Gravitation And Cosmology: Principles 
And Applications Of The General Theory Of Relativity

Steven Weinberg

“There are really four
dimensions, three which we
call the three planes of
Space, and a fourth, Time.
There is, however, a
tendency to draw an unreal
distinction between the
former three dimensions
and the latter, because it
happens that our conscious-
ness moves intermittently
in one direction along the
latter from the beginning
to the end of our lives.”

“That, said a very young
man, making spasmodic
efforts to relight his cigar
over the lamp; ‘that . . .
very clear indeed.’” H. G.
Wells, The Time Machine

2 SPECIAL
RELATIVITY

We now review Einstein’s Special Theory of Relativity. This chapter, while
self-contained, is only a brief summary, and aims primarily at establishing our
notation and collecting some formulas that will be useful later. The reader who
needs a more extensive introduction to special relativity is advised to turn to one
of the books listed at the end of this chapter, and then return. The reader who feels
completely at home with the subject may find it desirable to move on immediately
to Chapter 3.

1 Lorentz Transformations

The Principle of Special Relativity states that the laws of nature are invariant
under a particular group of space-time coordinate transformations, called Lorentz
transformations. We saw at the end of Chapter 1 that Newton’s laws of motion are
invariant under the Galilean coordinate transformations (1.3.2), but that Maxwell’s
equations are not, and that Einstein resolved this conflict by replacing Galilean invariance with Lorentz invariance. I shall not continue this discussion in historical terms, but shall simply define the Lorentz transformations, and then show how Lorentz invariance guides our search for the laws of nature.

A Lorentz transformation is a transformation from one system of space-time coordinates \( x^\gamma \) to another system \( x'^\gamma \), so that

\[
x'^\gamma = \Lambda^\gamma_\beta x^\beta + a^\gamma
\]  

(2.1.1)

where \( a^\gamma \) and \( \Lambda^\gamma_\beta \) are constants, restricted by the conditions

\[
\Lambda^\gamma_\gamma \Lambda^\gamma_\delta \eta_{\gamma \delta} = \eta_{\gamma \gamma}
\]

with

\[
\eta_{\alpha \beta} = \begin{cases} 
+1 & \alpha = \beta = 1, 2, 3, \text{ or } 3 \\
-1 & \alpha = \beta = 0 \\
0 & \alpha \neq \beta
\end{cases}
\]

(2.1.2)

(2.1.3)

In our notation \( \alpha, \beta, \gamma \), and so on, will always run over the four values 1, 2, 3, 0, with \( x^1, x^2, x^3 \) the Cartesian components of the position vector \( \mathbf{x} \) and \( x^0 \) the time \( t \). We shall use natural units in which the speed of light is unity, so all \( x^\gamma \) have the dimension of length. Any index, like \( \beta \) in Eq. (2.1.1), that appears twice, once as a subscript and once as a superscript, is understood to be summed over unless otherwise noted; that is, Eq. (2.1.1) is an abbreviation for

\[
x'^\gamma = \Lambda^\gamma_\gamma x^\gamma + \Lambda^\gamma_1 x^1 + \Lambda^\gamma_2 x^2 + \Lambda^\gamma_3 x^3 + a^\gamma
\]

The fundamental property that distinguishes the Lorentz transformations is that they leave invariant the "proper time" \( d\tau \), defined by

\[
d\tau^2 \equiv dt^2 - dx^2 = -\eta_{\alpha \beta} dx^\alpha dx^\beta
\]

(2.1.4)

In a new coordinate system \( x'^\gamma \), the coordinate differentials are given by (2.1.1) as

\[
dx'^\gamma = \Lambda^\gamma_\gamma dx^\gamma
\]

so the new coordinate time will be

\[
d\tau'^2 = -\eta_{\alpha \beta} dx'^\alpha dx'^\beta = -\eta_{\alpha \beta} \Lambda^\gamma_\gamma \Lambda^\gamma_\delta dx^\gamma dx^\delta = -\eta_{\gamma \delta} dx^\gamma dx^\delta
\]

and therefore

\[
d\tau'^2 = d\tau^2
\]

(2.1.5)

It is this property that accounts for the observation by Michelson and Morley that the speed of light is the same in all inertial systems. A light wave front will have \( dx/dt \) equal to the speed of light, which in our units is unity; hence the propagation of light is described by the statement that

\[
d\tau = 0
\]

(2.1.6)
Performing a Lorentz transformation does not change $dt$, so $d\tau^2 = 0$, and therefore $[dx'/dt'] = 1$; that is, the speed of light in the new coordinate system is still unity.

We can also show that the Lorentz transformations (2.1.1) are the only nonsingular coordinate transformations $x \to x'$ that leave $d\tau^2$ invariant. (Nonsingular means that $x'(x)$ and $x(x')$ are well-behaved differentiable functions, so that the matrix $\partial x'^\gamma / \partial x^\delta$ has a well-defined inverse $\partial x^\delta / \partial x^\gamma$.) A general coordinate transformation $x \to x'$ will change $d\tau$ into $d\tau'$, given by

$$d\tau'^2 = -\eta_{44} \, dx'^4 \, dx'^4 = -\eta_{44} \, \frac{\partial x'^4}{\partial x^4} \frac{\partial x^4}{\partial x^4} \, dx^4 \, dx^4$$

If this is equal to $d\tau^2$ for all $dx^i$, we must have

$$\eta_{44} = \eta_{44} \, \frac{\partial x^4}{\partial x^4} \frac{\partial x^4}{\partial x^4}$$

Differentiation with respect to $x^4$ gives

$$0 = \eta_{44} \, \frac{\partial^2 x^4}{\partial x^4 \partial x^4} - \frac{\partial x^4}{\partial x^4} \frac{\partial x^4}{\partial x^4}$$

To solve for the second derivatives, we add to this the same equation with $\gamma$ and $\varepsilon$ interchanged, and subtract the same with $\varepsilon$ and $\delta$ interchanged; that is,

$$0 = \eta_{44} \left[ \frac{\partial^2 x^4}{\partial x^4 \partial x^4} + \frac{\partial^2 x^4}{\partial x^4 \partial x^4} \right] - \frac{\partial x^4}{\partial x^4} \frac{\partial x^4}{\partial x^4}$$

The last term cancels the second, the penultimate cancels the fourth (because $\eta_{44} = \eta_{44}$), and the first equals the third, so we are left with

$$0 = 2\eta_{44} \, \frac{\partial^2 x^4}{\partial x^4 \partial x^4}$$

But both $\eta_{44}$ and $\partial x'^\gamma / \partial x^\delta$ are nonsingular matrices, so this immediately yields

$$0 = \frac{\partial^2 x^4}{\partial x^4 \partial x^4}$$

The general solution of (2.1.8) is of course just the linear function (2.1.1), and by inserting (2.1.1) in (2.1.7) we see that $A^{a}_{\beta}$ must be subject to the condition (2.1.2). This proof is an elementary example of the sort of thing we do in Chapter 13, on symmetric spaces. (Incidentally, if we had only assumed that the transformations
\( x \to x' \) leave \( dt \) invariant when \( dt = 0 \), that is, for a particle moving at the speed of light, then we would have found that these transformations are in general nonlinear, and form a 15-parameter group, the conformal group, which contains the Lorentz transformations as a subgroup. But the statement that a free particle moves at constant velocity would not be an invariant statement unless the velocity were that of light, and since there are massive particles in the world, we must reject the conformal group as a possible invariance of nature.)

The set of all Lorentz transformations of the form (2.1.1) is correctly called the inhomogeneous Lorentz group, or the Poincaré group. The subset with \( a^2 = 0 \) is called the homogeneous Lorentz group. Both the homogeneous and the inhomogeneous Lorentz groups have subgroups called the proper homogeneous and inhomogeneous Lorentz groups, defined by imposing on \( \Lambda^\alpha_\beta \) the additional requirements

\[
\Lambda^0_0 \geq 1; \quad \text{Det } \Lambda = +1 \tag{2.1.9}
\]

Note from (2.1.2) that:

\[
(\Lambda^0_\alpha)^2 = 1 + \sum_{i=1,2,3} (\Lambda^i_\alpha)^2 > 1 \tag{2.1.10}
\]

and

\[
(\text{Det } \Lambda)^2 = 1 \tag{2.1.11}
\]

([Equation (2.1.10) follows upon setting \( \gamma = \delta = 0 \) in (2.1.2). Equation (2.1.11) is derived by writing Eq. (2.1.2) as a matrix equation \( \eta = \Lambda^T \eta \Lambda \) and taking its determinant.] It follows that any \( \Lambda^\alpha_\beta \) that can be converted to the identity \( \delta^\alpha_\beta \) by a continuous variation of its parameters must be a proper Lorentz transformation, because it is impossible by a continuous change of parameters to jump from \( \Lambda^0_0 \leq -1 \) to \( \Lambda^0_0 \geq +1 \), or from \( \text{Det } \Lambda = -1 \) to \( \text{Det } \Lambda = +1 \), and the identity has \( \Lambda^0_0 = +1 \) and \( \text{Det } \Lambda = +1 \). The improper Lorentz transformations involve either space inversion (\( \text{Det } \Lambda = -1, \Lambda^0_0 \geq 1 \)), which is now known not to be an exact symmetry of nature,\(^1\) or time reversal (\( \text{Det } \Lambda = -1, \Lambda^0_0 \leq -1 \)), which is strongly suspected to be not an exact symmetry of nature,\(^2\) or their product. We are dealing almost exclusively with proper Lorentz transformations, and unless otherwise noted, any Lorentz transformation is assumed to satisfy Eq. (2.1.9).

The proper homogeneous Lorentz transformations have a further subgroup, consisting of the rotations, for which

\[
\Lambda^i_j = R_{ij}, \quad \Lambda^i_0 = \Lambda^0_i = 0, \quad \Lambda^0_0 = 1
\]

where \( R_{ij} \) is a unimodular orthogonal matrix (i.e., \( \text{Det } R = 1 \) and \( R^T R = 1 \)) and the indices \( i, j \) run over the values 1, 2, 3. With regard to both rotations and the space-time translations \( x' \to x' + a^\alpha \), there is no difference between the Lorentz group and the Galileo group discussed in Chapter 1. The difference arises only in those transformations, called boosts, that change the velocity of the coordinate
frame. Suppose that one observer O sees a particle at rest, and a second observer O' sees it moving with velocity \( \mathbf{v} \). From (2.1.1) we have

\[
dx^x = \Lambda^{x}_a \, dx^a \tag{2.1.12}
\]
or, since \( dx \) vanishes,

\[
dx^i = \Lambda^i_0 \, dt \quad (i = 1, 2, 3) \tag{2.1.13}
\]

\[
dt' = \Lambda^0_0 \, dt \tag{2.1.14}
\]

Dividing \( dx' \) by \( dt' \) gives the velocity \( \mathbf{v} \), so

\[
\Lambda^i_0 = v_i \Lambda^0_0 \tag{2.1.15}
\]

We can get a second relation between \( \Lambda^i_0 \) and \( \Lambda^0_0 \) by setting \( \gamma = \delta = 0 \) in Eq. (2.1.2):

\[
-1 = \Lambda^0_0 \, \Lambda^a_0 \eta_{a\bar{a}} = \sum_{i=1,2,3} (\Lambda^i_0)^2 - (\Lambda^0_0)^2 \tag{2.1.16}
\]

The solution of Eqs. (2.1.15) and (2.1.16) is

\[
\Lambda^0_0 = \gamma \tag{2.1.17}
\]

\[
\Lambda^i_0 = \gamma v_i \tag{2.1.18}
\]

where

\[
\gamma \equiv (1 - \mathbf{v}^2)^{-1/2} \tag{2.1.19}
\]

The other \( \Lambda^a_\bar{a} \) are not uniquely determined, because if \( \Lambda^a_\bar{a} \) carries a particle from rest to velocity \( \mathbf{v} \), then so does \( \Lambda^a_\bar{a} \, \bar{R}^\prime_\bar{a} \), where \( \bar{R} \) is an arbitrary rotation. One convenient choice that satisfies Eq. (2.1.2) is

\[
\Lambda^i_j = \delta_{ij} \cdot v_i v_j \frac{\gamma - 1}{\mathbf{v}^2} \tag{2.1.20}
\]

\[
\Lambda^0_j = \gamma v_j \tag{2.1.21}
\]

It can easily be seen that any proper homogeneous Lorentz transformation may be expressed as the product of a boost \( \Lambda(\mathbf{v}) \) times a rotation \( \bar{R} \).

