Chapter 17: Special relativity. (24 May 2021)

A. Background and summary. 1
B. Relativistic invariants; proper time and the velocity 4-vector. 5
C. Time, space, and frames. 7
D. 4-velocity physics. 10
E. Example 1: the relativistic mirror. 13
F. Example 2: synchrotron radiation 15
G. Example 3: Cerenkov radiation; shock waves. 16
H. Tensor formulation of special relativity. 18

A. Background and summary.

Relativity occupies a unique position in science. It has given the world $E = mc^2$, arguably the only scientific equation that is recognized by the public at large. It is counterintuitive: to an observer in a "laboratory" frame, events happen more slowly and lengths in the direction of motion are shorter in the "rest" frame of a moving object. Yet relativity accurately describes the universe in which we live. For example, about 10^{10} times more muons reach the surface of the earth than expected from classical calculations given their creation in the upper atmosphere about 15 km above the surface of the earth and their 2.2 μs average lifetime. Because time passes much more slowly in their rest frames, far more survive than what classical calculations predict. X-rays from synchrotrons form the basis of advanced spectroscopies currently underlying all aspects of science and technology. And relativistic corrections to the global positioning system (GPS) – equal to (-42) μsec/day for general and (+7) μs/day for special – keep location measurements accurate on earth. If these corrections appear to be minor, try multiplying them by 3×10^8 m/s to obtain a rough estimate of the magnitude of the errors that would accumulate in a day if they were neglected.

In this chapter we examine special relativity, but at a lower level than Jackson in Ch. 11 of his three editions. While the full tensor treatment is more elegant and definitely more powerful, working with 4-vectors and the Lorentz transformation directly provides more physical insight. It is worth noting that in his third edition Jackson converted from the complex-4-vector approach that was commonly used historically to the tensorial formulation, where the 4-vectors are real and the non-Euclidian aspects are encoded in the time-space metric tensor. There are sound reasons for doing this, mainly related to the fact that the wave equation is hyperbolic. These are described in Section D. Our task remains the same: get the math right, then extract the physics from the math. As seen below, special relativity provides some unique opportunities to do this.

Special relativity is typically covered in an electrodynamics course not only through its importance via the wave equation, but also through the speed of light. Poisson's 1818 calculation had the unintended consequence of proving Fresnel's conjecture that light was a wave, not a particle, phenomenon to be correct (even though Einstein brought us back full circle). The fact that light could be described as waves forced physicists to confront covariance from a non-mechanical perspective. This created dilemmas that were only exacerbated when Michelson and Morley measured the speed of light to sufficient accuracy to rule out the existence of an ether, although the new issues raised by this most famous failed experiment in history were not finally resolved until Einstein's work in the early part of the 1900's. Any lingering doubt about the existence of an ether can be eliminated by considering data obtained by LIGO, which shows that the speed of light is constant to within parts in 10^{24}.

A discussion of special relativity begins with coordinate transformations between a fixed "laboratory" inertial system and a second system moving relative to it with a constant velocity $\vec{v}$ (no accelerations). These can be considered as mappings of one system onto another, similar to conformal mapping in the theory of complex variables. By convention, the second system moves parallel to the *x* axis. Quantities in the "lab" system are unprimed, and quantities in the system moving with respect to it are primed. The coordinate systems themselves are referred to as "frames", and can be either moving or fixed. The "rest frame" of an object refers to the frame in which the object is stationary.

The equations of classical mechanics are covariant under the Galilean transformation

$$ct' = ct\,; \tag{17.1a}$$

$$x' = x - vt\,; \tag{17.1b}$$

$$y' = y\,; \tag{17.1c}$$

$$z' = z\,. \tag{17.1d}$$

Possible difficulties that would follow from combining quantities with different units are avoided by working with ct rather than *t*. The distinguishing equation is (17.1a), which states that the time t' in the moving frame is identical to the time *t* in the laboratory frame.

Equations (17.1) worked excellently well for the classical mechanics of the time, when nothing moved at speeds close to *c*. However, it was already well known that the wave equation is not covariant under a Galilean transformation, but required the Lorentz transformation

$$ct' = \gamma(ct - \beta x)\,; \tag{17.2a}$$

$$x' = \gamma(x - \beta ct)\,; \tag{17.2b}$$

$$y' = y\,; \tag{17.2c}$$

$$z' = z\,; \tag{17.2d}$$

where

$$\beta = v/c\,; \tag{17.2e}$$

$$\gamma = \left(1 - \beta^2\right)^{-1/2}. \tag{17.2f}$$

The inverse transformation, from the primed to the unprimed system, is given by the same equations but with the sign of the velocity reversed:

$$ct = \gamma(ct' + \beta x')\,; \tag{17.3a}$$

$$x = \gamma(x' + \beta ct')\,; \tag{17.3b}$$

$$y = y'\,; \tag{17.3c}$$

$$z = z'\,; \tag{17.3d}$$

as is easily verified by direct substitution of Eqs. (17.2) into (17.3) or vice versa.

If light were indeed a wave phenomenon, then the Lorentz transformation was not just a set of abstract mathematical relations but had to be taken seriously. None of the ideas put forward to deal with the conflict between $t = t'$ and Eqs. (17.2) and (17.3) were particularly attractive. Maybe the wave equation really was different in different inertial frames. This was hardly likely, because there was no reason to suppose that my inertial frame is preferable to yours. Alternatively, time might be different in different inertial systems. However, this appeared to bring back the idea of a preferred frame, and in any case, violated the common-sense relation $t = t'$ that works so well in mechanics.

Finally, c might be different in different frames. If this were the case we would simply replace $ct' \rightarrow c't$ in Eq. (17.2a), and life would continue as usual. This seemed to be the most plausible reason, because in every other case known up to that time, the speed of a wave in a medium was not a universal constant but depended on the speed of the medium. Why should light be any different? The experimental test was straightforward: compare values of c measured at different times of the day or seasons of the year, given that the rotational speed of the surface of the Earth is $1.66\ \text{km/h} \cong 460\,\text{m/s}$ at the equator and its orbital speed is approximately $108,000\ \text{km/h} \approx 30,000\ \text{m/s}$ relative to the nominally fixed stars. However, even well past the time that Maxwell showed that light was indeed a wave phenomenon, accuracies for measuring c were not up to the task, and the issue remained unresolved.

Accordingly, the history of relativity is inextricably connected to the history of our knowledge of the speed of light. In 1676 Rømer estimated a value $c = 2.2 \times 10^8$ m/s from observing eclipses of the moons of Jupiter. By 1848 measurements by Fizeau had improved this to $c = 3.14 \times 10^8$ m/s. Sadly, this number lasted only two years; it was superceded in 1850 by Foucault's improved estimate of 2.98×10^8 m/s. In 1856 Weber and Kohlrausch measured the ratio of unit electrostatic and electromagnetic charges, finding that this ratio, in modern notation, is $1/\sqrt{\mu_o \varepsilon_o} = 3.107 \times 10^8$ m/s.

The Weber-Kohlrausch result was the first indication that light might be related to electrodynamics. However, the significance of this result was not recognized until Maxwell's groundbreaking work published in 1861. Maxwell summarized what was then known about electrodynamics in 20 equations, including the wave equation (the Maxwell's Equations that we know did not reach their form until 1884.) Maxwell's work firmly established light as an electromagnetic phenomenon, and hence the covariance issue raised by the wave equation could no longer be dismissed. Even so, c could not be measured accurately enough to determine whether an ether was required or not.

