all invariably arise from one of several Newtonian traps. For example, one common pitfall is that of simultaneity: in relativity, two events that are simultaneous in one reference frame are not necessarily so in another. We saw this phenomenon at work in the previous example leading to geometric distortions. Yet,

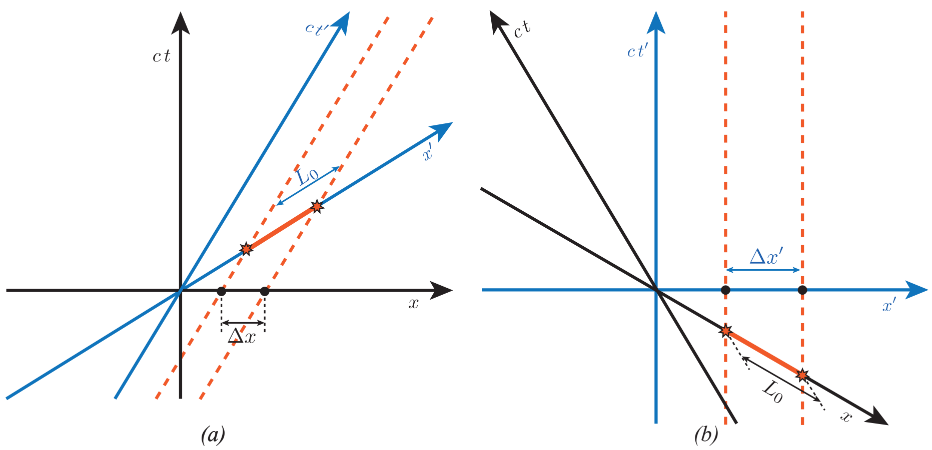


FIG 2.12 : The phenomenon of length contraction. (a) Shows the scenario of a meter stick carried by observer $\mathcal{O}'$. (b) shows the case of a stick carried by $\mathcal{O}$.

– based on our Newtonian daily experiences – we have no intuition for this, because we never encounter it in our normal experience. In general, once relativistic tinkering with the notion of time is taken into account, paradoxes are quickly resolved. And in resolving each paradox, one's intuition for relativity develops a bit more.

In this example we focus on the classic twin paradox. The scenario goes as follows. John lives in Claremont, CA and tracks time with his wristwatch. His twin, Jane, is on a trek to a nearby star a distance D away. Jane will travel along a straight path at constant speed V_0 relative to John, then will turn around and come back with the same constant speed. Figure 2.13(a) shows a Minkowski diagram of the setup with simultaneity lines according to John. If V_0 is large enough, time dilation effects will be at play. There are three segments of the trip, two of which last for a period T_1 to John, as shown in the figure, and the middle segment lasting a period T_0. The total time of the trip will be $T_0 + 2\,T_1$ on John's wristwatch. We want to compare this to the time elapsed on Jane's wristwatch during the same period. We can immediately tell that $T_1 = D/V_0$. However, Jane's clock-rate will necessarily be slow to John because of time dilation since she is moving relative to him. For the first and third segments of the trip, Jane's speed is constant and we simply have

$$T_1 = \gamma_0 \tau_1 \tag{2.121}$$

where τ_1 is the time elapsed on Jane's wristwatch while T_1 has elapsed on John's; and $\gamma_0 = (1 - V_0^2/c^2)^{-1/2}$. Hence, $T_1 > \tau_1$ and John ages more during these segments. The second segment is trickier since Jane is accelerating as she turns around to come back to Claremont. Let us assume, for simplicity, that Jane decelerates at a constant rate during the turnaround. From John's perspective, that is, from the perspective of observers at rest in his inertial frame, he can track what's happening to Jane using the relativistic form of Newton's second law. For constant acceleration, this is a problem we have studied already. We know Jane's trajectory