### 2 Time Dilation

Although the Lorentz transformations were invented to account for the invariance of the speed of light, the change from Galilean relativity to special relativity had immediate kinematic consequences for material objects moving at speeds less than that of light. The simplest and most important is the time dilation of moving clocks. An observer looking at a clock at rest will see two ticks separated by a space-time interval \( dx = 0 \), \( dt = \Delta t \), where \( \Delta t \) is the nominal period between
ticks intended by the manufacturer. He will calculate the proper time interval (2.1.4) as
\[ dt \equiv (dt^2 - dx^2)^{1/2} = \Delta t \]
A second observer, who sees the same clock moving with velocity \( v \), will observe that the two ticks are separated by a time interval \( dt' \) and also by a space interval \( dx' = v \, dt' \), and he will conclude that the proper time interval is
\[ dt' \equiv (dt'^2 - dx'^2)^{1/2} = (1 - v^2)^{1/2} \, dt' \]
But both observers are supposed to be using inertial coordinate systems, so their coordinate systems are related by a Lorentz transformation, and on comparing notes they must find that \( dt = dt' \), in accordance with Eq. (2.1.5). It follows that the observer who sees the clock in motion will see it tick with a period
\[ dt' = \Delta t(1 - v^2)^{-1/2} \quad (2.2.1) \]
[For an alternate derivation, use Eqs. (2.1.14), (2.1.17), (2.1.19).] This relation is literally being verified every day by experiments that measure the mean lifetime of rapidly moving unstable particles from cosmic rays and accelerators. Such particles of course do not tick; instead (2.2.1) tells us here that a moving particle will have a mean life larger than it has at rest by a factor \( (1 - v^2)^{-1/2} \), in perfect agreement with the lifetime measurements made electronically or by measuring the free path length.

The time dilation (2.2.1) is not to be confused with the apparent time dilation or contraction known as the Doppler effect. If our “clock” is a moving source of light of frequency \( v = 1/\Delta t \), then the time between emission of successive wavefronts (say, with a maximum value of some component of the electric field) is given by (2.2.1) as \( dt' = \Delta t(1 - v^2)^{-1/2} \). However, during this time the distance from the observer to the light source will have increased by an amount \( v_c \, dt' \), where \( v_c \) is the component of \( v \) along the direction from observer to light source. Hence the period between reception of wavefronts will be
\[ dt_0 = (1 + v_c) \, dt' = (1 + v_c)(1 - v^2)^{-1/2} \, \Delta t \]
That is, the ratio of the frequency of the light actually measured by the observer to the frequency of the light source at rest is
\[ \frac{v_{\text{obs}}}{v} = (1 + v_c)^{-1}(1 - v^2)^{1/2} \quad (2.2.2) \]
If the light source is moving away, then \( v_c > 0 \), and this is necessarily a red shift. If the light source is moving transversely, then \( v_c = 0 \), and we have the pure time dilation red shift discussed above. If the light source is moving directly toward the observer, then \( v_c = -v \), and (2.2.2) gives a violet shift by a factor
\[ (1 + v)^{1/2}(1 - v)^{-1/2} \]
The transition from violet to red shift occurs for a source moving at an angle between straight toward the observer and at right angles to the line of sight.

3 Particle Dynamics

Let us suppose that a particle moves in a field of force at a velocity so high that Newtonian mechanics does not suffice to calculate its motion. Let us also suppose, as in the case of electrodynamics, that we know how to calculate the force $\mathbf{F}$ on our particle in any Lorentz frame in which, at a given moment, it is at rest. Then we could compute the motion of our particle by performing a Lorentz transformation to a frame in which the particle is at rest at some time $t_0$, computing the velocity $d\mathbf{v} = \mathbf{F} \, dt/m$ at the time $t_0 + dt$, performing another Lorentz transformation to bring the velocity to zero again, and so on. Fortunately, there is an easier way.

Let us define the relativistic force $f^a$ acting on a particle with coordinates $x^a(\tau)$ by

$$f^a = \dot{m} \frac{d^2 x^a}{d\tau^2} \tag{2.3.1}$$

Clearly, if $f^a$ were known, we could compute the motion of our particle. We shall relate $f^a$ to the Newtonian force by noting two of its properties:

(A) If the particle is momentarily at rest, then the proper time interval $d\tau$ equals $dt$, so $f^a = F^a$, where $F^a$ are the Cartesian components of the nonrelativistic force $\mathbf{F}$, and

$$F^0 = 0 \tag{2.3.2}$$

(B) Under a general Lorentz transformation (2.1.1), the coordinate differentials transform according to $dx^a = \Lambda^a_\beta \, dx^\beta$, while $d\tau$ is invariant, so (2.3.1) tells us that $f^a$ has the Lorentz transformation rule:

$$f'^a = \Lambda^a_\beta f^\beta \tag{2.3.3}$$

Any quantity such as $dx^a$ or $f^a$ that transforms according to Eq. (2.3.3) is called a four-vector.

Now suppose that our particle has velocity $\mathbf{v}$ at some moment $t_0$, and introduce a new coordinate system $x'^a$, defined by

$$x'^a = \Lambda^a_\beta (\mathbf{v}) x^\beta$$

where $\Lambda(\mathbf{v})$ is the “boost” defined by Eqs. (2.1.17)–(2.1.21). Since $\Lambda(\mathbf{v})$ is constructed so as to carry a particle from rest to velocity $\mathbf{v}$, and since our particle has velocity $\mathbf{v}$ at time $t_0$ in the coordinate system $x^a$, it must be at rest at this moment in the coordinate system $x'^a$. Hence, according to (A), the force four-vector in the
coordinate system \( x^\alpha \) at time \( t_0 \) is equal to the nonrelativistic force \( f^\alpha \). And therefore, according to (B), the force in our original coordinate system is

\[
f^\alpha = \Lambda^\alpha_\beta (v) F^\beta
\]

(2.3.4)

or more explicitly, since \( F^0 = 0 \),

\[
f = F + (\gamma - 1) v \frac{(v \cdot F)}{v^2}
\]

(2.3.5)

\[
f^0 = \gamma v \cdot F = v \cdot f
\]

(2.3.6)

with \( v \) the instantaneous velocity.

Now that we know how to calculate \( f^\alpha \), we can use the differential equations (2.3.1) to calculate the four dependent variables \( x^\tau (\tau) \), and then eliminate \( \tau \) to determine \( x(\tau) \). However, the initial values of \( dx^\alpha /d\tau \) must be chosen so that \( d\tau \) really is the proper time, that is, so that

\[
-1 = \eta_{\alpha\beta} \frac{dx^\alpha}{d\tau} \frac{dx^\beta}{d\tau}
\]

(2.3.7)

Note that (2.3.7) will be true for all \( \tau \) if it is true at some initial \( \tau \), providing that its derivative vanishes, that is, providing that

\[
0 = 2\eta_{\alpha\beta} f^\alpha \frac{dx^\beta}{d\tau}
\]

(2.3.8)

That this is true can be seen either directly from (2.3.4), or more elegantly by noticing that the right-hand side is Lorentz-invariant:

\[
\eta_{\alpha\beta} f^\alpha \frac{dx^\beta}{d\tau} - \eta_{\alpha\beta} \Lambda^\alpha_\gamma \Lambda^\gamma_\delta f^\delta \frac{dx^\beta}{d\tau}
\]

\[
= \eta_{\alpha\beta} f^\alpha \frac{dx^\beta}{d\tau}
\]

and that it vanishes by virtue of (2.3.2) in a reference frame in which the particle is at rest.

4 Energy and Momentum

The relativistic form (2.3.1) of Newton’s second law immediately suggests that we define an energy-momentum four-vector

\[
p^\alpha \equiv m \frac{dx^\alpha}{d\tau}
\]

(2.4.1)
and write the second law as

$$\frac{dp^a}{d\tau} = f^a$$  \hfill (2.4.2)$$

Recall that

$$dt \equiv (d\tau^2 - dx^2)^{1/2} - (1 - v^2)^{1/2} dt$$

where

$$v \equiv \frac{dx}{dt}$$

Then the space components of $p^a$ form the momentum vector

$$p = m\gamma v$$  \hfill (2.4.3)$$

and its time component is the energy

$$p^0 \equiv E = m\gamma$$  \hfill (2.4.4)$$

where

$$\gamma \equiv \frac{d\tau}{dt} = (1 - v^2)^{-1/2}$$  \hfill (2.4.5)$$

For small $v$, these definitions give

$$p = mv + 0(v^3)$$  \hfill (2.4.6)$$

$$E = m + \frac{1}{2}mv^2 + 0(v^4)$$  \hfill (2.4.7)$$

in agreement with the nonrelativistic formulas, except for the term $m$ in $E$. (Recall that in our units 1 sec equals $3 \times 10^{10}$ cm, so 1 g equals $9 \times 10^{24}$ ergs.) Sometimes the factor $m\gamma$ is called the relativistic mass $\tilde{m}$, so that $p = \tilde{m}v$. I do not follow this custom here; for us, “mass” will always mean the constant $m$.

Why do we call $p$ and $E$ the relativistic momentum and energy? We can use these names for anything we like, but if the concepts of momentum and energy are to be useful, they must be reserved for quantities that are conserved. The unique feature of our $p$ and $E$ is that, if one observer says that they are conserved in a reaction, then so will any other observer related to the first by a Lorentz transformation. Note that $dx^a$ is a four-vector whereas $m$ and $dt$ are invariants, so the $p^a$ for any single particle is a four-vector; that is, it transforms under (2.1.1) like

$$p'^a = \Lambda^a_b p^b$$

Since $\Lambda$ does not depend on anything but the Lorentz transformation being performed, it follows that in any reaction, the change of the sum of the $p^a$ of all particles is also a four-vector:

$$\Lambda \sum_p p_p^a = \Lambda^a_b \sum_p p_p^b$$
(The sums run over all particles, and $\Delta$ denotes the difference between initial and final states.) The conservation of $p$ and $E$ in the original inertial frame tells us that $\Delta \sum_n p_n^j$ vanishes, so in any coordinate system related to the first by a Lorentz transformation they will still be conserved; that is, $\Delta \sum_n p_n^\alpha$ will vanish.

(I shall not show here that $p$ and $E$ are the only functions of velocity whose conservation is Lorentz-invariant. However, it is worth stressing that $E$ must be conserved if $p$ is. For suppose that momentum is conserved in two different coordinate systems related by a Lorentz transformation, that is,

$$\Delta \sum_n p_n = 0 \quad \Delta \sum_n p'_n = 0$$

Since $\Delta \sum_n p_n$ is a four-vector, we have

$$\Delta \sum_n p_n^i = \Lambda^i_\beta \Delta \sum_n p_n^\beta$$

and using momentum conservation in both coordinate systems, this gives

$$0 = \Lambda^i_0 \Delta \sum_n p_n^0$$

But $\Lambda^i_0$ is not necessarily zero, so $p^0 - E$ is conserved.)

At zero velocity the energy $E$ has the finite value $m$. For this reason we sometimes give the name "kinetic energy" to the quantity $E - m$, which for small $v$ is approximately $\frac{1}{2}mv^2$. If the total mass is conserved in a reaction (as in elastic scattering), then the kinetic energy is conserved, but if some mass is destroyed (as in radioactive decay or fusion or fission), then very large quantities of kinetic energy will be liberated, with consequences of well-known importance.