The situation changed dramatically in 1887, when Michelson and Morley obtained values of c accurate enough so the ether model could be tested critically. These workers expected to find approximately 0.40-fringe difference in their data, but instead found the difference to be less than 0.01 fringe, with a maximum probable uncertainty of ± 0.1 fringe. The ramifications were fundamental, and not surprisingly, it took considerable time for all the implications to sink in. Alternative explanations that were proposed included "frame dragging", where the ether was dragged along by the relatively large mass of the Earth. The issue was not fully resolved until Einstein's work in the early part of the 20th century.

Einstein's great insight was to realize that the proposition that the wave equation is the same in all inertial systems is far more plausible than to assume that a fundamental-reference coordinate system exists somewhere in the universe. And because *c* is an integral part of the wave equation, that meant that *c* itself had to be a universal constant, the same in any inertial system. This meant in turn that our perception of other systems, rather than the systems themselves, had to change. In particular, the Galilean perception of time and space had to be discarded. We have already encountered some surprising aspects of perception in the chapter on radiation, where the volume element d^3r' of the Green-function integral had to be modified to take account the motion of the source relative to the observer, although in this case the calculation was done only to order c^{-1}, i.e., entirely in the Galilean frame. But following Einstein's insight, attention was now focused on 4-vectors, the mathematical constructions that, like the wave equation, are also invariant in any inertial system. The correct math leads to other predictions, the most famous being $E = mc^2$.

Although we now take for granted that *c* is a universal constant, it is worth noting the test of this proposition by LIGO, which can be considered the Michelson-Morley experiment on steroids. This instrument is capable of measuring path-length differences of $\Delta d \approx 2\times10^{-18}$ m in the ca. 280 round trips of its 4 km arms. Converting this sensitivity to the uncertainty in the variation of the speed of light with time of day or season of the year yields a relative percentage difference of about 1 part in 10^{24}, or an uncertainty Δc of *c* of about 3×10^{-16} m/s.

Thus we can state with complete confidence that in our universe, at least, *c* is truly a constant, one that plays the role of infinity in the sense that anything added to infinity is still infinity. Already in 1983 the International Committee on Weights and Measures formalized this position by setting $c \equiv 299,782,458 \text{ m/s}$ *by definition*. While it appears strange to define the value of a measured quantity, recall that our inability to measure length to anywhere near the required accuracy means that it is more logical to express length in terms of the distance light travels in an interval of time, especially when time can now be measured to precisions of parts in 10^{18}.

In Sec. B we continue with a discussion of relativistic invariants. Proper time and the 4-velocity $(\gamma c, \gamma\vec{v})$ are particularly important, and are derived from basics. In Sec. C a detour explores counterintuitive aspects of coordinate transformations, particularly time. In Sec. D the 4-velocity is used as the basis for the momentum and current 4-vectors $(E/c, \vec{p})$ and $(c\rho, \vec{J})$, respectively. The former leads to the observed increase of mass with speed, $E = mc^2$, the Klein-Gordon Equation, and the dispersion relation for empty space. The latter yields the proof that $(\phi, \vec{A})$ is the 4-potential.

In later sections we apply these results to the relativistic mirror and to synchrotron radiation. These two applications are related, and the mirror is a good illustration of a situation where mathematics and physics go in different directions. A treatment of Cerenkov radiation is covered in Sec. G, even though shock waves are a consequence of retardation physics, not special relativity. Finally, in Sec. H we provide an introduction to the tensor formulation of special relativity, along with an explanation of why Jackson went to this approach in the third edition even though he used the complex-algebra version in the first and second editions. This section will be expanded in a future version of this chapter.

B. Relativistic invariants; proper time and the velocity 4-vector.

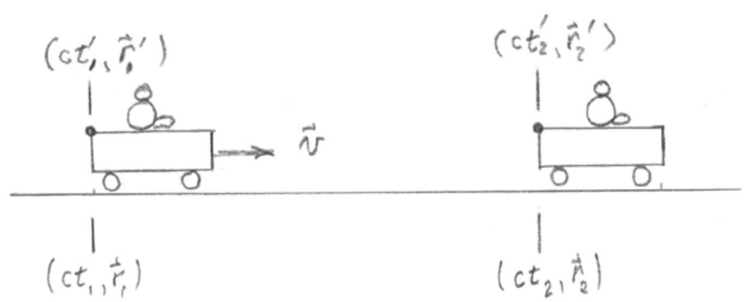
The Lorentz transformation is given in Eqs. (17.2) and (17.3). Given that c is a universal constant and that the wave equation is the same in any inertial system, it makes sense to ask whether there is anything else that is also invariant under the Lorentz transformation, and then anchor the physics to it. As it turns out, there is a group of mathematical constructs whose 4-lengths, defined below, are invariant under a Lorentz transformation. These constructs are termed 4-vectors, the prototypical example of which is the *position 4-vector* $(ct, \vec{r})$, in the standard notation of special relativity. While ct and $\vec{r}$ are not themselves relativistically invariant, the following function of ct and $\vec{r}$ is. This function is

$$s_4^2 = c^2t^2 - \vec{r}^{\,2} = c^2t'^2 - \vec{r}\,'^2 , \tag{17.4}$$

where s_4^2 is the square of the 4-length s_4. It is this invariance that identifies $(ct, \vec{r})$ as a 4-vector. The proof of Eq. (17.4) is easily done with a back-of-the-envelope calculation involving Eqs. (17.2a) and (17.2b), so we do not do it here.

Accordingly, a major goal is to identify other quantities whose 4-lengths are also relativistically invariant. Once the length of a 4-vector has been determined in one inertial system, it is the same in all inertial systems. This not only simplifies calculations, but also provides a way to introduce new physics. We shall see in the next section that a combination of the relativistic 4-velocity with other relativistic invariants such as the rest mass of an object leads to the description of a wide variety of physical phenomena, in addition to yielding, in two lines, Einstein's famous equation $E = mc^2$.

We start with the difference between two position 4-vectors. Another back-of-the-envelope calculation shows, not surprisingly, that this difference is also a 4-vector. We take advantage of this as follows. Consider a cart moving with a velocity $\vec{v} = v\hat{x}$ relative to a laboratory frame, as shown in the diagram. An observer in the lab frame observes the cart passing point $\vec{r}_1$ at time t_1 then point $\vec{r}_2$ at t_2. What does special relativity say about the observer's perception of the time and space experienced by a second observer riding the cart, for which the respective coordinates are t_1', $\vec{r}_1'$, t_2', and $\vec{r}_2'$?



Taking advantage of the fact that the difference of a 4-vector is also a 4-vector, we write for the laboratory frame

$$(ct_2 - ct_1, \vec{r}_2 - \vec{r}_1) = (c\Delta t, \Delta\vec{r}) , \tag{17.5}$$

and for the moving frame

$$(ct'_2 - ct'_1, \vec{r}\,'_2 - \vec{r}\,'_1) = (c\Delta t', \Delta\vec{r}\,') . \tag{17.6}$$

Because these are 4-vectors, their 4-lengths are relativistically invariant. Because they describe the same situation,

$$c^2\Delta t'^2 - \Delta\vec{r}'^2 = c^2\Delta t^2 - \Delta\vec{r}^2 . \tag{17.7}$$

We see, surprisingly, that the time increments Δt and $\Delta t'$ are not identical. Moreover, $\Delta t'$ has its smallest value if $\Delta\vec{r}' = 0$, that is, if the observer in the moving frame is sitting on the cart as it travels from position 1 to position 2, as drawn in the figure. This particular frame is termed the *rest frame* of an object. From the perspective of the observer in the lab, time moves slower for the observer riding the cart. Because an object has only one rest frame, the rest frame of any object is unique. Hence the minimum value $\Delta t'_{\min}$ is also unique, independent of any observer in any other frame. This relativistic invariant $t'_{\min} = \tau$ is called the *proper time*, and the interval equivalent is denoted as

$$\Delta t'_{\min} = \Delta\tau . \tag{17.8}$$

It can be appreciated that Eq. (17.7) is reciprocal: if the object is stationary in the lab frame, i.e., if $\Delta\vec{r} = 0$, then from the perspective of an observer at rest in the moving frame, $\Delta t^2 < \Delta t'^2$. We return to the perception issue below.