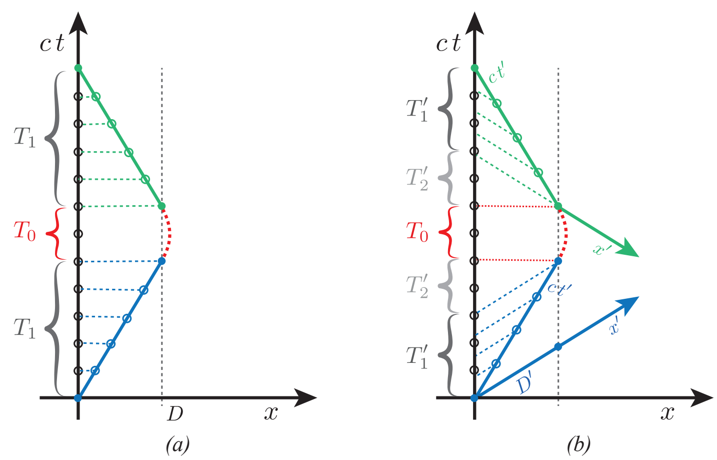


FIG 2.13 : Minkowski diagrams of the twin paradox. (a) shows simultaneity lines according to John. During the first and third part of the trip, a time $2 \times T_1$ elapse on John's clock; during the middle part, Jane is accelerating uniformly and the time elapsed is denoted by T_0(b) Shows simultaneity lines according to Jane, except for the two dotted lines sandwiching the accelerating segment. Jane's x' axis is also shown for two instants in time. The segment labeled T_0 is excised away and borrowed from John's perspective since Jane is a not an inertial frame during this period. T_1' and T_2' on the other hand can be computed from Jane's perspective. Notice how Jane's x' axis must smoothly flip around during the time interval T_0, as she turns around. Her simultaneity lines during this period will hence be distorted and require general relativity to fully unravel.

would be hyperbolic on a Minkowski diagram as shown in Figure 2.13. We also know that her speed will be evolving as

$$v(t) = \frac{a\,t}{\sqrt{1 + a^2 t^2/c^2}} \tag{2.122}$$

where a is some negative constant acceleration and $t = 0$ is taken as the moment when she has zero speed at the midpoint of the trip. Setting $v(T_0/2) = -V_0$, we can immediately deduce that

$$\frac{a\,T_0}{2} = -\gamma_0 V_0 \ , \tag{2.123}$$

where T_0 is the time it takes for Jane to change her speed from V_0 to $-V_0$ according to John, as shown in the figure. How much time passes on Jane's wristwatch during this period? At every instant in time, Jane is moving with some speed $v(t)$ and is subject to a time dilation effect

$$dt = \frac{d\tau}{\sqrt{1 - v(t)^2/c^2}} \ . \tag{2.124}$$

Substituting for $v(t)$ and integrating over the period $T_0/2$, we get

$$\sinh \frac{a\,\tau_0}{2\,c} = \frac{a\,T_0}{2\,c} \ , \tag{2.125}$$

where τ_0 is the time elapsed on Jane's wristwatch during the turnaround. Eliminating a using (2.123), we can then write

$$T_0 = \frac{(\gamma_0 V_0/c)}{\sinh^{-1}(\gamma_0 V_0/c)}\tau_0 > \tau_0 \ . \tag{2.126}$$

We thus have full control over the computation from John's perspective. We can tell that while Jane's wristwatch ticks for a period of

$$\tau_0 + 2\,\tau_1 \tag{2.127}$$

during the full trip, John's clock ticks

$$T_0 + 2\,T_1 = \frac{(\gamma_0 V_0/c)}{\sinh^{-1}(\gamma_0 V_0/c)}\tau_0 + 2\,\gamma_0 \tau_1 \tag{2.128}$$

during the same period. Hence, John ages more, since $T_0 > \tau_0$ and $T_1 > \tau_1$. Let us summarize the result: The travel time

$$\text{on John's watch} = T_0 + 2\,T_1 = \frac{2\,\gamma_0 V_0}{|a|} + \frac{2\,D}{V_0} \ , \tag{2.129}$$

$$\text{on Jane's watch} = \tau_0 + 2\,\tau_1 = \frac{2\,c\,\sinh^{-1}(\gamma_0 V_0/c)}{|a|} + \frac{1}{\gamma_0}\frac{2\,D}{V_0} \tag{2.130}$$

where we used $T_1 = D/V_0$ and $T_0 = 2\,\gamma_0 V_0/|a|$ to quote both results in terms of D (the distance of travel according to John), a (Jane's acceleration according to John), and V_0 (Jane's speed for most of the trip). Notice that for small speeds, $V_0 \ll c$, and the two periods become approximately the same, as expected, since $\sinh^{-1}(\gamma_0 V_0/c) \simeq (\gamma_0 V_0/c)$ and $\gamma_0 \simeq 1$.

Thus, Jane has aged less during the travel! This is fine and interesting until you try to reverse your perspective. From Jane's point of view, she was not moving. Instead, John traveled away while the star visited her! Both John and star traveled past Jane at speed V_0 in the opposite direction, as in watching trees move past you while you are driving a car. According to Jane, is it then John's time that is dilated? Hence, by the time the trek is over and the twins meet, would Jane expect that John has aged less during her travel period? Since John and Jane can now meet and compare notes, only one of the two can be correct, and hence the paradox.