The velocity can be eliminated from Eqs. (2.4.3) and (2.4.4), yielding a relation between energy and momentum

$$E(p) = (p^2 + m^2)^{1/2} \quad (2.4.8)$$

This can also be derived by noting from (2.4.1) and the definition of $dt$ that

$$\eta_{\alpha\beta} p^\alpha p^\beta = -m^2 \quad (2.4.9)$$

For a photon or neutrino we must set $v^2 = 1$ and $m = 0$, so (2.4.3) and (2.4.4) become indeterminate, but their ratio gives a relation useful for all particles

$$\frac{p}{E} = v \quad (2.4.10)$$

Note that for $m = 0$ Eq. (2.4.8) gives

$$E' = |p|$$

so $v$ is a unit vector, as of course it must be for a massless particle.
5 Vectors and Tensors

Next we go on to electrodynamics and relativistic hydrodynamics, but it is convenient first to pause and outline a notation that makes the Lorentz transformation properties of physical quantities transparent. This notation will be extended in Chapter 4, on tensor analysis, to encompass general coordinate transformations, but in fact few changes will be needed.

We have already introduced the term “four-vector” for any quantity such as $dx^a$ or $f^a$ or $p^a$ that undergoes the transformation

$$V^a \rightarrow V'^a = \Lambda^a_{\beta} V^\beta$$

when the coordinate system is transformed by

$$x^a \rightarrow x'^a = \Lambda^a_{\beta} x^\beta$$

(2.5.2)

More precisely, such a $V^a$ should be called a contravariant four-vector, to distinguish it from a covariant four-vector, defined as a quantity $U_a$ whose transformation rule is

$$U_a \rightarrow U'_a = \Lambda^\beta_a U_\beta$$

(2.5.3)

where

$$\Lambda^\beta_a = \eta_{\alpha\beta} \Lambda^\alpha_\delta$$

(2.5.4)

The matrix $\eta^{\alpha\beta}$ introduced here is numerically the same as $\eta_{\alpha\beta}$, that is,

$$\eta^{\alpha\beta} = i \eta_{\alpha\beta}$$

(2.5.5)

but we write it with indices upstairs to conform with our summation convention. Note that

$$\eta^{\alpha\beta} \eta_{\alpha\beta} = \delta^\beta_\alpha = \begin{cases} +1 & \alpha = \beta \\ 0 & \alpha \neq \beta \end{cases}$$

(2.5.6)

so $\Lambda^\beta_a$ is the inverse of the matrix $\Lambda^\alpha_\beta$, that is,

$$\Lambda^\gamma_a \Lambda^\alpha_\gamma - \eta_{\alpha\rho} \Lambda^\gamma_\rho \Lambda^\beta_\delta = \eta^\gamma_\rho \delta^\beta_\rho$$

(2.5.7)

It follows that the scalar product of a contravariant with a covariant four-vector is invariant, that is,

$$U_a' V'^a = \Lambda^\beta_a \Lambda^\alpha_\beta U_\alpha V^\alpha = U_\beta V^\beta$$

(2.5.8)

To every contravariant four-vector $V^a$ there corresponds a covariant four-vector

$$V_a = \eta_{a\beta} V^\beta$$

(2.5.9)

and to every covariant $U_a$ there corresponds a contravariant

$$U^a = \eta^{a\beta} U_\beta$$

(2.5.10)
Note that raising the index on $V_x$ simply gives back $V^x$, and lowering the index on $U_x$ simply gives back $U^x$,

$$\eta^a_\gamma V_\beta = \eta^a_\gamma \eta^\beta_\gamma V^x = V^x$$
$$\eta^a_\delta U^\beta = \eta^{\alpha}_\delta \eta^{\beta}_\gamma U^x = U_x$$

Note also that (2.5.9) does yield a covariant, because

$$V'_x = \eta^a_\alpha V'^x = \eta^a_\alpha \wedge^\alpha_\gamma V^\gamma = \eta^a_\alpha \eta^\gamma_\delta \wedge^\delta_\gamma V^\delta = \Lambda^\delta_\gamma \Lambda^\gamma_\delta V$$

in agreement with (2.5.3). Similarly, (2.5.10) does yield a contravariant.

Although any vector can be written in a contravariant or a covariant form, there are some vectors, such as $dx^x$, that appear more naturally contravariant and others that appear more naturally covariant. An example of the latter is the gradient $\partial/\partial x^x$, which obeys the transformation rule

$$\frac{\partial}{\partial x^x} \rightarrow \frac{\partial}{\partial x'^x} = \frac{\partial x^a}{\partial x'^x} \frac{\partial}{\partial x^x}$$

Multiplying (2.5.2) by $\Lambda^\gamma_\alpha$ gives

$$x'^\gamma = \Lambda^\gamma_\alpha x^\alpha$$

so

$$\frac{\partial x^a}{\partial x'^x} = \Lambda^a_\alpha$$

and therefore the gradient is covariant:

$$\frac{\partial}{\partial x'^x} = \Lambda^a_\alpha \frac{\partial}{\partial x^a} \quad (2.5.11)$$

One consequence is that the divergence of a contravariant vector $\partial V^x/\partial x^x$ is invariant. Another is that the scalar product of $\partial/\partial x^x$ with itself, the d'Alembertian operator

$$\Box = \eta^{a\beta} \frac{\partial}{\partial x^a} \frac{\partial}{\partial x^x} = \nabla^2 - \frac{\partial^2}{\partial t^2} \quad (2.5.12)$$

is also invariant.

Many physical quantities are not scalars or vectors, but more complicated objects called tensors. A tensor has several contravariant and/or covariant indices with corresponding Lorentz transformation properties, for example,

$$T'^{\gamma \beta} \rightarrow T'^{\gamma \beta} = \Lambda^\gamma_\delta \Lambda^\beta_\zeta T^{\delta \zeta}$$

A contravariant or covariant vector can be regarded as a tensor with one index, and a scalar is a tensor with no indices. There are several ways of forming tensors out of other tensors:
(A) **Linear Combinations.** A linear combination of tensors with the same upper and lower indices is a tensor with these indices. For instance, if \( R^\beta_\gamma \) and \( S^\beta_\gamma \) are tensors, and \( a \) and \( b \) are scalars, and we define

\[
T^\beta_\gamma \equiv aR^\beta_\gamma + bS^\beta_\gamma
\]

then \( T^\beta_\gamma \) is a tensor, that is,

\[
T^\beta_\gamma \equiv aR^\beta_\gamma + bS^\beta_\gamma = a\Lambda^\delta_\gamma \Lambda^\delta_\gamma R^\gamma_\delta + b\Lambda^\delta_\gamma \Lambda^\delta_\gamma S^\gamma_\delta = \Lambda^\gamma_\delta \Lambda^\delta_\gamma T^\gamma_\delta
\]

(B) **Direct Products.** The product of the components of two tensors yields a tensor whose upper and lower indices consist of all the upper and lower indices of the two original tensors. For instance, if \( A^\alpha_\delta \) and \( B^\gamma_\xi \) are tensors, and

\[
T^\beta_\gamma \equiv A^\alpha_\delta B^\gamma_\xi
\]

then \( T^\beta_\gamma \) is a tensor, that is,

\[
T^\beta_\gamma \equiv A^\alpha_\delta B^\gamma_\xi = \Lambda^\alpha_\delta \Lambda^\beta_\gamma \Lambda^\gamma_\delta \Lambda^\delta_\xi T^\delta_\gamma
\]

(C) **Contraction.** Setting an upper and lower index equal and summing it over its values 0, 1, 2, 3, yields a tensor with these two indices absent. For instance, if \( T^\alpha_\beta \) is a tensor and

\[
T^\alpha_\beta \equiv T^\alpha_\gamma \gamma
\]

then \( T^\alpha_\beta \) is a tensor, that is,

\[
T^\alpha_\beta \equiv T^\alpha_\gamma \gamma = \Lambda^\delta_\beta \Lambda^\gamma_\beta \Lambda^\gamma_\delta \Lambda^\delta_\xi T^\delta_\xi
\]

(D) **Differentiation.** The derivative \( \partial / \partial x^\alpha \) of any tensor is a tensor with one additional lower index \( \alpha \). For instance, if \( T^\beta_\gamma \) is a tensor and

\[
T^\beta_\gamma \equiv \partial / \partial x^\alpha \quad T^\beta_\gamma
\]

then \( T^\beta_\gamma \) is a tensor, that is,

\[
T^\beta_\gamma \equiv \partial / \partial x^\alpha \quad T^\beta_\gamma = \Lambda^\delta_\beta \partial / \partial x^\alpha \Lambda^\gamma_\alpha \Lambda^\gamma_\delta T^\delta_\xi
\]

\[
= \Lambda^\delta_\alpha \Lambda^\delta_\gamma T^\delta_\xi
\]
Note that the order of indices matters, even as between upper and lower indices. For instance, $T^{\alpha\beta}_{\gamma\delta}$ may or may not be the same as $T^{\beta\gamma}_{\alpha\delta}$.

Aside from the scalars, there are three special tensors whose components are the same in all coordinate systems:

(i) **The Minkowski Tensor.** The definition of Lorentz transformations tells us immediately that $\eta_{ab}$ is a covariant tensor,

$$\eta_{ab} = \Lambda^a_i \Lambda^b_j \eta_{ij}$$

Multiplying this equation by $\eta^{ac} \eta^{bd}$ and using (9.5.6) and (9.5.4), we find that

$$\eta^{ac} = \eta^{ac} \eta^{bd} \Lambda^i_a \Lambda^j_b \eta_{ij}$$

$$= \eta^{ak} \Lambda^i_k \Lambda^j_b$$

so $\eta^{ab}$ is a contravariant tensor. (Recall that $\eta_{ab}$ and $\eta^{ab}$ are numerically the same matrix, so this is a matrix that is both covariant and contravariant.) We can form a mixed tensor by lowering one index on $\eta^{ab}$ or raising one index on $\eta_{ab}$; this gives the Kronecker symbol

$$\delta^a_i = \eta^{ac} \eta_{ic}$$

That this is a tensor follows from rules (B) and (C) and the fact that $\eta^{ac}$ and $\eta_{ic}$ are tensors.

(ii) **The Levi-Civita Tensor.** This is a quantity $\varepsilon^{ab\gamma\delta}$ defined by

$$\varepsilon^{ab\gamma\delta} = \begin{cases} 
+1 & \text{if } \alpha\beta\gamma\delta \text{ even permutation of } 0123 \\
-1 & \text{if } \alpha\beta\gamma\delta \text{ odd permutation of } 0123 \\
0 & \text{otherwise} 
\end{cases} \quad (2.5.13)$$

Note that

$$\Lambda^a_\alpha \Lambda^b_\beta \Lambda^\gamma_\kappa \Lambda^\delta_\lambda \varepsilon^{\epsilon\kappa\lambda} \propto \varepsilon^{ab\gamma\delta}$$

because the left-hand side must be odd under any single permutation of the indices $\alpha\beta\gamma\delta$. To find the constant of proportionality, set $\alpha\beta\gamma\delta = 0123$. The left-hand side is then simply the determinant of $\Lambda$, which for proper Lorentz transformations is unity. (See Section 2.1.) Thus the constant of proportionality is unity, that is,

$$\Lambda^a_\alpha \Lambda^b_\beta \Lambda^\gamma_\kappa \Lambda^\delta_\lambda \varepsilon^{\epsilon\kappa\lambda} = \varepsilon^{ab\gamma\delta} \quad (2.5.14)$$

and therefore $\varepsilon^{ab\gamma\delta}$ is a tensor.

(iii) **The Zero Tensor.** We can define a tensor with an arbitrary pattern of upper and lower indices by setting all its components equal to zero.

Since $\eta^{ab}$ and $\eta_{ab}$ are tensors, we can use them to raise or lower indices on an arbitrary tensor; rules (B) and (C) tell us that this gives a new tensor with one
more upper or lower index and one less lower or upper index. For instance, if $T_{a\beta}$ is a tensor, then so is

$$T_{a\gamma} \equiv \eta^{a\delta} T_{\delta\gamma}$$

In particular, we can lower some or all of the indices on the Levi-Civita tensor $\varepsilon^{a\beta\gamma\delta}$. Lowering all the indices gives back the same numerical quantity except for a minus sign:

$$\varepsilon_{a\beta\gamma\delta} = -\varepsilon^{a\beta\gamma\delta} \quad (2.5.15)$$

The point of all this algebra is that it enables us to tell at a glance that an equation is Lorentz-invariant. The fundamental theorem is that if two tensors, with the same upper and lower indices, are equal in one coordinate system, then they are equal in any other coordinate system related to the first by a Lorentz transformation. For instance, if $T^{a}_{\beta} = S^{a}_{\beta}$, then

$$T^{\gamma}_{\beta} = \Lambda^{a}_{\gamma} \Lambda^{\beta}_{\delta} T^{a}_{\delta} = \Lambda^{a}_{\gamma} \Lambda^{\beta}_{\delta} S^{a}_{\delta} = S^{\gamma}_{\beta}$$

In particular, the statement that a tensor vanishes is Lorentz-invariant.