The connection between Δt and $\Delta\tau$ follows from Eqs. (17.7) and (17.8). We set $\Delta r'^2 = 0$, replace $\Delta t'^2 = \Delta\tau^2$, and recognize that $\Delta\vec{r}/\Delta t = \vec{v}_o = \vec{v}$ is the velocity of the cart in the laboratory frame, since it is stationary in the moving frame. The result is

$$\Delta\tau^2 = (c^2 - \vec{v}^2)\Delta t^2 , \tag{17.9}$$

or

$$\Delta t = \gamma\,\Delta\tau . \tag{17.10}$$

The connection between the proper time and the time in any other inertial system moving with a velocity $\vec{v}$ is now established.

With $\Delta\tau$ now identified as a system invariant, we next consider

$$\Delta s^2 = c^2\Delta t^2 - \Delta\vec{r}^2 . \tag{17.11}$$

In principle we can construct a 4-velocity by dividing $(c\Delta t, \Delta\vec{r})$ by Δt. However, this 4-length would not be a constant because, as we have seen above, Δt is not a relativistic invariant. Thus the associated construct is not a 4-vector. However, we can obtain a relativistically invariant result by dividing Δs by $\Delta\tau$. The result is the *velocity 4-vector*

$$\frac{\Delta s}{\Delta\tau} = \left(c\frac{\Delta t}{\Delta\tau}, \frac{\Delta\vec{r}}{\Delta t}\frac{\Delta t}{\Delta\tau}\right) = (\gamma c, \gamma\vec{v}) , \tag{17.12}$$

where we have taken advantage of Eq. (17.10). As required of a 4-vector, the 4-length of $(\gamma c, \gamma\vec{v})$ is a constant:

$$(\gamma c)^2 - (\gamma\vec{v})^2 = \gamma^2(c^2 - v^2) \tag{17.13a}$$

$$= c^2 . \tag{17.13b}$$

From a purely mathematical perspective we could have written Eq. (17.12) immediately by inspection, as we did with the radiation Green function in Ch. 15. However, the above derivation is important because it highlights the underlying physics.

C. Time, space, and frames.

The velocity 4-vector is the key to a very large bank. The entire next section is based on it. However, before getting into details, it is useful to gain insight into aspects of special relativity that are nonintuitive but completely verified experimentally and can be understood through Eqs. (17.2) and (17.3) together with the concept of frames. Examples to which the theory is applied include muon excess and GPS. In the former case, the number of muons created by cosmic-ray interactions in the upper atmosphere and reaching the surface of the earth is too large by a factor of about 10^{10}. In the latter case, the correction resulting from satellite motion is about 7 μs/day. Neither result can be explained classically.

Both applications involve time dilation and length contraction. Starting with time dilation, we showed in the previous section that time passes slowest for an object in the frame in which it is at rest. Because the applications both involve Eqs. (17.2) and (17.3), we first show that this conclusion follows from these equations as well. While these calculations are often done using differences, for example $\Delta x = x_2 - x_1$, we simplify the mathematics by assuming that everything is referenced to $x = x' = 0$ and $t = t' = 0$.

We start with the primed system. The definition of an object at rest at the origin in this system is $x' = 0$. Making this substitution in Eq. (17.3a), we find immediately that

$$ct = \gamma ct' = \frac{ct'}{\sqrt{1 - v^2/c^2}} \geq ct'. \tag{17.14}$$

Hence Eqs. (17.3) also show that time advances slowest in the rest system. The same result follows from Eqs. (17.2): setting $x' = 0$ in Eq. (17.2b) yields $x = vt$, as expected, which when substituted into Eq. (17.2a) yields Eq. (17.14).

Next, it is well known that lengths appear shortest in rest frames. To prove this proposition we obtain the relation connecting x and x', using the moving-frame condition that $t' = 0$. This condition ensures that all length measurements in the moving frame are taken simultaneously. Substituting $t' = 0$ in Eq. (17.3b) yields immediately

$$x = \gamma x' = \frac{x'}{\sqrt{1 - v^2/c^2}} \geq x', \tag{17.15}$$

proving the proposition. Again, the same result can be obtained from Eqs. (17.2): substituting $t' = 0$ in Eq. (17.2a) forces $ct = \beta x$, which when substituted into Eq. (17.2b) yields Eq. (17.15).

We now consider the muon. The average lifetime of a muon in its rest frame is about 2.2 μs. Muons created by cosmic rays travel at a speed of about 0.99*c*, so the average distance traveled by a muon after creation is about 650 m. Consequently, the chances of a muon making it the ~15 km from where it was created in the upper atmosphere to the surface of the earth is effectively zero. Yet a significant number survive the trip.

To understand this result, we examine the situation from two perspectives: first, from our rest frame, and second, from the rest frame of the muon. To make the calculation simple, we assume $\gamma \simeq 10$, corresponding to $v/c = 0.995$ (the actual number is $\gamma = 9.14$.) From the lab perspective, Eq. (17.14) shows that time in the rest frame of the muon is slowed down by a factor of 10. Thus from our perspective the muon survives about 22 μs on the average. At the speed of light the average distance that the muon can travel is therefore more than 6 km, which is more than enough for a substantial fraction to make their way to the surface of the Earth.

Next, consider the situation from the perspective of the muon. According to Eq. (17.15), distances in the lab scale are reduced by the same factor of 10. Thus in the 2.2 μs average lifetime of the muon, it can travel more than 6 km of laboratory lengths. Thus the same conclusion is obtained: a significant fraction of the muons generated in the upper atmosphere make it to the surface of the earth, whether we view the situation from the perspective of someone on the earth, or the muon traveling at highly relativistic speeds.

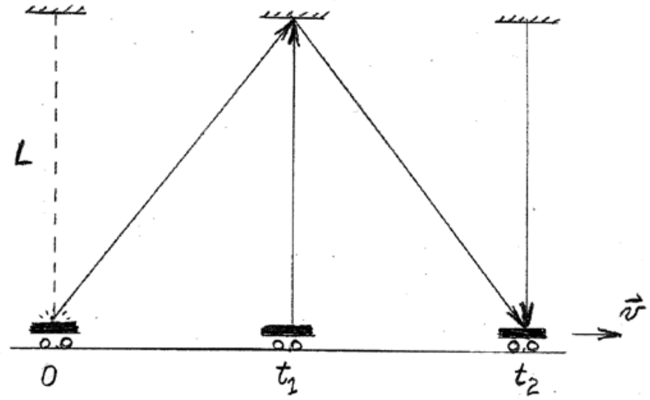
Lest this be viewed as an interesting but academic exercise, we now consider GPS. Because the satellites in orbit are moving relative to the surface of the earth, time is evolving more slowly in their rest frame. The amount can be calculated and is found that this effect causes the atomic clocks in the satellites to lose about 7 μs/day relative to the same clocks on the surface of the earth. To accommodate this effect and maintain synchronization with the atomic clocks in our frame on earth, the clocks in satellites are sped up by 7 μs/day so they can remain synchronized with our clocks on the ground. And the effect is cumulative – as the satellite continues in orbit, an uncorrected clock would continue to lose time relative to the same clock on the surface of the earth.