The resolution lies in the realization that Jane is not at rest in any single inertial reference frame throughout the trip, while John *is*. This is because Jane has to decelerate and turn around at the star to come back to Claremont. During the turnaround period, Jane is not an inertial observer, and John and Jane are *not* equivalent observers as far as the laws of physics are concerned. For example, Jane can hold a pendulum and notice that it sways while she is turning around to come back home. To find out the outcome from Jane's perspective, we would need to learn how to handle the point of view of accelerating observers; we need to know how space and time are affected in Jane's reference frame when she is decelerating. This quickly gets us into the territory of Einsteins's **general theory of relativity**, and we want to avoid doing this. Fortunately, we can immediately deduce that John's conclusion must be the correct one since he is indeed inertial: Jane ages less. However, we can still analyze Jane's point of view to a certain satisfactory extent, and we will do so using graphical methods. Figure 2.13(b) shows the same setup with simultaneity lines according to Jane. The middle segment of the trek where Jane is not inertial has been excised: for this period, we still need to rely on John's perspective. We then take as given

$$T_0 = \frac{(\gamma_0 V_0/c)}{\sinh^{-1}(\gamma_0 V_0/c)}\tau_0 > \tau_0 \ . \tag{2.131}$$

The question is now to determine T_1' and T_2' as shown in the figure. T_1' corresponds to the time elapsed on John's wristwatch during the first segment *according to Jane*. Time dilation tells us that it is given by

$$\tau_1 = \gamma_0 T_1' \Rightarrow T_1' = \frac{\tau_1}{\gamma_0} \tag{2.132}$$

since John is doing the moving according to Jane. This makes sense since Jane sees the distance D contracted to $D' = D/\gamma_0$. So, her travel time must be $\tau_1 = (D/\gamma_0)/V_0 = T_1/\gamma_0$ as we found before from John's perspective. To find T_2', we need to look at the figure and do a bit of geometry. The slope of Jane's x' axis on the figure is $\pm V_0/c$. We can then immediately read off

$$c\,T_2' = \frac{V_0}{c} \times D \ . \tag{2.133}$$

Putting things together we find the total time of the trip on John's wristwatch from Jane's perspective

$$\begin{aligned} & \text{Travel time on John's wristwatch} = 2\,T_1' + 2\,T_2' + T_0 \\ = & \frac{2\,\tau_1}{\gamma_0} + \frac{2\,V_0 D}{c^2} + \frac{2\,\gamma_0 V_0}{|a|} = \frac{2\,D}{\gamma_0^2 V_0} + \frac{2\,V_0 D}{c^2} + \frac{2\,\gamma_0 V_0}{|a|} \\ = & \frac{2\,D}{V_0} + \frac{2\,\gamma_0 V_0}{|a|} \end{aligned} \tag{2.134}$$

where we used $\tau_1 = (D/\gamma_0)/V_0$ and $T_0 = (2\,\gamma_0 V_0)/|a|$. We see that the conclusion is identical to John's, equation (2.129): John ages more. From Jane's perspective, we relied on her notion of simultaneity during the first and third segments of the trip (computations of T_1' and T_2'), during the intervals when Jane is an inertial observer. However, we did borrow John's conclusion about his and Jane's clock rates (computation of T_0), since his perspective was the inertial one – a framework where the laws of special relativity can be applied. During this acceleration phase, the laws of physics are altered from Jane's perspective, and to carry the computation of T_0 from her reference frame requires us to learn how special relativity is modified in an accelerated frame. We will see in Chapter 4 that the **principle of equivalence** plays a central role in such settings. However, a full treatment necessitates excursions into the subject of **general relativity** — a fully relativistic theory of gravity. Although we will delve a bit into general relativity in Chapter 10, a full treatment of this beautiful subject goes beyond the scope of our book.

Problems

PROBLEM 2-1: Clock A is placed at the origin of the primed frame; it reads time $t' = 0$ just as the origins of the primed and unprimed frames coincide. At a later time t to observers in the unprimed frame, clock A has moved a distance $x = Vt$. Using the Lorentz transformation, find the reading t' of clock A.

PROBLEM 2-2: A primed frame moves at $V = (3/5)c$ relative to an unprimed frame. Just as the origins pass, clocks at the origins of both frames read zero, and a flashbulb explodes at that point. Later, the flash is seen by observer A at rest in the primed frame, whose position is $(x', y', z') =$ (3 m, 0, 0). (a) What does A's clock read when A sees the flash? (b) When A sees the flash, where is A located according to unprimed observers? (c) To unprimed observers, what do their clocks read when A sees the flash?

PROBLEM 2-3: A stick of length L_0 is placed at rest in the primed frame, along the x' axis. Observers in the unprimed frame measure both ends of the stick at the same time t according to unprimed clocks. Using the Lorentz transformation, find the length $L \equiv (x_2 - x_1)$ of the stick in the unprimed frame, in terms of L_0 and the relative frame velocity V. Here x_1 and x_2 are the end locations in the unprimed frame.