The formalism outlined in this section is nothing but a description of the representations of the homogeneous Lorentz group. We shall explore these representations in greater generality in Section 2.12.

6 Currents and Densities

Suppose that we have a system of particles with position $x_n(t)$ and charges $e_n$. The current and charge densities are usually defined by

$$J(x, t) = \sum_n e_n \delta^3(x - x_n(t)) \frac{dx_n(t)}{dt} \quad (2.6.1)$$

$$\varepsilon(x, t) = \sum_n e_n \delta^3(x - x_n(t)) \quad (2.6.2)$$

Here $\delta^3$ is the Dirac delta function, defined by the statement that for any smooth function $f(x)$,

$$\int d^3x f(x) \delta^3(x - y) = f(y)$$

We can unite $J$ and $\varepsilon$ into a four-vector $J^a$ by setting

$$J^0 \equiv \varepsilon \quad (2.6.3)$$

that is

$$J^a(x) = \sum_n e_n \delta^3(x - x_n(t)) \frac{dx_n^a(t)}{dt} \quad (2.6.4)$$
To show that this is a four-vector, define \( \mathbf{x}_n^0(t) = t \), and write (2.6.4) as

\[
J^*(x) = \int dt' \sum_n \epsilon_a \delta^4(x - \mathbf{x}_n(t')) \frac{dx_n^a(t')}{dt'}
\]

The differentials \( dt' \) cancel, and hence can be replaced with an invariant \( d\tau \):

\[
J^*(x) = \int dt \sum_n \epsilon_a \delta^4(x - \mathbf{x}_n(t)) \frac{dx_n^a(t)}{d\tau} \quad (2.6.5)
\]

But \( \delta^4(x - \mathbf{x}_n(t)) \) is a scalar (because \( \text{Det } \Lambda = 1 \)) and \( dx_n^a \) is a four-vector, so \( J^* \) is a four vector.

We also note that

\[
\nabla \cdot \mathbf{J}(x, t) = \sum_n \epsilon_a \frac{\partial}{\partial x^a} \delta^4(x - \mathbf{x}_n(t)) \frac{dx_n^a(t)}{dt}
\]

\[
= -\sum_n \epsilon_a \frac{\partial}{\partial x_n^a} \delta^4(x - \mathbf{x}_n(t)) \frac{dx_n^a(t)}{dt}
\]

\[
= -\sum_n \epsilon_a \frac{\partial}{\partial t} \delta^4(x - \mathbf{x}_n(t))
\]

\[
= -\frac{\partial}{\partial t} \epsilon(x, t)
\]

or, in four-dimensional language

\[
\frac{\partial}{\partial x^a} J^*(x) = 0 \quad (2.6.6)
\]

The Lorentz invariance of this statement is evident.

Whenever any current \( J^*(x) \) satisfies the invariant conservation law (2.6.6), we can form a total charge

\[
Q = \int d^3x J^0(x) \quad (2.6.7)
\]

This quantity is time-independent, because (2.6.6) and Gauss's theorem give

\[
\frac{dQ}{dt} = \int d^3x \frac{\partial}{\partial x^0} J^0(x) = -\int d^3x \nabla \cdot \mathbf{J}(x) = 0
\]

If \( J^*(x) \) is a four-vector, then \( Q \) is not only constant but a scalar. To see this, write \( Q \) as

\[
Q = \int d^4x J^*(x) \delta^2(\mathbf{n} \cdot \mathbf{r}) \quad (2.6.8)
\]
where \( \theta \) is the step function

\[
\theta(s) = \begin{cases} 
1 & s > 0 \\
0 & s < 0 
\end{cases}
\]

and \( n_2 \) is defined by

\[
n_1 \equiv n_2 \equiv n_3 \equiv 0, \quad n_0 \equiv +1
\]

The effect of a Lorentz transformation on \( Q \) is then evidently simply to change \( n \):

\[
Q' = \int d^4x J^\alpha(x) \partial_\alpha \theta(n'_\mu x^\mu)
\]

\[
n'_\mu = \Lambda^\alpha_\mu n_\alpha
\]

and using (2.6.6), the change in \( Q \) is then

\[
Q' - Q = \int d^4x \partial_\mu [J^\mu(x)\{\theta(n'_\mu x^\mu) - \theta(n_\mu x^\mu)\}]
\]

The current \( J^\mu(x) \) can be presumed to vanish if \( |x| \to \infty \) with \( t \) fixed, whereas the function \( \theta(n'_\mu x^\mu) - \theta(n_\mu x^\mu) \) vanishes if \( |t| \to \infty \) with \( x \) fixed. Hence we can apply the four-dimensional Gauss theorem, and find \( Q' - Q = 0 \); that is, \( Q \) is a scalar. (For the current density \( J^0 \) defined by (2.6.2) the charge (2.6.7) is

\[
Q = \sum_n e_n
\]

which of course is a constant scalar; however, in dealing with the charge and current distributions of extended particles it is important to realize that (2.6.7) defines a time-independent scalar for any conserved four-vector \( J^\mu \).)

7 Electro dynamics

Maxwell’s equations for the electric and magnetic fields \( \mathbf{E}, \mathbf{B} \) produced by a given charge density \( \rho \) and current density \( \mathbf{J} \) are

\[
\nabla \cdot \mathbf{E} = \rho \\
\n(2.7.1)
\]

\[
\nabla \times \mathbf{B} = \frac{\partial \mathbf{E}}{\partial t} + \mathbf{J} \\
\n(2.7.2)
\]

\[
\nabla \cdot \mathbf{B} = 0 \\
\n(2.7.3)
\]

\[
\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \\
\n(2.7.4)
\]
To uncover the Lorentz transformation properties of \( E \) and \( B \), we introduce a matrix \( F^{\alpha \beta} \), defined by

\[
\begin{align*}
F^{12} &= B_3 \\
F^{23} &= B_4 \\
F^{01} &= E_4 \\
F^{02} &= E_4 \\
F^{03} &= E_3 \\
F^{\alpha \beta} &= -F^{\beta \alpha}
\end{align*}
\] (2.7.5)

Then (2.7.1) and (2.7.2) can be written as

\[
\frac{\partial}{\partial x^\alpha} F^{\alpha \beta} = -J^\beta
\] (2.7.6)

(recall that \( J^0 \equiv \varepsilon \)) whereas (2.7.3) and (2.7.4) give

\[
\varepsilon^{\alpha \beta \gamma \delta} \frac{\partial}{\partial x^\delta} F_{\gamma \delta} = 0
\] (2.7.7)

where \( \varepsilon^{\alpha \beta \gamma \delta} \) is the Levi-Civita symbol defined in Section 2.5, and \( F_{\gamma \delta} \) is the covariant defined as usual by

\[
F_{\gamma \delta} \equiv \eta_{\gamma \nu} \eta_{\delta \mu} F^{\mu \nu}
\]

Since \( J^\alpha \) is a four-vector, we conclude that \( F^{\alpha \beta} \) is a tensor,

\[
F^{\alpha \beta} = \Lambda^\alpha_\gamma \Lambda^\beta_\delta F^{\gamma \delta}
\] (2.7.8)

because if \( F^{\alpha \beta} \) is a solution of (2.7.6) and (2.7.7), then (2.7.8) will be a solution in a Lorentz transformed coordinate system.

The electromagnetic force on a charged particle is

\[
f^\gamma = \varepsilon \eta_{\beta \delta} F^{\alpha \delta} \frac{dx^\gamma}{d\tau} = \varepsilon F^\gamma \frac{dx^\gamma}{d\tau}
\] (2.7.9)

That this is correct may be seen by repeating the arguments of Section 3. Equation (2.7.9) is correct in a reference system in which the particle is at rest because in this frame it gives \( f = eE, f^\gamma = 0 \), and it transforms like a four-vector, so it is correct for all velocities. Note incidentally that (2.7.9) and (2.4.2) give

\[
\frac{dp}{dt} = e[\mathbf{E} + \mathbf{v} \times \mathbf{B}]
\]

so the formula for magnetic force follows as a consequence of special relativity.

There is a useful alternate form to the homogeneous equations (2.7.7):

\[
\frac{\partial}{\partial x^\alpha} F_{\beta \gamma} + \frac{\partial}{\partial x^\beta} F_{\gamma \alpha} + \frac{\partial}{\partial x^\gamma} F_{\alpha \beta} = 0
\] (2.7.10)

Note that for \( \alpha, \beta, \gamma \) all different, Eq. (2.7.10) is the same as (2.7.7); for instance, setting \( \alpha = 0 \) in Eq. (2.7.7) gives the same result as setting \( \alpha \beta \gamma = 123 \) in Eq.
(2.7.10). On the other hand, for two indices equal, Eq. (2.7.10) is an identity; for instance, if \( \beta = \gamma \) then (2.7.10) reads

\[
\frac{\partial}{\partial x^\beta} F_{\beta \alpha} + \frac{\partial}{\partial x^\gamma} F_{\gamma \alpha} = 0 \quad \text{(not summed)}
\]

and this is identically true because \( F_{\gamma \beta} = -F_{\beta \gamma} \).

Equation (2.7.7) allows us to represent \( F_{\gamma \delta} \) as a "curl" of a four-vector \( A_\gamma \):

\[
F_{\gamma \delta} = \frac{\partial}{\partial x^\gamma} A_\delta - \frac{\partial}{\partial x^\delta} A_\gamma \tag{2.7.11}
\]

(See section 4.11.)

We can change \( A_\gamma \) by a term \( \partial_\phi \) without affecting \( F_{\gamma \delta} \); so \( A_\gamma \) may be defined so that

\[
\partial^\gamma A_\gamma = 0 \tag{2.7.12}
\]

With (2.7.11) and (2.7.12), the rest of Maxwell's equations reduce to

\[
\square A_\alpha = -J_\alpha \tag{2.7.13}
\]

8 Energy-Momentum Tensor

In Section 5 we introduced the density \( \varepsilon \) and current \( J \) of electric charge. We now give a similar definition for the density and current of the energy-momentum four-vector \( p^\alpha \). First consider a system of particles labeled \( n \), with energy-momentum four-vectors \( p_n^{\alpha}(t) \). The density of \( p^\alpha \) is defined by

\[
T^{\alpha 0}(x^t) = \sum_n p_n^{\alpha}(t) \delta^3(x - x_n(t)) \tag{2.8.1}
\]

and its current is defined by

\[
T^{\alpha i}(x^t) = \sum_n p_n^{\alpha}(t) \frac{dx_n^i(t)}{dt} \delta^3(x - x_n(t)) \tag{2.8.2}
\]

These two definitions can be united into a single formula,

\[
T^{\alpha \beta}(x) = \sum_n p_n^{\alpha} \frac{dx_n^\beta(t)}{dt} \delta^3(x - x_n(t)) \tag{2.8.3}
\]

where \( x_n^{0}(t) \equiv t \). We note from (2.4.10) that

\[
p_n^\beta = E_n \frac{dx_n^\beta}{dt}
\]
so (2.8.3) can also be written as

\[ T^{\alpha\beta}(x) = \sum_{n} \frac{p_{n}^{\alpha} p_{n}^{\beta}}{E_{n}} \delta^{3}(x - x_{n}(t)) \quad (2.8.4) \]

and we see that \( T^{\alpha\beta} \) is symmetric:

\[ T^{\alpha\beta}(x) = T^{\beta\alpha}(x) \quad (2.8.5) \]

We can also write (2.8.3) in analogy with (2.6.5) as

\[ T^{\alpha\beta}(x) = \sum_{n} \int d\tau p_{n}^{\alpha} \frac{dx_{n}^{\beta}}{d\tau} \delta^{4}(x - x_{n}(\tau)) \quad (2.8.5a) \]

and we see that \( T^{\alpha\beta} \) is a tensor, that is,

\[ T^{\alpha\beta} = \Lambda^{\alpha}_{\gamma} \Lambda^{\beta}_{\delta} T^{\gamma \delta} \]

under a Lorentz transformation (2.1.1).