However, this is only part of the GPS story. Being deeper in the gravitational well, the clocks on the surface of the earth experience a stronger gravitational pull, which according to general relativity equals a stronger acceleration. This relative difference in accelerations causes the clocks on the surface of the earth to run slower than those in the satellites by about 42 μs/day. When both corrections are taken into account, the net result is that the clocks in satellites are programmed to lose 35 μs/day relative to what it would be accurate in the absence of motion and acceleration. Lest you feel that this correction is insignificant, multiply 35 μs by 3×10^{8} m/s to find out how much error would accumulate in a day if these corrections were not taken into account.

In the above discussions, the different frames describe different but quite real universes, with the Lorentz transformation giving the mapping of one universe onto another. Consequently, in relativistic calculations, the first step is to carefully define the frames. You may have noticed that the combination of the effect of special and general relativity ensures that you age 35 μs less per day than the GPS satellites overhead. With respect to the twin paradox, if the frames are properly defined the universes are also properly defined, and the paradox does not exist.

While these results are superficially similar to the perception-induced change of length $\left(1-\hat{R}\cdot\vec{v}/c\right)$ of the volume differential in the direction of motion encountered in Ch. 15, the time and distance dilation results here are of second order in (v/c), and hence do not reverse sign when the direction of motion is reversed. From another perspective, the calculation in Ch. 15 uses the same clock for both moving and laboratory frames. Finally, the retardation physics discussed in Ch. 15 is described by the Galilean transformation. Nevertheless, the two are connected in that the change of retardation length is described by the linear part of Eq. (17.2b).

A common thought experiment arrives at Eqs. (17.14) and (17.15) in a different manner. Suppose that the cart in the figure on the right carries a flashbulb and a mirror mounted vertically at a distance L above the cart as shown in the next figure. At $t = t' = 0$ the flashbulb fires, sending light in all directions. An observer riding on the cart sees the ray that travels straight up and is reflected straight down. An observer standing off to the side sees a different ray going up at an angle, reflecting off the mirror, then reaching the observer after traveling back at the angle of reflection.



Both observers use the clocks in their respective inertial frames to determine the elapsed times of their respective rays making the trip. With c being the same for both observers, the observer on the car calculates

$$ct' = 2L . \qquad (17.16)$$

The observer off to the side calculates

$$ct = \sqrt{(2L)^2 + (vt)^2} . \qquad (17.17)$$

Since $2L$ is a common reference for both frames, it can be eliminated from the two equations and determine how t and t' are related. The calculation shows

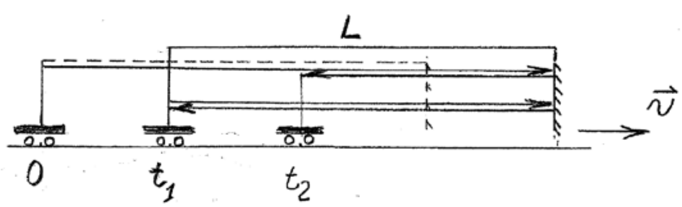
$$c^2t'^2 = (2L)^2 = c^2t^2 - v^2t^2 , \qquad (17.18)$$

or

$$ct' = \sqrt{1-\frac{v^2}{c^2}}\, ct = \frac{ct}{\gamma} , \qquad (17.19)$$

the same as that given in Eq. (17.14). Thus the only way that the lab observer can make sense of the situation is to conclude that time is progressing slower in the rest frame of the car.

Now the observers repeat the experiment, but with the mirror hung at a distance L in front of the car, as shown in the diagram on the right. Both observers now watch the same ray travel first to the mirror then back to the moving observer.



For the moving observer the elapsed time is again given by Eq. (17.16). The calculation for the fixed observer is more complicated. Because the car is moving in the same direction of the ray, it covers a distance

$$ct_1 = L + vt_1 \tag{17.20}$$

by the time the ray reaches the mirror. On the return, the path is shorter. The distance is now

$$ct_2 = L - vt_2 . \tag{17.21}$$

Thus for the lab observer the elapsed time is

$$t_1 + t_2 = t = \frac{L}{c+v} + \frac{L}{c-v}$$

$$= \frac{2cL}{c^2 - v^2} . \tag{17.22}$$

Taking the ratios as written yields

$$\frac{t'}{t} = 1 - \frac{v^2}{c^2} = \frac{1}{\gamma^2} . \tag{17.23}$$

But this is inconsistent with Eq. (17.19), where the ratio is the square root of the ratio given in Eq. (17.23). The only way that the observer at the side can make sense out of the situation is to conclude that in the moving system the perceived length of the distance to the mirror is less than its nominal length, or

$$L' = \sqrt{1 - \frac{v^2}{c^2}} L = \frac{L}{\gamma} . \tag{17.24}$$

The result agrees with Eq. (17.15).

D. 4-velocity physics.

We now pull physics out of the math. First, if a frame is moving at a relative speed $v = c$, Eq. (17.19) shows that $\Delta t = 0$, that is, time is frozen in that frame: it cannot advance. That is why the observation of the neutrino deficit from the Sun was so important: if neutrinos can change character in the ca. 8 min transit time from the Sun to the earth, as the data show, then they cannot be moving at the speed of light. Therefore, they must have mass. The quantitative connection will be established in the next paragraphs.

Putting the velocity 4-vector to work, the rest mass m_o of an object is another obvious relativistic invariant, since a rest frame is a rest frame independent of the inertial frame of an observer. Consequently, multiplying Eq. (17.12) by m_o generates another relativistic invariant, the *momentum 4-vector*:

$$(\gamma c, \gamma \vec{v}) m_o = (\gamma m_o c, \gamma m_o \vec{v}) = (mc, m\vec{v}) . \tag{17.25}$$

The vector part of Eq. (17.25) is clearly the momentum $\vec{p} = m\vec{v}$ as seen in the laboratory frame. But to the laboratory-frame observer, the mass of the object is no longer m_o but has increased to $m = \gamma m_o$ as a result of its motion. This is another well-known consequence of special relativity. To interpret the scalar part, expand γ for small v/c. The result is

$$\gamma\, m_o c = \frac{m_o c}{\sqrt{1-(v/c)^2}} \tag{17.26a}$$

$$= \frac{1}{c}\left(m_o c^2 + \frac{1}{2} m_o v^2 + ... \right) = \frac{E}{c}\,. \tag{17.26b}$$

We immediately recognize the second term in the expansion as the Newtonian kinetic energy of m_o. Thus the scalar part of the momentum 4-vector is the energy E of m_o divided by c.

However, the leading term is new. It informs us that the energy of m_o is not simply its kinetic energy, but that the object has an intrinsic energy $m_o c^2$ as a result of its mass. Einstein's famous equation is obtained.

The Lorentz transformation for energy and momentum is therefore

$$\frac{E'}{c} = \gamma\left(\frac{E}{c} - p_x c \right); \tag{17.27a}$$

$$p_x' = \gamma\left(p_x - \beta \frac{E}{c} \right); \tag{17.27b}$$

$$p_y' = p_y; \quad p_z' = p_z\,. \tag{17.27c,d}$$

The length of the associated 4-vector is

$$\frac{E^2}{c^2} - \vec{p}^{\,2} = m_o^2 c^2\,. \tag{17.28}$$

This is the energy-momentum relation of special relativity.