PROBLEM 2-4: Two clocks A and B are placed at rest in the primed frame, both on the x' axis, at x'_A and x'_B. Using the Lorentz transformation, find the difference $t'_A - t'_B$ of these clock readings when they are observed simultaneously in the unprimed frame, both at some time t. The result shows that events which are simultaneous in one frame may not be simultaneous in another frame.

PROBLEM 2-5: Two spaceships are approaching one another. According to observers in our frame, (a) the left-hand ship moves to the right at (4/5)c, and the right-hand ship moves to the left at (3/5)c. How fast is the right-hand ship moving in the frame of the left-hand ship? (b) The left-hand ship moves to the right at speed $(1 - \epsilon)c$ and the right-hand ship moves to the left at $(1 - \epsilon)c$, where ϵ is in the range $0 < \epsilon < 1$. How fast is the right-hand ship moving in the frame of the left-hand ship? Show that this is less than c, no matter how small ϵ is.

PROBLEM 2-6: By differentiating the velocity transformation equations one can obtain transformation laws for acceleration. Find the acceleration transformations for $a_x, a_y,$ and a_z.

PROBLEM 2-7: Al and Bertha are identical twins. When she is 18 yrs old, Bertha travels to a distant star at constant speed (24/25)c, turns quickly around, and returns at the same speed. When she arrives home she is 25 yrs old. (a) How old is Al when she returns? (b) How far away was the star in Al's frame?

PROBLEM 2-8: A particle moves at speed 0.99c. How fast must it move to double its momentum?

PROBLEM 2-9: A photon of momentum p_γ strikes an atomic nucleus at rest, and is absorbed. If the mass of the final (excited) nucleus is M, calculate its velocity.

PROBLEM 2-10: Two particles make a head-on collision, stick together and stop dead. The first particle has mass m and speed $(3/5)c$, and the second has mass M and speed $(4/5)c$. Find M in terms of m.

PROBLEM 2-11: A particle moves in the x, y plane with velocity $v = (4/5)c$, at an angle of 30° to the x axis. (a) Find all four components of the particle's four-vector velocity v_μ and evaluate the invariant square of its components $\eta_{\mu\nu} v_\mu v_\nu$. (b) Find all four components of the particle's four-vector velocity in a frame moving to the right at velocity $V = (3/5)c$. (c) Evaluate explicitly the invariant square of the components in this frame.

PROBLEM 2-12: Prove that the sign of the zeroth component of a timelike four-vector is invariant under Lorentz transformations.

PROBLEM 2-13: (a) An unstable particle of mass m decays in time $\tau = 10^{-10}$ s in its own rest frame. If its energy is $E = 1000mc^2$ in the lab, how far (in meters) will it move before decaying? (b) The kinetic energy of a particular newly-created particle in the laboratory happens to equal its mass energy. If it travels a distance d before decaying, find an expression for how long it lived in its own rest-frame.

PROBLEM 2-14: A photon of energy $E =$ 5000 MeV is absorbed by a nucleus of mass M_0 originally at rest. Afterwards, the excited nucleus has mass M and is moving at speed $(5/13)c$. (a) In units MeV$/c$, find the momentum of M. Then in units MeV$/c^2$, find (b) the mass M (c) the mass M_0.

PROBLEM 2-15: The Large Hadron Collider (LHC) at CERN near Geneva, Switzerland, accelerates protons of mass energy mc^2 to an energy $E \gg mc^2$. (a) In terms of mc^2 and E, write down a simple expression for the γ-factor of LHC protons. (b) The velocity of these protons can be expressed in the form $v/c = 1 - \epsilon$, where $\epsilon \ll 1$. Derive a simple expression for ϵ in terms of mc^2 and E. (c) Each proton has mass energy 938 MeV and may eventually with energy as large as 7 TeV $= 7 \times 10^{12}$ eV. Find ϵ for these protons.

PROBLEM 2-16: A 1.0 kg space probe is ejected from the Moon by a powerful laser that pushes on the probe with the constant force 1000 N. How long, according to Moon clocks, does it take the probe to reach speed $c/2$?

PROBLEM 2-17: Astronomers and cosmologists use the symbol z for the redshift of distant galaxies or quasars, where $z \equiv \Delta\lambda/\lambda = (\lambda_{\rm ob} - \lambda_{\rm em})/\lambda_{\rm em}$. Here $\lambda_{\rm ob}$ is the wavelength of a spectral line as observed on Earth and $\lambda_{\rm em}$ is the wavelength emitted by the distant object in its own rest frame. (a) Show that, in terms of z, the recessional velocity of a distant quasar according to the Doppler formula is given by

$$v/c = \frac{(1+z)^2-1}{(1+z)^2+1}. \tag{2.135}$$

(b) If the Lyman alpha line ($\lambda = 121.6$ nm) emitted by hydrogen atoms in a quasar is observed on Earth to have the wavelength $\lambda = 243.2$ nm, how fast is the quasar moving away from us?