The conservation law for \( T^{\alpha\beta} \) will take a little more thought. Returning to (2.8.1) and (2.8.2) we see that

\[ \frac{\partial}{\partial x^{i}} T^{\gamma i}(x, t) = -\sum_{n} p_{n}^{\gamma}(t) \frac{dx_{n}^{i}(t)}{dt} \frac{\partial}{\partial x_{n}^{i}} \delta^{3}(x - x_{n}(t)) \]

\[ = -\sum_{n} p_{n}^{\gamma}(t) \frac{\partial}{\partial t} \delta^{3}(x - x_{n}(t)) \]

\[ = -\frac{\partial}{\partial t} T^{\gamma 0}(x, t) + \sum_{n} \frac{dp_{n}^{\gamma}(t)}{dt} \delta^{3}(x - x_{n}(t)) \]

and so

\[ \frac{\partial}{\partial x^{\alpha}} T^{\alpha\beta} = G^{\gamma} \quad (2.8.6) \]

where \( G^{\gamma} \) is the density of force.

\[ G^{\gamma}(x, t) \equiv \sum_{n} \delta^{3}(x - x_{n}(t)) \frac{dp_{n}^{\gamma}(t)}{dt} = \sum_{n} \delta^{3}(x - x_{n}(t)) \frac{dx_{n}^{\gamma}(t)}{dt} \frac{d}{dt} f_{n}^{\gamma}(t) \]

If the particles are free, then \( p_{\gamma}^{n} \) is constant and \( T^{\alpha\beta} \) is conserved, that is,

\[ \frac{\partial}{\partial x^{\alpha}} T^{\alpha\beta}(x) = 0 \quad (2.8.7) \]

The same is also true if the particles interact only during collisions that are strictly localized in space. In this case (2.8.6) gives

\[ \frac{\partial}{\partial x^{\alpha}} T^{\alpha\beta}(x) = \sum_{\epsilon} \delta^{3}(x - x_{\epsilon}(t)) \frac{d}{dt} \sum_{n \in \epsilon} p_{n}^{\gamma}(t) \]
where \( x_c(t) \) is the location of the \( c \)th collision going on at time \( t \), and \( n \in c \) means we sum only over the particles participating in the \( c \)th collision. But each collision conserves momentum, so \( \sum_{n \in c} p_n^a(t) \) must be time-independent, yielding the conservation equation (2.8.7).

The energy-momentum tensor (2.8.3) will not be conserved if the particles are subject to forces that act at a distance. For instance, consider a gas of charged particles, with charges \( e_n \). Then (2.8.6), (2.4.1), and (2.7.9) give

\[
\frac{\partial}{\partial x^\theta} T^{\alpha\beta}(x) = \sum_n e_n K_\gamma^\beta(x) \frac{dx_n^\gamma}{dt} \delta^3(x - x_n(t))
\]

and, using (2.6.4), this gives

\[
\frac{\partial}{\partial x^\theta} T^{\alpha\beta}(x) = F^\gamma_\gamma(x) J^\gamma(x) \tag{2.8.8}
\]

Although this is not conserved, we can construct a conserved tensor by adding a purely electromagnetic term

\[
T_{em}^{\alpha\beta} \equiv F^\gamma_\gamma F^\beta_\gamma - \frac{1}{4} \eta^{\gamma\delta} F_{\gamma\theta} F^{\theta\delta} \tag{2.8.9}
\]

That is, the electromagnetic energy and momentum densities are given by

\[
T_{em}^{00} = \frac{1}{2} (E^2 + B^2) \quad T_{em}^{i0} = (E \times B)_j \tag{2.8.10}
\]

We note that

\[
\frac{\partial}{\partial x^\theta} T_{em}^{\alpha\beta} = \frac{\partial}{\partial x^\theta} F^\gamma_\gamma F^\beta_\gamma + F^\delta_\gamma \frac{\partial}{\partial x^\theta} F^\gamma_\delta - \frac{1}{4} \eta^{\gamma\delta} \frac{\partial}{\partial x^\theta} F^{\theta\delta}
\]

[Here \( \partial/\partial x^\theta = \eta^{\gamma\delta}(\partial/\partial x^\gamma). \)] With a little reshuffling of indices, this becomes

\[
\frac{\partial}{\partial x^\theta} T_{em}^{\alpha\beta} = F^\gamma_\gamma \frac{\partial}{\partial x^\theta} F^\beta_\gamma - \frac{1}{4} F^\gamma_\gamma \left( \frac{\partial}{\partial x^\theta} F^\delta_\gamma + \frac{\partial}{\partial x^\delta} F^\gamma_\theta + \frac{\partial}{\partial x^\gamma} F^\theta_\delta \right)
\]

Using the Maxwell equations (2.7.6) and (2.7.10), we find

\[
\frac{\partial}{\partial x^\theta} T_{em}^{\alpha\beta} = -F^\gamma_\gamma J^\gamma \tag{2.8.11}
\]

Comparing (2.8.8) with (2.8.11), we are led to redefine the energy-momentum tensor as

\[
T^{\alpha\beta} = \sum_n p_n^a \frac{dx_n^\beta}{dt} \delta^3(x - x_n(t)) + T_{em}^{\alpha\beta} \tag{2.8.12}
\]

This is again a symmetric tensor, and is now conserved

\[
\partial^\alpha T^{\alpha\beta} = 0 \tag{2.8.13}
\]
We can continue to add more and more terms to \( T^{x}{}^{y} \) to account for other fields and keep \( T^{x}{}^{y} \) conserved. A systematic method for constructing these terms is presented in Chapter 12.

Just as the integral of the charge density \( J^{0} \) is the total charge, the integral of the density \( T^{x}{}^{0} \) of \( p^{x} \) is the total \( p^{x} \):

\[
p^{x}_{\text{total}} = \int d^{3}x T^{x}{}^{0}(x, t) \quad (2.8.14)
\]

That this is a constant four-vector can be shown in the same way that we showed in Section 6 that the total charge \((2.6.7)\) is a constant scalar.

9 Spin

One important use we can make of the energy-momentum tensor \( T^{x}{}^{y} \) is to define angular momentum and spin. Consider first an isolated system, for which the total energy-momentum tensor \( T^{x}{}^{y} \) is conserved

\[
\frac{\partial}{\partial x^{y}} T^{y}{}^{y} = 0
\]

We can use \( T \) to construct another tensor,

\[
M^{x}{}^{y} \equiv x^{x}T^{y}{}^{y} - x^{y}T^{x}{}^{y}
\]

and because \( T \) is conserved and symmetric, \( M \) is also conserved:

\[
\frac{\partial M^{x}{}^{y}}{\partial x^{y}} = T^{x}{}^{y} - T^{y}{}^{x} = 0 \quad (2.9.2)
\]

We can then form a total angular momentum

\[
J^{x}{}^{y} = \int d^{3}x M^{0}{}^{x} = -J^{x}{}^{0} \quad (2.9.3)
\]

From (2.9.2) we see (by following the arguments of the last section) that \( J^{x}{}^{y} \) is constant in time and is a tensor. We further note that

\[
J^{ij} = \int d^{3}x (x^{i}T^{j}{}^{0} - x^{j}T^{i}{}^{0})
\]

and since \( T^{j}{}^{0} \) is the density of the \( j \)th component of momentum, we may regard \( J^{x}{}^{0}, J^{x}{}^{1}, \) and \( J^{x}{}^{2} \) as the 1-, 2-, and 3-components of the angular momentum. The other components of \( J^{x}{}^{y} \) are

\[
J^{0}{}^{i} = tp^{i} - \int x^{i}T^{00} \, d^{3}x
\]
These components have no clear physical significance, and in fact can be made to vanish if we fix the origin of coordinates to coincide with the “center of energy” at \( t = 0 \), that is, if at \( t = 0 \) the moment \( \int x^i T^{00} d^3x \) vanishes.

Although a tensor with regard to the homogeneous Lorentz transformations \( x^\alpha \to \Lambda^\alpha_\beta x^\beta \), the total angular momentum behaves peculiarly under the translation \( x^\alpha \to x'^\alpha = x^\alpha + a^\alpha \). From (2.9.3) and (2.8.13) we find that

\[
J^{\alpha \beta} \to J'^{\alpha \beta} = J^{\alpha \beta} + \alpha^\alpha p^\beta - \alpha^\beta p^\alpha
\]

(2.9.4)

This is of course because \( J^{\alpha \beta} \) includes the orbital angular momentum, which is always defined with respect to some center of rotation. In order to isolate the internal part of \( J^{\alpha \beta} \), it is convenient to define a spin four-vector

\[
S_\alpha \equiv \frac{1}{2} \epsilon_{\alpha \beta \gamma \delta} J^{\beta \gamma} U^\delta
\]

(2.9.5)

where \( \epsilon_{\alpha \beta \gamma \delta} \) is the completely antisymmetric tensor discussed in Section 5, and \( U^\alpha \equiv p^\alpha/(-p_\mu p^\mu)^{1/2} \) is the four-vector velocity of the system. Because of the antisymmetry of \( \epsilon_{\alpha \beta \gamma \delta} \), the translation \( x^\alpha \to x'^\alpha = x^\alpha + a^\alpha \), which changes \( J^{\alpha \beta} \) according to the rule (2.9.4), does not change \( S_\alpha \). Furthermore, \( S_\alpha \) is obviously a vector and is constant for a free particle

\[
\frac{dS_\alpha}{dt} = 0
\]

(2.9.6)

Finally we note that in the center-of-mass frame of the system \( U^i = 0 \) and \( U^0 = 1 \), so in this frame

\[
S_1 = J^{23}, \quad S_2 = J^{31}, \quad S_3 = J^{12}, \quad S_0 = 0
\]

(2.9.7)

This justifies us in regarding \( S_\alpha \) as the internal angular momentum of the system. Even when the velocity \( U \) is not zero, \( S_\alpha \) really has only three independent components, because (2.9.5) gives

\[
U^\alpha S_\alpha = 0
\]

(2.9.8)

We use these properties of \( S_\alpha \) later, when we discuss the precession of a gyroscope in free fall.

10 Relativistic Hydrodynamics

A great many macroscopic physical systems, including perhaps the universe itself, may be approximately regarded as perfect fluids. A perfect fluid is defined as having at each point a velocity \( \mathbf{v} \), such that an observer moving with this velocity sees the fluid around him as isotropic. This will be the case if the mean free path between collisions is small compared with the scale of lengths used by the observer. (For instance, a sound wave will propagate in air if its wavelength is large compared
with the mean free path, but at very short wavelengths viscosity becomes important and the air stops acting like a perfect fluid. We shall translate the above definition of a perfect fluid into a statement about the energy-momentum tensor. First suppose that we are in a reference frame (distinguished by a tilde) in which the fluid is at rest at some particular position and time. At this space-time point, the perfect fluid hypothesis tells us that the energy-momentum tensor takes the form characteristic of spherical symmetry:

\[ T^{ij} = p \delta_{ij} \]  \hspace{1cm} (2.10.1)
\[ T^{i0} = T^{0i} = 0 \]  \hspace{1cm} (2.10.2)
\[ T^{00} = \rho \]  \hspace{1cm} (2.10.3)

The coefficients \( p \) and \( \rho \) are called the pressure and the proper energy density, respectively. Now go into a reference frame at rest in the laboratory, and suppose that the fluid in this frame appears to be moving (at the given space-time point) with velocity \( \mathbf{v} \). The connection between the comoving coordinates \( \tilde{x}^\beta \) and the lab coordinates \( x^\alpha \) is then

\[ x^\alpha = \Lambda^\alpha_\beta(\mathbf{v}) \tilde{x}^\beta \]

with \( \Lambda^\alpha_\beta(\mathbf{v}) \) the "boost" defined by Eqs. (2.1.17)-(2.1.21). But \( T^{\alpha\beta} \) is a tensor, so in the lab frame it is

\[ T^{\alpha\beta} = \Lambda^\alpha_\gamma(\mathbf{v}) \Lambda^\beta_\delta(\mathbf{v}) \tilde{T}^{\gamma\delta} \]

or explicitly

\[ T^{ij} = p \delta_{ij} + (p + \rho) \frac{\mathbf{v}_i \mathbf{v}_j}{1 - \mathbf{v}^2} \]  \hspace{1cm} (2.10.4)
\[ T^{i0} = (p + \rho) \frac{v_i}{1 - \mathbf{v}^2} \]  \hspace{1cm} (2.10.5)
\[ T^{00} = \frac{(p + \rho \mathbf{v}^2)}{1 - \mathbf{v}^2} \]  \hspace{1cm} (2.10.6)

To check that this is a tensor, we note that (2.10.4)-(2.10.6) can be integrated into a single equation:

\[ T^{\alpha\beta} = \rho \eta^{\alpha\beta} + (p + \rho) U^\alpha U^\beta \]  \hspace{1cm} (2.10.7)

where \( U^\alpha \) is the velocity four-vector.

\[ U = \frac{dx}{dt} = (1 - \mathbf{v}^2)^{-1/2} \mathbf{v} \]  \hspace{1cm} (2.10.8)
\[ U^0 = \frac{dt}{d\tau} = (1 - \mathbf{v}^2)^{-1/2} \]

normalized so that

\[ U_\alpha U^\alpha = -1 \]  \hspace{1cm} (2.10.9)
Indeed, Eq. (2.10.7) could have been derived very easily by noting that the quantity on the right-hand side is a tensor, which equals the tensor \( T^{\alpha\beta} \) in a Lorentz frame moving with the fluid, and hence must equal \( T^{\alpha\beta} \) in all Lorentz frames.