Next, substituting the quantum mechanical operators

$$E = -i\hbar \frac{\partial}{\partial t}; \quad \vec{p} = i\hbar\nabla \tag{17.29a,b}$$

in Eq. (17.28) yields the Klein-Gordon equation

$$\hbar^2\left(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} \right)\psi = \frac{1}{c^2}(mc^2)^2\psi\,. \tag{17.30}$$

This equation is unphysical because it leads to negative probabilities, a consequence of its quadratic nature. However, this inconsistency was fixed by Dirac, who introduced noncommuting 4×4 matrices to write Eq. (17.30) in factored form. The elements of the associated 4-element column vector represent electrons and holes with spins-up and spins-down.

Returning to the neutrino mass, note first that the momentum part of Eq. (17.25) is

$$\vec{p} = \gamma m_o c = \frac{m_o c}{\sqrt{1-(v/c)^2}}\,. \tag{17.31}$$

If $v = c$, then if the momentum of a particle is to be finite, it *must* have zero mass. Next, if $x = vt = ct$, then Eq. (17.2a) shows that t' remains frozen at zero. Hence if the neutrino truly has zero mass, then it could not evolve among the different versions. Therefore, neutrinos have mass. At present none of these masses are known, although data show that the masses of the electron, mu, and tau neutrinos are different. The only reasonably certain aspect is an upper limit of the order of 10^{-6} that of the electron mass.

Pushing the zero-mass situation further, substitute the quantum-mechanical relations $E = \hbar\omega$ and $\vec{p} = \hbar\vec{k}$ in Eq. (17.28) with $m_o = 0$. The result is

$$\frac{\hbar^2\omega^2}{c^2} - \hbar^2 k^2 = 0, \tag{17.32a}$$

or

$$\frac{c^2k^2}{\omega^2} = 1. \tag{17.32b}$$

This is the dispersion relation of photons in empty space. Photons are therefore described by 4-vectors $(\omega, c\vec{k})$. Thus photons satisfy the Lorentz transformation, specifically

$$\omega' = \gamma\left(\omega - \beta c k_x\right); \tag{17.33a}$$

$$ck'_x = \gamma\left(ck_x - \beta\omega\right); \tag{17.33b}$$

$$k_y' = k_y; \quad k_z' = k_z; \tag{17.33c,d}$$

and, with the appropriate substitutions, the inverse relations Eqs. (17.3). In particular, Eq. (17.33a) gives the relativistic Doppler shift, and Eqs. (17.3) describe synchrotron radiation. The former will be discussed in the next section, and synchrotron radiation in the next after that.

What else can we learn? Another relativistic invariant is the charge density ρ_o in its rest frame. Repeating the above development for ρ_o, the associated 4-vector is

$$\rho_o(\gamma c, \gamma\vec{v}) = (\gamma\rho_o c, \gamma\rho_o\vec{v}) = (\rho, \vec{J}). \tag{17.34}$$

Thus $\rho = \gamma\rho_o$ increases with speed, which is partially the result of shortening the dimension in the direction of motion discussed above. At the same time $\vec{J} = \rho\vec{v}$, so $(c\rho, \vec{J})$ is a 4-vector. We have already shown that

$$\phi = \int_V d^3r' \frac{\rho(r')}{|\vec{r} - \vec{r}'|} = \frac{1}{c}\int_V d^3r' \frac{c\rho(r')}{|\vec{r} - \vec{r}'|}, \tag{17.35a}$$

$$\vec{A} = \frac{1}{c}\int_V d^3r' \frac{\vec{J}(r')}{|\vec{r} - \vec{r}'|}. \tag{17.35b}$$

Therefore, $(\phi, \vec{A})$ is a 4-vector as well. Its Lorentz transformation is

$$\phi' = \gamma(\phi - \beta A_x); \tag{17.36a}$$

$$A_x' = \gamma(A_x - \beta\phi); \tag{17.36b}$$

$$A_y' = A_y; \quad A_z' = A_z\,, \tag{17.36c,d}$$

with the inverse transformation given by Eqs. (17.3) Its length in the radiation zone, where $\phi = \hat{k}\cdot\vec{A}$, is

$$s_4 = \phi^2 - \vec{A}^2 = \vec{A}_\perp^2 \; s_4 = \phi^2 - \vec{A}^2 = -\vec{A}_\perp^2\,. \tag{17.37}$$

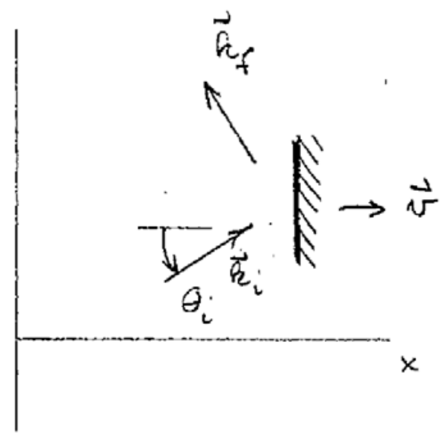
Operating on $(\phi, \vec{A})$ with the 4-divergence $(-\frac{1}{c}\frac{\partial}{\partial t}, \nabla\cdot)$ yields the definition of the Lorentz gauge,

$$\nabla\cdot\vec{A} + \frac{1}{c}\frac{\partial\phi}{\partial t} = 0\,, \tag{17.38}$$

another Lorentz invariant. The proof that this is an invariant is more difficult than the proof for the wave equation, because the operator $(-\frac{\partial}{c\partial t}, \nabla\cdot)$ cannot be separated from the operand $(\phi, \vec{A})$. The proof that Eq. (17.39) is relativistically invariant will be left as a homework assignment.

E. Example 1: the relativistic mirror.

An illustrative application involving Lorentz transformations is the mirror moving at relativistic speeds. The calculation is straightforward, but requires careful bookkeeping. It also has a built-in cross-check in that the result must satisfy the dispersion equation.



Suppose a plane wave of frequency ω and wave vector $\vec{k} = k(\hat{x}\cos\theta_i + \hat{y}\sin\theta_i)$ is reflected from a mirror moving at a relativistic velocity $\vec{v} = v\hat{x}$, where θ is measured with respect to the velocity of the mirror, as shown. The objectives are to determine the frequency ω'' of the reflected wave in the laboratory frame, and the angle θ'' between the wave vector $\vec{k}''$ of the reflected wave and the *x* axis.

The first step is to Lorentz-transform the incoming wave to the rest frame of the mirror, since we know exactly how to describe reflectance in this frame: we simply replace k'_x with $k'_{rx} = -k'_x$. Accordingly, in the laboratory frame start with ω and

$$c\vec{k} = \hat{x}ck_x + \hat{y}ck_y = ck(\hat{x}\cos\theta + \hat{y}\sin\theta)$$
$$= \omega(\hat{x}\cos\theta + \hat{y}\sin\theta)\,, \tag{17.39}$$

noting that $ck = \omega$. Transforming the above into the rest frame of the mirror yields

$$\omega' = \gamma\left(\omega - \beta c k_x\right) = \omega\gamma(1-\beta\cos\theta)\,; \tag{17.40a}$$

$$ck'_x = \gamma\left(ck_x - \beta\omega\right) = \omega\gamma\left(\cos\theta - \beta\right). \tag{17.40b}$$