PROBLEM 2-18: A light source glows uniformly in all directions, in a frame at rest relative to the light source. Show that if the source is moving at speed v in our frame, half of the emitted photons are radiated into a forward cone whose half-angle is $\theta = \cos^{-1}(v/c)$. This is called the "headlight effect".

PROBLEM 2-19: Suppose that primed and unprimed inertial frames have relative velocity V in the x direction. Suppose also that the transformation between them has the linear form $t = At' + Bx'$, $x = Ct' + Dx'$, $y = y'$, $z = z'$. Using the meaning of V and the two postulates of relativity, derive the Lorentz transformation by evaluating the four constants A, B, C, D in terms of V.

PROBLEM 2-20: A straight stick is placed at rest in the x', y' plane of the primed frame, at angle θ' to the x' axis. As observed in the unprimed frame, what is the angle of the stick relative to the x axis?

PROBLEM 2-21: Show that the momentum and velocity four-vectors are both timelike, and that the force four-vector is spacelike.

PROBLEM 2-22: We wish to show that the wave equation for light is invariant under Lorentz transformations: *i.e.*, , observers $\mathcal{O}$ and $\mathcal{O}'$ would write the same equation in their respective coordinate systems to describe light propagation. Using the chain rule of partial differentiation, we start by noting

$$\frac{\partial}{\partial t} = \frac{\partial t'}{\partial t}\frac{\partial}{\partial t'} + \frac{\partial x'}{\partial t}\frac{\partial}{\partial x'} + \frac{\partial y'}{\partial t}\frac{\partial}{\partial y'} + \frac{\partial z'}{\partial t}\frac{\partial}{\partial z'} = \gamma\frac{\partial}{\partial t'} - \gamma\beta\, c\frac{\partial}{\partial x'}\ . \tag{2.136}$$

Similarly, we have

$$\frac{\partial}{\partial x} = -\gamma\beta\, c\frac{\partial}{\partial t'} + \gamma\frac{\partial}{\partial x'},\quad \frac{\partial}{\partial y} = \frac{\partial}{\partial y'},\quad \frac{\partial}{\partial z} = \frac{\partial}{\partial z'}\ . \tag{2.137}$$

Using these relations, show that

$$\frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2} = 0 \Rightarrow \frac{\partial^2\phi}{\partial x'^2} + \frac{\partial^2\phi}{\partial y'^2} + \frac{\partial^2\phi}{\partial z'^2} - \frac{1}{c^2}\frac{\partial^2\phi}{\partial t'^2} = 0, \tag{2.138}$$

confirming the invariance of the wave equation under Lorentz transformations. Both observers $\mathcal{O}$ and $\mathcal{O}'$ then write the same wave equation — with the same speed parameter c — despite their relative motion.

PROBLEM 2-23: The wave equation for light is

$$\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} + \frac{\partial^2 \phi}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2 \phi}{\partial t^2} = 0. \tag{2.139}$$

Show that the set of all linear transformations of the spacetime coordinates that permit this wave equation to be written as we did, correspond to (i) four possible translations in space and time, (ii) three constant rotations of space, and (iii) three Lorentz transformations. Collectively, these are called the **Poincaré transformations** of spacetime.

PROBLEM 2-24: Prove that the time order of two events is the same in all inertial frames if and only if they can be connected by a signal traveling at or below speed c.

A Paradox: The train in the tunnel (Thanks to D. C. Petersen)

Cast of characters:

A relativistic commuter train, rest length 300 m, traveling at $(4/5)c$:
a dark tunnel, rest length 400 m;
the trains diligent crew;
some nefarious saboteurs.

The situation:

The train must pass through the tunnel. The saboteurs have decided to blow up the train in the tunnel. They put photodetectors on the top of the train at the front and the back. When either photodetector is in darkness, it sends a laser beam to a third detector on the top of the train, located exactly midway between the front and back of the train. If this detector sees both laser beams simultaneously it sends a signal to detonate a bomb on the train.

The paradox:

As the train is hurtling down the track at $(4/5)c$ toward the tunnel, the engineer is informed of the plan in detail. Not knowing any relativity, he foresees both ends of the train are soon to be in the dark at the same time and panics. Stop the train!

His fireman, who has been reading a little relativity during breaks between shoveling coal into the firebox is even more worried. Its even worse than you think, he tells the engineer. Were moving at $(4/5)c$, so our train is only (300 m)$\sqrt{1-(4/5)^2} =$ 180 m long. Clearly, we'll perish! Put on the brakes!