Apart from energy and momentum, a fluid will in general carry one or more conserved quantities, such as the charge, the number of baryons minus the number of antibaryons, or, at normal temperatures, the number of atoms. Let us consider one such conserved quantity, and refer to it for brevity as the "particle number." If \( n \) is the particle number density in a Lorentz frame that moves with the fluid at a given space-time point, then in this frame the particle current four-vector at this point is

\[
\vec{N}^i = 0 \quad \vec{N}^0 = n \tag{2.10.10}
\]

In any other Lorentz frame, in which the fluid at this point moves with velocity \( \mathbf{v} \), the particle current is related to (2.10.10) by the "boost" \( \Lambda(\mathbf{v}) \):

\[
N^i = \Lambda^i_\beta(\mathbf{v}) \vec{N}^\beta = (1 - \mathbf{v}^2)^{-1/2} v^i n \tag{2.10.11}
\]

\[
N^0 = \Lambda^0_\beta(\mathbf{v}) \vec{N}^\beta = (1 - \mathbf{v}^2)^{-1/2} n \tag{2.10.12}
\]

or, more concisely,

\[
N^a = n U^a \tag{2.10.13}
\]

The motion of the fluid will be governed by the equations of conservation of energy and momentum,

\[
0 = \frac{\partial T^{a\beta}}{\partial x^\beta} = \frac{\partial p}{\partial x^\beta} + \frac{\partial}{\partial x^\beta} [(\rho + p)U^a U^\beta] \tag{2.10.14}
\]

and of the particle number:

\[
0 = \frac{\partial N^a}{\partial x^a} = \frac{\partial}{\partial x^a} (n U^a) = \frac{\partial}{\partial t} (n(1 - \mathbf{v}^2)^{-1/2}) + \mathbf{v} \cdot \nabla (n(1 - \mathbf{v}^2)^{-1/2}) \tag{2.10.15}
\]

It is convenient to write (2.10.14) as separate three-vector and scalar equations. The three-vector equation is obtained by setting \( a = i \) in Eq. (2.10.14), writing \( U^i = v^i U^0 \), and then using Eq. (2.10.14) with \( a = 0 \); this gives

\[
\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla) \mathbf{v} = -\frac{(1 - \mathbf{v}^2)}{(\rho + p)} \left[ \nabla p + \mathbf{v} \frac{\partial \rho}{\partial t} \right] \tag{2.10.16}
\]

The scalar equation is obtained by multiplying Eq. (2.10.14) by \( U^a \); using the relation

\[
0 = \frac{\partial}{\partial x^a} (U_a U^a) = 2 U^a \frac{\partial U^a}{\partial x^a} \tag{2.10.17}
\]
we then have

\[0 = U^a \frac{\partial T^{a\theta}}{\partial x^\theta} = U^a \frac{\partial p}{\partial x^\theta} - \frac{\partial}{\partial x^\theta} \left[(p + \rho)U^\theta\right]\]

Using Eq. (2.10.15), we can write this as

\[0 = -nU^\theta \left[p \frac{\partial}{\partial x^\theta} \left(\frac{1}{n}\right) + \frac{\partial}{\partial x^\theta} \left(\frac{\rho}{n}\right)\right]\]

(2.10.17a)

The second law of thermodynamics tells us that the pressure \(p\), the energy density \(\rho\), and the volume per particle \(1/n\) may be expressed as functions of the temperature \(T\) and the entropy per particle \(\sigma k\), in such a way that

\[kT \, d\sigma = p d \left(\frac{1}{n}\right) + d \left(\frac{\rho}{n}\right)\]

(2.10.18)

(Boltzmann’s constant \(k\) is introduced here to make \(\sigma\) dimensionless.) Our scalar equation (2.10.17a) can now be written

\[0 = U^\theta \frac{\partial \sigma}{\partial x^\theta} + \frac{\partial \sigma}{\partial t} + (\nabla \cdot \nabla)\sigma\]

(2.10.19)

The specific entropy \(\sigma\) is therefore constant in time at any point that moves along with the fluid. The fundamental equations of relativistic hydrodynamics are the “continuity equation” (2.10.15), the “Euler equations” (2.10.16), the “energy equation” (2.10.19), together with equations of state that give \(p\) and \(\rho\) in terms of \(n\) and \(\sigma\).

In order to gain some insight into the possible equations of state, we may consider a fluid composed of structureless point particles that interact only in spatially localized collisions. As shown in Section 2.8, the energy-momentum tensor is

\[T^{a\theta} = \sum_N \frac{p_N x^\theta}{E_N} \delta^3(x - x_N)\]

(2.10.20)

[See Eq. (2.8.4.)] In a comoving Lorentz frame, \(T^{a\theta}\) will have the isotropic form (2.10.1)–(2.10.3), so the pressure and energy density will be given in this frame by

\[p = \frac{1}{3} \sum_{i=1}^3 T_{ii} = \frac{1}{3} \sum_N \frac{p_N x^\theta}{E_N} \delta^3(x - x_N)\]

\[\rho = T^{00} = \sum_N \frac{E_N \delta^3(x - x_N)}{\sum_N E_N}\]

(2.10.21)

(2.10.22)
whereas the particle number density is, in analogy with (2.6.2),

\[ n = \sum_N \delta^3(x - x_N) \]  

(2.10.23)

It follows that in general

\[ 0 \leq p \leq \frac{\rho}{3} \]  

(2.10.24)

For a cool, nonrelativistic gas, we can approximate

\[ E_N \sim m + \frac{p_N^2}{2m} \]

so (2.10.22) gives

\[ \rho \simeq nm + \frac{2}{3}p \]  

(2.10.25)

For a hot, extremely relativistic gas, we have

\[ E_N \sim |p_N| \gg m \]

so (2.10.22) gives

\[ \rho \simeq 3p \gg nm \]  

(2.10.26)

Both (2.10.25) and (2.10.26) can be incorporated into a single equation,

\[ \rho - nm \simeq (\gamma - 1)^{-1}p \]  

(2.10.27)

with

\[ \gamma = \begin{cases} 
\frac{5}{3} & \text{nonrelativistic} \\
\frac{4}{3} & \text{extreme relativistic} 
\end{cases} \]  

(2.10.28)

Equation (2.10.18) then gives

\[ kT \, d\sigma = \rho d \left( \frac{1}{n_j} \right) + (\gamma - 1)^{-1} d \left( \frac{p}{n_j} \right) = \frac{n_j^{\gamma - 1}}{\gamma - 1} d \left( \frac{p}{n_j^{\gamma}} \right) \]  

(2.10.29)

Thus Eq. (2.10.19) takes the form

\[ 0 = \frac{\partial}{\partial t} \left( \frac{p}{n \gamma} \right) + (\mathbf{v} \cdot \nabla) \left( \frac{p}{n \gamma} \right) \]  

(2.10.30)

and (2.10.27) is to be used to express \( \rho \) in terms of \( n \) and \( p \) in Eq. (2.10.16). The proportionality expressed in Eq. (2.10.27) between internal energy and pressure actually holds, with various values of \( \gamma \), over a class of fluids much wider than the simple gas of point particles discussed here. For all such fluids, the energy equation can be put in the form (2.10.30).

As an example, let us calculate the speed of sound in a static homogeneous relativistic fluid. In the unperturbed state, we have \( n, \rho, p \), and \( \sigma \) constant in space.
and time, and \( v = 0 \). The sound wave produces small changes \( n_1, \rho_1, p_1, \) and \( v_1 \) in \( n, \rho, p, \) and \( v \). but according to (2.10.19), it leaves \( \sigma \) unchanged. To first order in small quantities. Eqs. (2.10.15) and (2.10.16) then read

\[
\frac{\partial n_1}{\partial t} + n \mathbf{v} \cdot \mathbf{v}_1 = 0
\]

\[
\frac{\partial v_1}{\partial t} = -\frac{v p_1}{p + \rho}
\]

But with \( d\sigma = 0 \), Eq. (2.10.18) gives

\[
0 = \frac{(p + \rho)}{n} n_1 + \rho_1
\]

so that

\[
\frac{\partial v_1}{\partial t} = -\frac{v_s^2 n_1}{n}
\]

where

\[
v_s^2 = \frac{p_1}{\rho_1} = \left(\frac{\partial p}{\partial \rho}\right)_{\text{const}}
\]

Combining the equations for \( n_1 \) and \( v_1 \), we obtain a wave equation

\[
0 = \left[ \frac{\partial^2}{\partial t^2} - \frac{v_s^2}{2} \nabla^2 \right] n_1
\]

that shows that sound waves travel with the speed \( v_s \), just as in a nonrelativistic fluid. The speed of sound is much less than the speed of light (i.e., unity) for a nonrelativistic fluid, but it increases with temperature, so it is worth checking whether \( v_s \) might exceed unity for a fluid of highly relativistic point particles, such as hydrogen above \( 10^{13} \) °K. In this case, (2.10.26) and (2.10.31) give a sound speed

\[
v_s = \frac{1}{\sqrt{3}}
\]

which is still safely less than unity. This conclusion would not be affected if electromagnetic forces were taken into account, because Eqs. (2.10.7) and (2.8.9) impose on the electromagnetic pressure \( p_{em} \) and energy density \( \rho_{em} \) the relation

\[
0 = T_{em}^x = 3p_{em} - \rho_{em}
\]

so the inclusion of \( p_{em} \) and \( \rho_{em} \) would not invalidate (2.10.26) or (2.10.32). It is an open question whether \( v_s \) remains less than unity when nonelectromagnetic forces are taken into account.\(^4\)
11 Relativistic Imperfect Fluids

The last section dealt with a perfect fluid, in which mean free paths and times are so short that perfect isotropy is maintained about any point moving with the fluid. In practice, one often has to deal with somewhat imperfect fluids, in which the pressure, density, or velocity vary appreciably over distances of the order of a mean free path, or over times of the order of a mean free time, or both. In such fluids, thermal equilibrium is not strictly maintained, and the fluid kinetic energy is dissipated as heat.

The correct treatment of dissipative effects for relativistic fluids raises certain delicate questions of principle, which do not arise in the nonrelativistic case. For this reason, and also because dissipation plays an increasingly important role in theories of the early universe (see Sections 10.8, 10.10, 15.11), it will be worth our while here to develop the outlines of the general theory of relativistic imperfect fluids.