The y component $ck_y = ck'_y = \omega\sin\theta$ of $\vec{k}$ is of course unchanged. As an intermediate cross-check we leave as an exercise the proof that

$$\omega'^2 - c^2 k'^2_x - c^2 k_y^2 = 0\,. \tag{17.41}$$

Reflection leaves ω' and ck'_y unchanged, but converts $k'_x \to -k'_x = k'_{xr}$. Thus after reflection

$$\omega' = \omega\gamma\left(1-\beta\cos\theta\right); \tag{17.42b}$$

$$ck'_{rx} = -\omega\gamma\left(\cos\theta - \beta\right). \tag{17.42b}$$

Returning to the laboratory frame, we use Eqs. (17.3). After some algebra

$$\omega'' = \omega\gamma^2(1+\beta^2 - 2\beta\cos\theta)\,; \tag{17.43a}$$

$$ck''_x = \omega\gamma^2\left((1+\beta^2)\cos\theta - 2\beta\right); \tag{17.43b}$$

$$ck''_y = \omega\sin\theta\,. \tag{17.43c}$$

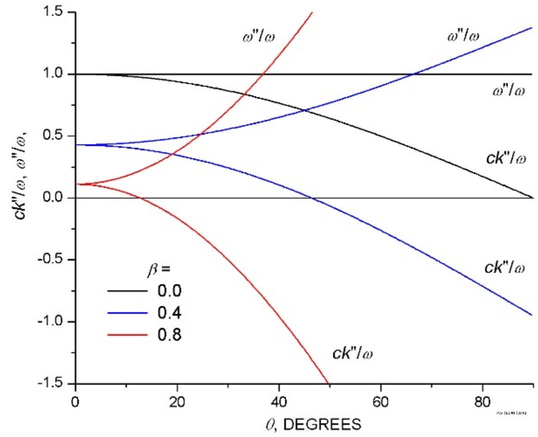
The Doppler-shifted frequency is given by Eq. (17.43a), and the emerging angle by the combination of Eqs. (17.43b) and (17.43c). Again, we leave as an exercise to show that $\omega'' - c\vec{k}''^2 = 0$.

At this stage the mathematics collides with physics. While the mathematics passes all tests in both frames, reflectance can occur only if the wave is propagating toward the mirror, not away from it. This basic physical principle separates the physics from the math. Returning to Eq. (17.42b), this requires

$$\cos\theta > \beta\,. \tag{17.44}$$

For greater speeds in the reference frame of the mirror, the “incoming” wave is propagating backward. While this perspectives issue is certainly legitimate and Eqs. (17.43) are mathematically correct, if $\cos\theta < \beta$ Eqs. (17.43) have no physical meaning. At the break point $\cos\theta = \beta$ the incoming wave is propagating parallel to the mirror. It is easily verified that in this situation the phase velocity ω'/k'_y remains equal to c, as it must, and that Eq. (17.43a) reduces to $\omega'' = \omega$, again as it must.

This is illustrated in the figure at the top of the next page, which shows ω''/ω and ck''_x/ω for $\beta = 0$, 0.4, and 0.8 as a function of θ from 0 to 90°. For $\beta = 0$ the mirror is not moving, the relative back-reflected frequency ω''/ω is unchanged, and the x-axis projection of $c\vec{k}''/\omega$ is simply $\cos\theta$. For the moving mirrors the relativistic shift at normal incidence is evident, and with no orthogonal component at $\theta = 0$, the dispersion relation requires $ck''_x = \omega''$, as seen. As θ increases the relativistic shift is reduced, since the orthogonal component undergoes no shift. At $\cos\theta = \beta$ the normal component can no longer "catch up" to the mirror, and $\omega'' = \omega$. This occurrence at $\cos\theta = \beta$ is also evident in the figure for both $\beta = 0.4$ and $\beta = 0.8$.



Equations (17.40) return in the next section, although with the reversed sign for β since the transformation there is from the moving frame to the lab frame.

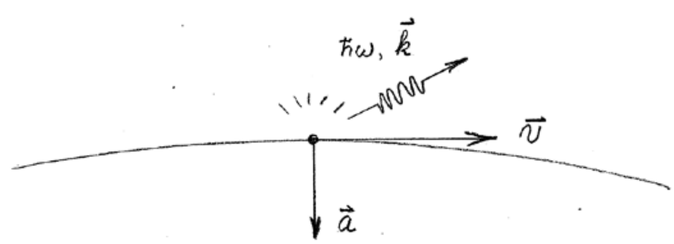
F. Example 2: synchrotron radiation.

In the last 30 years, spectroscopy has benefitted enormously from synchrotron radiation, which provides tunable sources of radiation from the far infrared to well into the X-ray region of the spectrum. Tunable X-rays have found a host of applications ranging from structural analysis through material characterization to trace-element detection with any number of applications in biology. From the perspective of Ch. 17, it offers an excellent opportunity to put theory into practice.

The geometry is shown in the figure at the top of the next page. A charged particle (electron) moving at relativistic speeds is accelerated radially in a "bending" magnet, which generates a $\frac{1}{c}\vec{v} \times \vec{B}$ force on the electron, causing a radial acceleration. The Liénard-Wiechert form derived in Ch. 13 is correct to third order in v/c if we do not expand the denominator, which means that

$$\vec{A}(\vec{r},t) = \left.\frac{q\vec{v}}{Rc\left(1 - \hat{R}\cdot\vec{v}\right)}\right|_{t=t_{ret}} . \tag{17.45}$$

where $\vec{R} = \vec{r} - \vec{r}\,'$. Jackson performs a Fourier analysis of this time dependence in his Sec. 14.6, but the physics can be illustrated without going into these details. We simply assume that these photons exist, then consider the relativistic projection of their 4-vectors $(\omega', c\vec{k}\,')$ in the rest frame of the electron to $(\omega/c, \vec{k})$ in the lab. Because in the far-field limit the electric field is



parallel to the transverse component of the acceleration vector, the emerging radiation is mainly polarized in the plane of the orbit.

Our working equations are (17.3). Consider a photon of frequency $\omega' = ck'$ and wave vector $\vec{k}' = k'(\hat{x}\cos\theta + \hat{y}\sin\theta)$, where θ is measured in the direction of $\vec{v}$ in the rest frame of the electron. By Eq. (17.3), the parameters of the emerging radiation are

$$\omega = \gamma(\omega' + \beta ck'\cos\theta) = \gamma\omega'(1+\beta\cos\theta); \qquad (17.46a)$$

$$ck_x = \gamma\omega'(\cos\theta + \beta); \qquad (17.46b)$$

$$ck_y = \omega'\sin\theta. \qquad (17.46c)$$

A short calculation shows that $c^2k^2/\omega^2 = 1$, as found for Eqs. (17.43). Given that $\gamma = 1/\sqrt{1-v^2/c^2}$ and $v \cong c$, the frequency ω of the radiation emerging in the forward direction $\cos\theta \sim 1$ can be enormously larger than ω'. This is the relativistic Doppler shift. An examination of Eqs. (17.46b) and (17.46c) shows that the radiation is strongly concentrated in the forward direction, and that very little appears in the backward direction ($\cos\theta \sim -1$). This can be formalized as

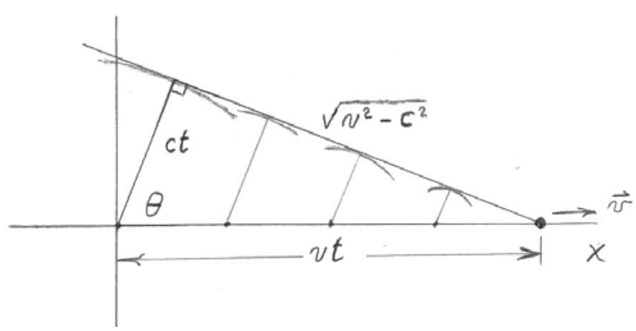
$$\theta = \tan^{-1}(k_y/k_x). \qquad (17.47)$$

In a more informal perspective, because the emerging radiation travels at *c* and the electron is traveling essentially at *c*, the oscillations pile up in the forward direction as described by γ.