The brakeman, being somewhat lazy and not wanting to slow the train, has also been reading his relativity. Dont worry, he says. When we go through the tunnel we'll see that the tunnel, which is, after all, moving relative to us, is only (400 m)$\sqrt{1-(4/5)^2} =$ 240 m long. By the time the tail end of the train gets to the tunnel, the front will already be out. The detector won't see both laser beams simultaneously and won't detonate the bomb. In fact the faster we go, the safer we'll be. More coal!

The question: Does the train blow up or not?

PROBLEM 2-25: A particle decays into two particles of unequal mass. (a) Prove that in the rest frame of the initial particle, it is the less-massive final particle that carries away most of the final kinetic energy. (b) Then show that in the nonrelativistic limit, in which mass is nearly conserved, the ratio of the final particle kinetic energies is equal to the inverse ratio of the final particle masses.

PROBLEM 2-26: It is possible to create antiprotons ($\overline{p}$) by the reaction $p{+}p \to p{+}p{+}\overline{p}{+}p$, where one of the initial protons comes from a high-energy accelerator, and the other is at rest in the lab. Find the threshold energy, the minimum energy of the incident proton needed to make the reaction go. It is helpful to first explain why for this energy all of the final particles move in the forward direction at the same velocity.

PROBLEM 2-27: A particle of mass M_0 decays into two particles, with masses m_1 and m_2. (a) In the rest-frame of M_0 find the energy, the kinetic energy, the momentum, and the speed of each particle, in terms of the masses and the speed of light. (b) An Ω^- particle sometimes decays into a Ξ^0 hyperon and a pion, written $\Omega^- \to \Xi^0 + \pi^-$. The mass energies are Ω^-: 1676 MeV; Ξ^0: 1311 MeV; π^-: 140 MeV. Using energy units, in the rest-frame of the Ω^- find the pions (a) total energy (b) kinetic energy (c) momentum (d) velocity, as a multiple of c.

PROBLEM 2-28: K^+ mesons can be photoproduced by the reaction $\gamma + p \to K^+ + \Lambda$ where the initial proton is at rest in the lab. From the conservation laws, discover if it is possible for either the K^+ or the Λ to be at rest in the lab, and for what photon energy (in terms of the particle masses) this could happen. The particle mass-energies are (in MeV) p : 938.3, K^+: 493.7, Λ: 1115.7.

PROBLEM 2-29: The tachyon, a hypothetical particle that travels faster than light, has imaginary energy and momentum using the traditional formulas. (a) Show that these quantities can be made real by assigning tachyons the imaginary mass im, where $i = \sqrt{-1}$ and m is real. How then do the resulting real momentum and energy depend upon velocity? In terms of m and c, what is the quantity $E^2 - p^2c^2$ for tachyons? Now suppose an ordinary particle of mass M at rest decays into an ordinary particle of mass m and an unseen particle that may or may not be a tachyon. Knowing M and m, show how you could tell from measurements of the energy of the final ordinary particle m, whether the unseen particle is a tachyon, a massless particle, or an ordinary massive particle. Assume that energy and momentum are conserved.

PROBLEM 2-30: (a) Prove that two colliding particles cannot transform into a single photon. (b) Explain why a photon that strikes a free electron cannot be absorbed: $\gamma{+}e^- \to e^-$.

PROBLEM 2-31: (a) Show that the components of the net force acting on a particle in the usual primed and unprimed frames are related by

$$F_x = \frac{F_{x'} + V/c^2(\boldsymbol{v'} \cdot \boldsymbol{F'})}{1 + v_{x'}V/c^2}, \quad F_y = \frac{F_{y'}\sqrt{1 - V^2/c^2}}{1 + v_{x'}V/c^2}, \quad F_z = \frac{F_{z'}\sqrt{1 - V^2/c^2}}{1 + v_{x'}V/c^2} \tag{2.140}$$

where $\boldsymbol{v'} \cdot \boldsymbol{F'} = v_{x'}F_{x'} + v_{y'}F_{y'} + v_{z'}F_{z'}$. (b) Then show that if the motion is entirely in the x direction, it follows that $F_x = F_{x'}$, just as it would be if the Galilean transformation were valid.

PROBLEM 2-32: Prove that the four-vector force is spacelike.

PROBLEM 2-33: The J/ψ meson was discovered in the 1970s at Brookhaven National Laboratory and at the Stanford Linear Accelerator Center, using different reactions. SLAC used $e^- + e^+ \to J/\psi$, in which the initial electron and positron had equal but opposite velocities, and each had total energy $E = 1.55$ GeV. (a) What is the mass-energy of the J/ψ meson? (b) What is the velocity of the original electron, expressed in the form $v/c = 1 - \epsilon$? (c) Suppose the SLAC researchers had planned to create the meson by firing a positron at an electron at rest. What must the total energy of the positron have been in that case?