We suppose that the presence of weak space-time gradients in an imperfect fluid has the effect of modifying the energy-momentum tensor and particle current vector by terms $\Delta T^{\alpha \beta}$ and $\Delta N^\alpha$, which are of first order in these gradients. Instead of (2.10.7) and (2.10.13), we then have

\begin{align}
T^{\alpha \beta} &= \rho n^{\alpha \beta} + (\rho + p)U^\alpha U^\beta + \Delta T^{\alpha \beta} \\
N^\alpha &= n U^\alpha + \Delta N^\alpha
\end{align}

Once we allow such correction terms, the definitions of the pressure $p$, energy density $\rho$, particle density $n$, and fluid velocity $U^\alpha$ become somewhat ambiguous. The general practice is to define $\rho$ and $n$ as the total energy density and particle number density in a comoving frame:

\begin{align}
T^{00} &\equiv \rho \\
N^0 &\equiv n
\end{align}

a comoving frame being characterized by the condition that at a given point, the velocity four-vector is

\begin{align}
U^t &\equiv 0 \\
U^0 &\equiv 1
\end{align}

In addition, the pressure $p$ is generally defined to be the same function of $\rho$ and $n$ [e.g., (2.10.27)] as in the case where all fluid gradients are negligible and dissipation is absent. Finally, it is necessary in a relativistic fluid to specify whether $U^t$ is the velocity of energy transport or particle transport. In the approach of Landau and Lifshitz, $U^t$ is taken to be the velocity of energy transport, so that $T^{0t}$ vanishes in a comoving frame. In the approach of Eckart, $U^t$ is taken to be the velocity of

*This section lies somewhat out of the book’s main line of development, and may be omitted in a first reading.
particle transport, so that it is \( N^i \) that vanishes in a comoving frame. The two approaches are perfectly equivalent, but Eckart's seems to me to be slightly more convenient, and will be adopted here. With this definition of \( U^a \), we then have in a comoving frame

\[
N^i = 0 \quad (2.11.6)
\]

A comparison of (2.11.3)-(2.11.6) with (2.11.1) and (2.11.2) shows that in a comoving frame, the dissipative terms \( \Delta T^{ab} \) and \( \Delta N^a \) are subject to the constraints

\[
\Delta T^{00} = \Delta N^0 = \Delta N^i = 0 \quad (2.11.7)
\]

and therefore, in a general Lorentz frame,

\[
U^a U^b \Delta T_{ab} = 0 \quad (2.11.8)
\]

\[
\Delta N^a = 0 \quad (2.11.9)
\]

All effects of dissipation thus show up as contributions to \( \Delta T^{ab} \). Our task is now to construct the most general possible dissipative tensor \( \Delta T^{ab} \) allowed by Eq. (2.11.8) and by the second law of thermodynamics.

To this end, let us calculate the entropy produced by fluid motions. As in the last sections, we start by contracting the conservation law (2.8.7) with \( U_a \):

\[
0 = U_z \frac{\partial}{\partial x^z} T^{ab} \quad (2.11.10)
\]

By following the same reasoning that was used to derive (2.10.10) for a perfect fluid, one sees that in general

\[
U_z \frac{\partial}{\partial x^a} [p n^{xb} + (p + \rho) U^a U^b] = -kT \frac{\partial}{\partial x^z} (\sigma U^z)
\]

where \( T \) and \( \sigma k \) are the temperature and entropy per particle, defined by Eq. (2.10.18). Hence (2.11.10) now reads

\[
\frac{\partial}{\partial x^z} (\sigma U^z) - \frac{1}{kT} U_z \frac{\partial}{\partial x^z} \Delta T^{ab}
\]

or equivalently

\[
\frac{\partial S^a}{\partial x^z} \frac{\partial}{\partial x^z} - \frac{1}{T} \frac{\partial U_z}{\partial x^z} \Delta T^{ab} + \frac{1}{T^2} \frac{\partial T}{U_z \Delta T^{ab}} \quad (2.11.11)
\]

where

\[
S^a = \sigma k U^a \quad (2.11.12)
\]

The entropy density in a comoving frame is \( \sigma k - S^0 \), so we may interpret \( S^a \) as the entropy current four-vector, and Eq. (2.11.11) thus gives the rate of entropy
production per unit volume. The second law of thermodynamics then requires that $\Delta T^{a\beta}$ be a linear combination of velocity and temperature gradients, such that the right-hand side of (2.11.11) is positive for all possible fluid configurations. Note that this is only possible because we have included the second term in Eq. (9.11.19); without this term, $\partial S^a/\partial x^a$ would not be simply quadratic in first derivatives, and hence could not be positive for all fluid configurations. Note also that $\Delta T^{a\beta}$ is not allowed to involve gradients of $p$, $\rho$, $n$, and so on, because if it did then (2.11.11) would contain products of pressure or density gradients with velocity or temperature gradients, and, again, these products would not be positive for all fluid configurations.

It is convenient at this point to go over to a comoving frame, in which $U^a$ has the form (2.11.5) at a given space-time point $P$. From (2.10.17), it follows that in this frame, all gradients of $U^0$ vanish at $P$. Setting $U^i$, $\partial U^a/\partial x^a$, and $\Delta T^{00}$ equal to zero in Eq. (2.11.11), we find that in a Lorentz frame comoving at $P$, the rate of entropy production per unit volume at $P$ is

$$\frac{\partial S^a}{\partial x^a} = \left(1 - \frac{T}{U_i} \right) \Delta T^{i0} - \frac{1}{T} \frac{\partial U_i}{\partial x^i} \Delta T^{ij}$$

(2.11.13)

In order for this to be positive for all possible fluid configurations, we must have

$$\Delta T^{i0} = -\chi \left( \frac{\partial T}{\partial x^i} + T \mathbf{U}_i \right)$$

(2.11.14)

$$\Delta T^{ij} = -\eta \left( \frac{\partial U_i}{\partial x^j} + \frac{\partial U_j}{\partial x^i} - \frac{2}{3} \mathbf{V} \cdot \mathbf{U} \delta_{ij} \right) - \zeta \mathbf{V} \cdot \mathbf{U} \delta_{ij}$$

(2.11.15)

with positive coefficients

$$\chi \geq 0, \quad \eta \geq 0, \quad \zeta \geq 0$$

(2.11.16)

so that (2.11.13) reads

$$\frac{\partial S^a}{\partial x^a} - \frac{\chi}{T^2} \left( \mathbf{V} \cdot \mathbf{V} + T \mathbf{U} \right)^2$$

$$+ \frac{\eta}{2T} \left( \frac{\partial U_i}{\partial x^j} + \frac{\partial U_j}{\partial x^i} - \frac{2}{3} \delta_{ij} \mathbf{V} \cdot \mathbf{U} \right) \left( \frac{\partial U_i}{\partial x^j} + \frac{\partial U_j}{\partial x^i} - \frac{2}{3} \delta_{ij} \mathbf{V} \cdot \mathbf{U} \right)$$

$$+ \frac{\zeta}{T} \left( \mathbf{V} \cdot \mathbf{U} \right)^2 \geq 0$$

(2.11.17)

Except for the relativistic correction $T \mathbf{U}$ in (2.11.14), the form of (2.11.14) and (2.11.15) is the same as in the nonrelativistic theory of imperfect fluids, and we
therefore may identify $\chi$, $\eta$, and $\zeta$ as the coefficients of heat conduction, shear viscosity, and bulk viscosity.

It now only remains to translate our results from the forms (2.11.5), (2.11.7), (2.11.14), (2.11.15), which are valid only in comoving frames, to forms valid in general Lorentz frames. Let us define a shear tensor,

$$W_{\alpha\beta} \equiv \frac{\partial U_{\beta}}{\partial x^\alpha} + \frac{\partial U_{\alpha}}{\partial x^\beta} - \frac{2}{3} \eta_{\alpha\beta} \frac{\partial U^\gamma}{\partial x^\gamma}$$

(2.11.18)

a heat flow vector,

$$Q_\alpha = \frac{\partial T}{\partial x^\alpha} \left[ T \frac{\partial U_\alpha}{\partial x^\beta} U^\beta \right]$$

(2.11.19)

and a projection tensor on the hyperplane normal to $U^\alpha$:

$$H_{\alpha\beta} \equiv \eta_{\alpha\beta} + U_\alpha U_\beta$$

(2.11.20)

It is straightforward to check that in a comoving Lorentz frame, our formulas (2.11.1), (2.11.14), (2.11.16) for $\Delta T^\alpha$ are satisfied by the tensor

$$\Delta T^\alpha = -\eta H^{\alpha\gamma} H^{\beta\delta} W_{\gamma\delta}$$

$$-\chi (H^{\alpha\gamma} U^\beta + H^{\beta\gamma} U^\alpha) Q_\gamma - \zeta H^{\alpha\beta} \frac{\partial U^\gamma}{\partial x^\gamma}$$

(2.11.21)

Since this formula is Lorentz-invariant, and valid in a comoving Lorentz frame, it is valid in all Lorentz frames.

In general, the coefficients $\chi$, $\eta$, and $\zeta$ might be expected on dimensional grounds to be of the order of the pressure, or the thermal energy density, times some sort of mean free time. However, there are important special cases in which the bulk viscosity $\zeta$ is much smaller than $\eta$ or $\chi T$. To see when this applies, note that (2.11.1) and (2.11.21) give the trace of the total energy-momentum tensor as

$$T^\alpha_\alpha = 3p - \rho - 3\zeta \frac{\partial U^\gamma}{\partial x^\gamma}$$

(2.11.22)

Suppose that we are dealing with a medium for which this trace can be expressed as a function of $\rho$ and $n$ alone:

$$T^\alpha_\alpha = f(\rho, n)$$

(2.11.23)

For instance, for the simple gas characterized by (2.10.20), this trace is

$$T^\alpha_\alpha = -\sum \frac{m^2}{E_N} \delta^3(x - x_N)$$
In the extreme relativistic case, we have \( E_N \gg m \), so in this case (2.11.23) is satisfied, with
\[
f(\rho, n) \simeq 0
\]

In the nonrelativistic case, we have
\[
\frac{1}{E_N} \simeq \frac{1}{m} - \left( \frac{E_N - m}{m^2} \right)
\]
so in this case (2.11.23) is again satisfied, with
\[
f(\rho, n) \simeq -mn \mid (\rho - mn)
\]

In the absence of velocity gradients, Eqs. (2.11.22) and (2.11.23) would give a formula for the pressure
\[
p = \frac{1}{3} \left[ \rho + f(\rho, n) \right]
\]
(2.11.24)

But we have agreed to define \( p \) in general as the same function of \( \rho \) and \( n \) as in the absence of dissipation, so (2.11.24) must hold even in the presence of velocity gradients, and therefore (2.11.22), (2.11.23), and (2.11.24) give
\[
\zeta = 0
\]
(2.11.25)

It would be wrong, however, to conclude that \( \zeta \) is generally negligible. As we have seen, the trace of the energy-momentum tensor for a simple gas is a function of \( \rho \) and \( n \) only in the extreme relativistic or extreme nonrelativistic limit; for \( kT \) of order \( m \), \( T_{\text{rad}} \) cannot be expressed in the form (2.11.23), and the bulk viscosity is of the same order as the shear viscosity. The bulk viscosity is also important in a fluid that allows an easy exchange of energy between translational and internal degrees of freedom, as in the case of a gas of rough spheres. Another case, of particular importance to cosmology, is that of a material medium with very short mean free times, interacting with radiation quanta with a finite mean free time \( \tau \). In this case, the coefficients of heat conduction, shear viscosity, and bulk viscosity are calculated to be
\[
\chi = \frac{3}{2} a T^3 \tau
\]
(2.11.26)
\[
\eta = \frac{1}{4} a T^4 \tau
\]
(2.11.27)
\[
\zeta = 4 a T^4 \tau \left[ \frac{1}{3} - \left( \frac{\partial p}{\partial \rho} \right)_{\text{rad}} \right]^2
\]
(2.11.28)

where \( a \) is the Stefan-Boltzmann constant, defined so that the radiation energy density is \( a T^4 \), and \( p \) and \( \rho \) are the total pressure and energy density of the matter and radiation. Note that, in general, \( \chi, \eta, \) and \( \zeta \) are comparable, but if the pressure and thermal energy are dominated by radiation, then \( \left( \frac{\partial p}{\partial \rho} \right)_{\text{rad}} \simeq \frac{1}{3} \) and, as expected, the bulk viscosity will be small.
12 Representations of the Lorentz Group*

The tensor formalism described in Section 2.5 is perfectly adequate for handling the problems of relativistic classical physics. However, there are certain formal advantages in looking at these Lorentz transformation rules in a more general way, from the perspective of the theory of the representations of the homogeneous Lorentz group. We shall see in Section 12.5, that this approach allows an elegant reformulation of the effects of gravitation on arbitrary physical systems. Also, it is only in this way that we can deal with fields of half-integer spin.