Considerable engineering has gone into synchrotron-radiation technology, although it has not yet been brought to laboratory scales. Consequently, to perform experiments that require tunable X-rays, experimentalists must still travel to these facilities. However, EUV and soft-X-ray photons can be generated by tabletop sources via the method of high-harmonic generation, as described in Ch. 15. These provide photons at much lower frequencies, and are not yet (at least) competitive with the major facilities.

G. Example 3: Cerenkov radiation; shock waves.

Cerenkov radiation is a consequence of retardation physics, not special relativity. However, the reason for including its treatment in a chapter on special relativity is that it results when the speed of a charged particle exceeds the speed of light in the medium in which the particle is traveling. Whenever this occurs the result is a shock wave, whether the wave is electromagnetic, acoustic, or something else. Although the discussion here is centered on electromagnetic waves, the math is basically the same for other propagating disturbances.



The geometry is shown in the figure. Let a charged particle be moving parallel to the *x* axis at a speed that exceeds the speed of light in the medium. In contrast to the situation where $c > v$, the particle now outruns the wave fronts. The diagram illustrates the history of the spherical wavefronts emitted by the particle shown at the right at various times in the past. As the waves radiate outward, their fronts form a line, as illustrated by the triangles. Because the ratio of *c* to *v* is constant, all triangles are similar. The angle θ between the *x* axis and the shortest ray to the propagation locus follows from geometry, considering that a photon travels a distance ct when the particle moves vt. From the diagram

$$\cos\theta = \frac{ct}{vt} = \frac{c}{v}, \tag{17.48}$$

recalling that $v > c$ in the medium.

Next, consider the expression for the 4-potential, which we derived in Ch.13:

$$(\phi, \vec{A}) = \frac{1}{Rc}\int_V d^3r' \left.\frac{(\rho c, \rho\vec{v})}{1-\frac{\hat{R}\cdot\vec{v}}{c}}\right|_{t=t_{ret}} . \tag{17.49}$$

For $\rho(\vec{r}) = q\delta(\vec{r}-\vec{r}_o)$, the vector-potential part reduces to

$$\vec{A} = \frac{q\vec{v}}{Rc}\left.\frac{1}{1-\frac{\hat{R}\cdot\vec{v}}{c}}\right|_{t=t_{ret}} = \frac{q\vec{v}}{Rc}\left.\frac{1}{1-\frac{v\cos\theta}{c}}\right|_{t=t_{ret}} , \tag{17.50}$$

where $R = |\vec{r} - \vec{v}t|$, assuming that the particle is at the origin when $t = 0$. Up to now we have treated the denominator to first order, but with $\vec{v} > c$ this assumption is no longer valid. Of particular interest is the vanishing of the denominator at

$$\cos\theta = \frac{c}{v}. \tag{17.51}$$

Thus $\vec{A}$ is singular precisely in the direction of the wavefront. This, together with all waves contributing along the shock front, is the reason for the apparently anomalous strength of shock waves.

While the above treatment highlights the essential physics, a correct derivation, as done by Jackson, avoids the singularity. Jackson approaches this result by evaluating the energy dissipated in the material by the relativistic particle, in which case the infinity is replaced by a more gradual enhancement. The mathematics is somewhat on the formidable side, and as a result the basic mechanism is obscured.

A third direction follows from evaluation of the retarded time. Assuming that *q* is in uniform motion,

$$t_{ret} = t - \frac{1}{c}|\vec{r} - \vec{r}_o(t_{ret})| = t - \frac{1}{c}|\vec{r} - \vec{v}_o t_{ret}| \tag{17.52}$$

Writing $\vec{r} = \vec{r}_o + \vec{v}_o t$, Eq. (17.) can be rewritten as

$$t - t_{ret} = \Delta t = \frac{1}{c}|\vec{r} - \vec{r}_o - \vec{v}_o(t - t_{ret})| \qquad \text{(17.53a}$$

$$= \frac{1}{c} | \Delta\vec{r} - \vec{v}_o \Delta t | \tag{17.53b}$$

where $\Delta\vec{r} = \vec{r} - \vec{r}_o$. Solving this quadratic equation for Δt yields

$$\Delta t = \frac{-\Delta\vec{r} \cdot \vec{v}_o \pm \sqrt{(\Delta\vec{r} \cdot \vec{v}_o)^2 + \Delta r^2 (c^2 - v_o^2)}}{c^2 - v_o^2} . \tag{17.54}$$

If $c > v$, which is the usual case, then the positive sign must be selected.

If $v_o > c$, it is useful to rewrite Eq. (17.54) as

$$\Delta t = t - t_{ret} = \frac{\Delta\vec{r} \cdot \vec{v}_o \pm \sqrt{(\Delta\vec{r} \cdot \vec{v}_o)^2 - \Delta r^2 (v_o^2 - c^2)}}{v_o^2 - c^2} \Delta r \tag{17.55a}$$

$$= \frac{v_o \cos\beta \pm \sqrt{c^2 - v_o^2 \sin^2\beta)}}{v_o^2 - c^2} \Delta r . \tag{17.55b}$$

where $\cos\beta$ is the angle between $\Delta\vec{r}$ and $\vec{v}_o$. Equations (17.55) have a range where two solutions exist and another where none exist. The boundary between the two is the degeneracy point

$$\sin\beta = \pm \frac{c}{v_o} . \tag{17.56}$$

Thus β is the complement of θ. In particular, it defines the angle that the shock front makes with the reference axis.

H. Tensor formulation of special relativity.

The wave equation is hyperbolic, with one of the four terms having a sign opposite the other three. Thus the relevant space is non-Euclidean. Up to now this has not been a problem, but to show that it could become a problem, we contrast the demonstration of the covariance of the Lorentz gauge with that of the wave equation.

Considering the wave equation first, the objective is to show that if

$$\left(\nabla'^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t'^2} \right) (\phi', \vec{A}') = 0 , \tag{17.57a}$$

then

$$\left(\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) (\phi, \vec{A}) = 0 . \tag{17.57b}$$

The moving (primed)-vs.-fixed (unprimed) order makes no difference with the wave equation, but when considering the definition of the Lorentz gauge, the calculation is easier going from the moving to the rest system, and we work both calculations the same way. Because $(\phi, \vec{A})$ is a 4-vector, the components transform according to

$$\phi' = \gamma(\phi - \beta A_x)\,; \tag{17.58a}$$

$$A'_x = \gamma(A_x - \beta\phi)\,. \tag{17.58b}$$

For simplicity we ignore the *y* and *z* components of $\vec{A}$. Then by the chain rule of differentiation

$$\frac{\partial}{\partial x'} = \frac{\partial x}{\partial x'}\frac{\partial}{\partial x} + \frac{\partial t}{\partial x'}\frac{\partial}{\partial t}$$

$$= \gamma\frac{\partial}{\partial x} + \frac{\gamma\beta}{c}\frac{\partial}{\partial t}\,; \tag{17.59a}$$

$$\frac{\partial}{\partial t'} = \frac{\partial x}{\partial t'}\frac{\partial}{\partial x} + \frac{\partial t}{\partial t'}\frac{\partial}{\partial t}\,,$$

$$= \beta c\frac{\partial}{\partial x} + \gamma\frac{\partial}{\partial t} \tag{17.59b}$$

where we have used Eqs. (17.3).