PROBLEM 2-34: A team of particle physicists wants to create a new particle Q in the collision $p + p \to Q$. Each initial proton has mass m, and particle Q has a mass M that is unknown, except one expects that $M \gg m$. Suppose that one or both of the protons can be accelerated up to energy E_0. (a) Show that if both protons achieve energy E_0 in a colliding-beam experiment, the largest mass energy that can be created is $Mc^2 = 2E_0$. That is, in a colliding-beam experiment, Mc^2 increases linearly with E_0. (b) Show that if only one of the protons achieves E_0, while the other is a stationary target, the largest mass energy that can be created is $Mc^2 \cong \sqrt{2mc^2E_0}$. That is, in a stationary-target experiment, Mc^2 increases only as the square root of E_0. (c) If $E_0 = 72mc^2$, find the maximum mass that can be created in each type of experiment.

PROBLEM 2-35: Consider an observer $\mathcal{O}'$ traveling in the usual arrangement at speed V relative to $\mathcal{O}$ along their mutual x' and x axes. The two observers are receeding. $\mathcal{O}'$ shines a flashlight of frequency ν' at an angle θ' to her x' axis. The light ray is seen by $\mathcal{O}$ at an angle θ for his x axis and at frequency ν. Show that

$$\nu = \gamma_V \left(1 - \frac{V}{c}\cos\theta'\right)\nu' \tag{2.141}$$

and

$$\cos\theta = \frac{\cos\theta' - \frac{v}{c}}{1 - \frac{v}{c}\cos\theta'} \ . \tag{2.142}$$

PROBLEM 2-36: In the text, we derived the Doppler formulae for light. Using the same strategy, find the relativistic Doppler formulae for waves traveling at a speed $v \neq c$. For example, the waves may be sound waves in some very very stiff material whose sound speed is a few percent that of c.

PROBLEM 2-37: Ultra high-energy cosmic rays consist primarily of protons that may have originated in far-away active galactic nuclei. As they zip through space they will inevitably encounter low-energy photons in the cosmic background radiation (CBR) that was set loose in the early universe. CBR photons have a wide range of wavelengths, peaked at approximately 1 mm. These photons have energies that are way below the threshold to cause pion photoproduction off protons that are at rest in the CM frame of our galaxy, but can have very high energies in the rest frame of the cosmic-ray protons themselves. If these energies exceed the threshold, pions will be produced and the proton energy in the frame of our galaxy will be reduced, leading to an upper limit in the cosmic-ray proton energies we can observe. This is called the GZK limit (for the physicists Greisen, Kuzmin, and Zatsepin who predicted it). Estimate the GZK limit (in eV) by pretending that the CBR consists entirely of photons with wavelength 1 mm in our galactic frame of reference. *Hint*: Let the unprimed frame be the rest frame of the cosmic-ray proton and the primed frame be the frame of our galaxy. Then show that the gamma factor $\gamma = (1 - V^2/c^2)^{-1/2}$ between these two frames is given by $\gamma = E'_p/m_p c^2 \cong E_\gamma/(2E'_\gamma)$ where E_γ and E'_γ are the photon energies in the unprimed and primed frames. The actual result of GZK was 6×10^{19} eV. Nevertheless, cosmic-ray protons with energies of up to 3×10^{20} eV have apparently been observed. The reason for the discrepancy is unclear.

PROBLEM 2-38: An algebraic expression is said to be **Lorentz covariant** if its form is the same in all inertial frames: the expression differs in two inertial frames $\mathcal{O}$ and $\mathcal{O}'$ only by putting prime marks on the coordinate labels. For example, $A_\mu \eta_{\mu\nu} B_\nu = K$ is a Lorentz covariant expression, where A_μ and B_ν are four-vectors and K a constant. Under the Lorentz transformation, $A_\mu \eta_{\mu\nu} B_\nu = A_{\mu'} \Lambda_{\mu\mu'} \eta_{\mu\nu} B_{\nu'} \Lambda_{\nu\ \nu'} = A_{\mu'} \eta_{\mu'\nu'} B_{\nu'} = K$, where we used $\eta_{\mu'\nu'} = \Lambda_{\mu\mu'} \eta_{\mu\nu} \Lambda_{\nu\ \nu'}$. Because the indices come matched in pairs across a metric factor $\eta_{\mu\nu}$, the expression preserves its structural form. The quantity is also a **Lorentz scalar**: its *value* is unchanged under a Lorentz transformation. Which of the following quantities are Lorentz scalars, given that K is a constant and any quantity with a single superscript is a four-vector? (a) $K A_\mu \eta_{\mu\nu}$ (b) $C_\mu = D_\mu (A_\lambda \eta_{\lambda\nu} B_\nu)$. (c) $K A_\mu \eta_{\mu\nu} B_\lambda \eta_{\lambda\sigma} D_\nu F_\sigma$

PROBLEM 2-39:

Consider a Lorentz covariant expression that is *not* a Lorentz scalar, $C_\lambda = K_\lambda h(A_\mu \eta_{\mu\nu} B_\nu)$, where h is any function of the quantity in parentheses. Here quantities with a single subscript are four-vectors. Under a Lorentz transformation, $A_\mu \eta_{\mu\nu} B_\nu$ is Lorentz covariant and is also a Lorentz scalar. Hence, its form and value are unchanged; which means the function $h(A_\mu \eta_{\mu\nu} B_\nu)$ is unchanged in form or value as well. K_μ on the other hand is a four-vector; this means that it transforms as $K_\mu = \Lambda_{\mu\mu'} K_{\mu'}$. The right-hand side of the equation for C_λ transforms as a four-vector as whole, which implies that C_λ also transforms as a four-vector and observer $\mathcal{O}'$ would write $C_{\lambda'} = K_{\lambda'} h(A_{\mu'} \eta_{\mu'\nu'} B_{\nu'})$. This quantity is said to be a **Lorentz vector** (instead of a scalar) since it transforms as a four-vector: its components change, but through the well-defined prescription for a four-vector. Which of the following quantities are Lorentz vectors, given that K is a Lorentz scalar and any quantity with a single subscript is a Lorentz vector? (a) $K \eta_{\mu\nu}$ (b) $C_\lambda = D_\mu A_\lambda \eta_{\mu\nu} B_\nu$. (c) $K A_\mu \eta_{\mu\nu} B_\lambda \eta_{\lambda\sigma} D_\nu F_\sigma$

PROBLEM 2-40:

The concept of Lorentz covariance is important because it allows one to quickly determine the transformation properties of expressions under change of inertial reference frames. The principle of relativity requires that all laws of physics are unchanged as seen by different inertial observers. Hence, we need to insure that expressions reflecting statements of a law of physics are Lorentz covariant: they retain structural form under Lorentz transformations. A useful application of this comes from the second modified law of dynamics

$$f_\mu = \frac{dp_\mu}{d\tau} \ . \tag{2.143}$$

Forces that we insert on the left hand side of this equation must be Lorentz covariant expressions that transform as four-vectors. This insures that observer $\mathcal{O}'$ would simply write

$$f_{\mu'} = \frac{dp_{\mu'}}{d\tau} \ . \tag{2.144}$$

For example, we could write $f_\mu = K_\mu$ with a constant four-vector K_μ. (a) Is a "relativistic spring law" $f_\mu = -(0, k\boldsymbol{x})$ for some constant k, a Lorentz covariant expression? (b) What about a modified spring law $f_\mu = -Kx_\mu = -k(c\,t, \boldsymbol{x})$? (c) What about Newtonian gravity $\boldsymbol{F} = -(k/r^3)\boldsymbol{r}$?

PROBLEM 2-41: *Leading clocks lag*. Minkowski diagrams of primed and unprimed frames are shown in the figure, corresponding to a relative velocity $V = (3/5)c$. (a) Place two dots on it indicating the spacetime positions of two clocks at rest in the primed frame (and synchronized in that frame) when each reads the same time (say, 1:00 pm). Note that these dots must both lie on a line that is parallel to the x' axis, since all points on such a line have the same value of t'. (b) Show from the diagram that these clocks are not synchronized in the unprimed frame. (c) Then show that leading clocks lag; that is, that the clock with the larger value of x (which is leading the other clock in space as they move together in the unprimed frame) lags the other clock in time. (d) Now place two additional dots on the diagram, indicating the events when two clocks at rest in the unprimed frame (and synchronized in that frame) both read some definite time (say, 1:00 pm). Show from the diagram that these two clocks are not synchronized in the primed frame. (e) Then show that according to observers in the primed frame, the leading clock lags. *Hint*: As seen in the primed frame, is it the clock with the larger value of x' or the smaller value of x' that leads the other clock as they move along together?

PROBLEM 2-42: Show that the most general Lorentz transformation can be written as a four by four matrix $\hat{\boldsymbol{\Lambda}}$ satisfying

$$\hat{\boldsymbol{\Lambda}}^t \cdot \hat{\boldsymbol{\eta}} \cdot \hat{\boldsymbol{\Lambda}} = \hat{\boldsymbol{\eta}} \tag{2.145}$$

and

$$|\hat{\boldsymbol{\Lambda}}| = 1 \ . \tag{2.146}$$

Since a Lorentz transformation is by definition a linear transformation of time and space that preserves the speed of light, you simply need to show that these two properties as necessary and sufficient for this. Note also that reflections get ruled out by the second condition by choice.