With this background, we evaluate

$$\begin{aligned}
\left(\frac{\partial^2}{\partial x'^2} - \frac{1}{c^2}\frac{\partial^2}{\partial t'^2}\right)(\phi', \vec{A}') &= \left(\gamma^2\left(\frac{\partial}{\partial x} + \frac{\beta}{c}\frac{\partial}{\partial t}\right)\left(\frac{\partial}{\partial x} + \frac{\beta}{c}\frac{\partial}{\partial t}\right)\right. \\
&\quad \left. -\gamma^2\left(\beta\frac{\partial}{\partial x} + \frac{1}{c}\frac{\partial}{\partial t}\right)\left(\beta\frac{\partial}{\partial x} + \frac{1}{c}\frac{\partial}{\partial t}\right)\right)(\phi', \vec{A}) \\
&= \left(\gamma^2\left(\frac{\partial^2}{\partial x^2} + 2\frac{\beta}{c}\frac{\partial^2}{\partial x\partial t} + \frac{\beta^2}{c^2}\frac{\partial^2}{\partial t^2}\right)\right. \\
&\quad \left. -\gamma^2\left(\beta^2\frac{\partial^2}{\partial x^2} + 2\frac{\beta}{c}\frac{\partial^2}{\partial x\partial t} + \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\right)\left(\phi', A'_x\right) \\
&= \left(\frac{\partial^2}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\left(\phi', \vec{A}'\right).
\end{aligned} \tag{17.60}$$

Because ϕ' and $\vec{A}'$ are linear combinations of ϕ and $\vec{A}$, and because ϕ and $\vec{A}$ are themselves solutions of the wave equation, the proposition is proved. The important point here is that covariance is established within the operators themselves. It is not necessary to get the potentials involved directly.

The situation is different when first derivatives are involved, as is the case with the definition of the Lorentz gauge. Here the distinction between transformations between functions and derivatives is critical. We start with

$$\nabla'\cdot\vec{A}' + \frac{1}{c}\frac{\partial\phi'}{\partial t'} = 0 \tag{17.61}$$

and seek to determine whether this expression remains valid when everything is transformed to the unprimed system. Again ignoring the y and z components, the above expression, written in terms of the unprimed variables and potentials, is

$$\frac{\partial A'_x}{\partial x'}+\frac{1}{c}\frac{\partial \phi'}{\partial t'}=\left(\gamma\frac{\partial}{\partial x}+\frac{\gamma\beta}{c}\frac{\partial}{\partial t}\right)\gamma\left(A_x-\beta\phi\right)+\frac{1}{c}\left(\gamma\beta c\frac{\partial}{\partial x}+\gamma\frac{\partial}{\partial t}\right)\gamma\left(\phi-\beta A_x\right)$$

$$=\gamma^2\left(\frac{\partial A_x}{\partial x}-\beta\frac{\partial \phi}{\partial x}+\frac{\beta}{c}\frac{\partial A_x}{\partial t}-\frac{\beta^2}{c}\frac{\partial \phi}{\partial t}\right)$$

$$+\gamma^2\left(\beta\frac{\partial \phi}{\partial x}-\beta^2\frac{\partial A_x}{\partial x}+\frac{1}{c}\frac{\partial \phi}{\partial t}-\frac{\beta}{c}\frac{\partial A_x}{\partial t}\right).$$

$$=\gamma^2\left(1-\beta^2\right)\left(\frac{\partial A_x}{\partial x}+\frac{1}{c}\frac{\partial \phi}{\partial t}\right)=0. \tag{17.62}$$

Thus covariance is achieved, but only because the transformations must be applied to both derivatives and potentials. The differences in sign between direct and derivative operations are clearly important.

Given the difficulty in dealing with covariance, it can be appreciated that evaluation of the transformation rules for electric and magnetic fields can be much more difficult. Two approaches are in common use. One is to write 4-vectors in complex form, for example $s_4=(\vec{r},ict)$. The advantages of this approach are first, that the length of the 4-vector can be considered to be the sum of the squares of the components without including a minus sign for the scalar part, and second, in evaluating the chain rule for derivatives, *I* appears in the expression for the time derivative, thereby allowing the inverse transformation to be bypassed. However, there are costs involved: the 4-vectors are complex, and the calculation of the transformation rules for fields is no simpler than that of the approach using the metric tensor. In this regard we agree with Jackson, which is to keep the 4-vectors real and use metrics.

Jackson is in his element summarizing the metric-tensor approach in Secs. 11.6 and 11.7, and I follow his development here. He begins with group-theoretic terminology, noting that special relativity is described by the homogeneous Lorentz group, the collection of all operators that leave the 4-length

$$s^2=x_0^2-x_1^2-x_2^2-x_3^2 \tag{17.63}$$

invariant. The inhomogeneous Lorenz group, or the Poincaire group, is the set of operations that leaves the difference vector

$$s^2=(x_0-y_0)^2-(x_1-y_1)^2-(x_2-y_2)^2-(x_3-y_3)^2. \tag{17.64}$$

These are not considered in his treatment, or ours.

Operations are done in a four-dimensional space that is non-Euclidian, since one of the four dimensions enters the length with an opposite sign. The operations themselves are written as

$$x'^{\alpha}=x'^{\alpha}(x^0,x^1,x^2,x^3), \tag{17.65}$$

where $\alpha = 0, 1, 2, 3$. The rank of the tensor depends on its transformation properties. A scalar (tensor of rank zero) is a single quantity that is unchanged by the transformation. The 4-length obviously falls into this category.

4-vectors come in two varieties, contravariant and covariant. The contravariant vector A^α has components A^0, A^1, A^2, A^3 and transforms according to

$$A'^\alpha = \frac{\partial x'^\alpha}{\partial x^\beta} A^\beta \tag{17.66}$$

where summation over repeated indices is understood. The derivatives obviously fall into this category. Covariant vectors transform as

$$B'_\alpha = \frac{\partial x^\beta}{\partial x'^\alpha} B^\beta . \tag{17.67}$$

Tensors of rank 2 are 16-element objects that transform according to

$$F'^{\alpha\beta} = \frac{\partial x'^\alpha}{\partial x^\gamma}\frac{\partial x'^\beta}{\partial x^\delta} F^{\gamma\delta} . \tag{17.68}$$

The description of special relativity in this context is then built around the metric tensor $g_{\alpha\beta}$, which is diagonal and has elements $g_{00} = 1$, $g_{11} = g_{22} = g_{33} = -1$, which in the matrix representation is

$$g_{11} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} . \tag{17.69}$$

As can be guessed, the covariant and contravariant forms are identical, or $g_{\alpha\beta} = g^{\alpha\beta}$.

From this perspective the square of the 4-length is

$$s_4^2 = g_{\alpha\beta} x^\alpha x^\beta , \tag{17.70a}$$

or

$$s_4^2 = \begin{pmatrix} x^0 & x^1 & x^2 & x^3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} . \tag{17.70b}$$

Various identities can be proven much more easily with this approach than with the brute-force calculations given at the beginning of the section. For example the definition of the Lorentz gauge can be written

$$\partial^\alpha A_\alpha = \partial_\alpha A^\alpha = \frac{\partial A^0}{\partial x^0} + \nabla \cdot \vec{A} = 0\,. \tag{17.71}$$

Without going into detail, evaluation of the definitions of the electric and magnetic fields yields the second-rank field tensor

$$F^{\alpha\beta} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix}. \tag{17.72}$$

The transformation properties are given in vector form in Jackson's Eqs. (11.143):

$$\vec{E}\,' = \gamma(\vec{E} + \vec{\beta} \times \vec{B}) - \frac{\gamma^2}{\gamma+1}\vec{\beta}(\vec{\beta}\cdot\vec{E})\,; \tag{17.73a}$$

$$\vec{B}\,' = \gamma(\vec{B} - \vec{\beta} \times \vec{E}) - \frac{\gamma^2}{\gamma+1}\vec{\beta}(\vec{\beta}\cdot\vec{B})\,. \tag{17.73b}$$