Under the general Lorentz transformation rule, a set of quantities $\psi_m$ transform under a Lorentz transformation $\Lambda^\mu_\rho$ into the new quantities:

$$\psi'_m = \sum_n [D(\Lambda)]_{mn}\psi_n \quad (2.12.1)$$

In order for a Lorentz transformation $\Lambda_1$ followed by a Lorentz transformation $\Lambda_2$, to give the same result as the Lorentz transformation $\Lambda_1^2\Lambda_2$, it is necessary that the matrices $D(\Lambda)$ should furnish a representation of the Lorentz group, that is,

$$D(\Lambda_1)D(\Lambda_2) = D(\Lambda_1^2\Lambda_2) \quad (2.12.2)$$

with matrix multiplication now understood. For instance, if $\psi$ is a contravariant vector $V^\rho$, then $D(\Lambda)$ is simply

$$[D(\Lambda)]^\rho_\sigma = \Lambda^\rho_\sigma \quad (2.12.3)$$

whereas for a covariant tensor $T_{\alpha\beta}$ the corresponding $D$-matrix is

$$[D(\Lambda)]_{\alpha\beta} = \Lambda_{\alpha}^{\gamma} \Lambda_{\beta}^{\delta} \quad (2.12.4)$$

It is easy to check that (2.12.3) and (2.12.4) do satisfy the group multiplication rule (2.12.2). We can compile a catalogue of all possible Lorentz transformation rules by constructing the most general representation of the homogeneous Lorentz group.

In fact, the most general true representations of the homogeneous Lorentz group are provided by the tensor representations, such as (2.12.3) and (2.12.4), so we might expect that all quantities of physical interest should be tensors. However, there are additional representations of the infinitesimal Lorentz group, the spinor representations, that play an important role in relativistic quantum field theory. The infinitesimal Lorentz group consists of Lorentz transformations infinitesimally close to the identity, that is,

$$\Lambda^\rho_\sigma = \delta^\rho_\sigma + \omega^\rho_\sigma \quad (2.12.5)$$

$$|\omega^\rho_\sigma| \ll 1$$

* This section lies somewhat out of the book's main line of development, and may be omitted in a first reading.
12 Representations of the Lorentz Group

In order for this to satisfy the fundamental condition (2.1.2) for a Lorentz transformation, we must have

$$ (\delta_\gamma^\epsilon + \omega_\gamma^\epsilon) (\delta_\delta^\beta + \omega_\delta^\beta) = \eta_{\epsilon \delta} $$

or, to first order in \( \omega \),

$$ \omega_{\gamma \delta} = -\omega_{\delta \gamma} \quad (2.12.6) $$

with the indices on \( \omega \) of course lowered with \( \eta \):

$$ \omega_{\gamma \delta} \equiv \eta_{\gamma \rho} \omega_{\rho \delta} $$

For such a transformation, the matrix representation \( D(\Lambda) \) must be infinitesimally close to the identity

$$ D(1 + \omega) = 1 + \frac{1}{2} \omega_{\beta \delta} \sigma_{\alpha \beta} \quad (2.12.7) $$

where \( \sigma_{\alpha \beta} \) are a fixed set of matrices, which by virtue of (2.12.6) can always be chosen to be antisymmetric in \( \alpha \) and \( \beta \):

$$ \sigma_{\alpha \beta} = -\sigma_{\beta \alpha} \quad (2.12.8) $$

For instance, for the tensor representations (2.12.3) and (2.12.4), we have

$$ \left[ \sigma_{\alpha \beta} \right]_{\gamma \delta}^\epsilon = \delta_\gamma^\epsilon \eta_{\beta \delta} - \delta_\delta^\epsilon \eta_{\gamma \beta} \quad (2.12.9) $$

$$ \left[ \sigma_{\alpha \beta} \right]_{\gamma \delta}^\epsilon = \eta_{\gamma \alpha} \delta_\delta^\epsilon - \eta_{\gamma \beta} \delta_\delta^\epsilon $$

$$ + \eta_{\gamma \epsilon} \delta_\delta^\alpha - \eta_{\delta \epsilon} \delta_\delta^\alpha \gamma $$

$$ = \eta_{\gamma \epsilon} \delta_\delta^\alpha - \eta_{\delta \epsilon} \delta_\delta^\alpha \gamma $$

The matrices \( \sigma_{\alpha \beta} \) are not allowed to be just any set of constant matrices, but must be constrained so that \( D(\Lambda) \) satisfies the group multiplication rule (2.12.2). It is convenient first to apply this rule to the product \( \Lambda [1 + \omega] \Lambda^{-1} \):

$$ D(\Lambda) D(1 + \omega) D(\Lambda^{-1}) = D(1 + \omega \Lambda^{-1}) $$

To zero order in \( \omega \), this simply says that \( 1 - 1 \), whereas to first order, we must equate the coefficients of \( \omega_{\alpha \beta} \) on both sides:

$$ D(\Lambda) \sigma_{\alpha \beta} D(\Lambda^{-1}) = \sigma_{\gamma \delta} \Lambda^\gamma_{\alpha} \Lambda^\delta_{\beta} \quad (2.12.10) $$

If we now set \( \Lambda = 1 + \omega \) (not necessarily with the same \( \omega \)) and \( \Lambda^{-1} = 1 - \omega \), then this will be satisfied to first order in \( \omega \) provided that \( \sigma \) satisfies the commutation relations,

$$ [\sigma_{\alpha \beta}, \sigma_{\gamma \delta}] = \eta_{\gamma \delta} \sigma_{\alpha \beta} - \eta_{\gamma \beta} \sigma_{\alpha \delta} + \eta_{\delta \beta} \sigma_{\gamma \alpha} - \eta_{\delta \alpha} \sigma_{\gamma \beta} \quad (2.12.11) $$

with square brackets denoting the usual matrix commutator

$$ [u, v] \equiv uv - vu $$
The reader can easily check that the matrices (2.12.9) and (2.12.10) do satisfy Eq. (2.12.12). The problem of finding the general representations of the infinitesimal homogeneous Lorentz group is thus reduced to the problem of finding all matrices that satisfy the commutation relations (2.12.12).

These commutation relations can be put in a somewhat more familiar form by defining the matrices

\[
\begin{align*}
  a_1 &= \frac{1}{4}[-i\sigma_{23} + \sigma_{10}] \\
  b_1 &= \frac{1}{4}[-i\sigma_{23} - \sigma_{10}] \\
  a_2 &= \frac{1}{4}[-i\sigma_{31} + \sigma_{20}] \\
  b_2 &= \frac{1}{4}[-i\sigma_{31} - \sigma_{20}] \\
  a_3 &= \frac{1}{4}[-i\sigma_{12} + \sigma_{30}] \\
  b_3 &= \frac{1}{4}[-i\sigma_{12} - \sigma_{30}]
\end{align*}
\]

Equation (2.12.12) then takes the form

\[
\begin{align*}
  a \times a &= i a \\
  b \times b &= i b \\
  [a_i, b_j] &= 0
\end{align*}
\] (2.12.14) (2.12.15) (2.12.16)

Equations (2.12.14)–(2.12.16) are simply the commutation relations for a pair of independent angular momentum matrices. The rules for constructing such matrices can be found in any book on nonrelativistic quantum mechanics. In the most general case, \(a\) and \(b\) are a direct sum of "irreducible" components, each characterized by an integer or half-integer \(A\) or \(B\), with

\[
a^2 = A(A + 1) \quad b^2 = B(B + 1)
\] (2.12.17)

and with dimensionality \(2A + 1\) or \(2B + 1\). Thus the most general objects \(\psi_n\), which transform linearly under infinitesimal homogeneous Lorentz transformations, can be decomposed into "irreducible" pieces, characterized by a pair of integers \(A\) and \(B\) or half-integers \((A, B)\), each piece having \((2A + 1)(2B + 1)\) components.

A straightforward calculation shows that the contravariant vector representation (2.12.9), as well as its covariant counterpart, has \(A = B = \frac{1}{4}\). Any tensor representation, such as (2.12.10), can be regarded as a direct product of vector representations, so it consists only of irreducible components with \(A + B\) an integer; for instance, the general second-rank tensor representation (2.12.10) consists of irreducible components with \((A, B)\) equal to \((1, 1), (1, 0), (0, 1), \) and \((0, 0)\). The representations in which \(A + B\) is a half-integer are quite distinct from the tensors and are called spinor representations. The most familiar example is the Dirac electron field, which consists of components with \((A, B)\) equal to \((\frac{1}{2}, 0)\) and \((0, \frac{1}{2})\).

The transformation property of any object under ordinary spatial rotations is determined by its behavior with respect to infinitesimal Lorentz transformations (2.12.5) for which \(\omega_{10} = 0\), and hence by the structure of the purely spatial com-
ponents $\sigma_{12}, \sigma_{23}, \sigma_{31}$ of $\sigma_{sp}$. From these components, we can construct a matrix vector

$$s = a + b = -i\{\sigma_{23}, \sigma_{31}, \sigma_{12}\}$$

(2.12.18)

that according to (2.12.14)-(2.12.16) has the commutation relations of an angular momentum.

$$s \times s = is$$

(2.12.19)

Any irreducible representation $(A, B)$ of the homogeneous Lorentz group can be decomposed\(^{11}\) into pieces with $s^2$ equal to $s(s + 1)$, where $s$ is an integer or half-integer between $|A - B|$ and $A + B$; each term describes excitations (e.g., particles) of spin $s$. It follows then from (2.12.18) that the tensor representations can describe only excitations with integral spin, whereas the spinor representations describe only excitations with half-integer spin.

Finite Lorentz transformations can be built up by multiplying together an infinite number of infinitesimal Lorentz transformations. In the same way, the tensor representations of the infinitesimal Lorentz group can be used to construct the tensor representations, such as (2.12.3) and (2.12.4), of the group of finite Lorentz transformations. However, if we try to construct spinor representations of the finite Lorentz transformations, we find that we can only get "representations up to a sign";\(^{12}\) that is, the group multiplication law (2.12.2) will occasionally have a minus sign on the right-hand side. For instance, the product of two successive $180^\circ$ rotations about a given axis does not give the unit matrix, but minus the unit matrix. The appearance of these minus signs means that a spinor field itself cannot be a physical observable, though even functions of spinors fields can be observables.

13 Temporal Order and Antiparticles*

One of the most striking features of the Lorentz transformations is that they do not leave invariant the order of events. For instance, suppose that in one reference frame an event at $x_2$ is observed to occur later than one at $x_1$, that is, $x_2^0 > x_1^0$. A second observer who sees the first observer moving with velocity $v$ will see the events separated by a time difference

$$x_2^0 - x_1^0 = \Lambda^0_a(v)(x_2^a - x_1^a)$$

where $\Lambda^a_a(v)$ is the "boost" defined by (2.1.17)-(2.1.21). Applying (2.1.17) and (2.1.21) gives then

$$x_2^0 - x_1^0 = \gamma(x_2^0 - x_1^0) + \gamma v \cdot (x_2 - x_1)$$

* This section lies somewhat out of the book's main line of development, and may be omitted in a first reading.
and this will be negative if

$$v \cdot (x_2 - x_1) < -(x_2^0 - x_1^0) \tag{2.13.1}$$

At first sight this might seem to raise the danger of a logical paradox. Suppose that the first observer sees a radioactive decay $A \rightarrow B + C$ at $x_1$, followed at $x_2$ by absorption of particle $B$, for example, $B + D \rightarrow E$. Does the second observer then see $B$ absorbed at $x_2$ before it is emitted at $x_1$? The paradox disappears if we note that the speed $|v|$ characterizing any Lorentz transformation $A(v)$ must be less than unity, so that (2.13.1) can be satisfied only if

$$|x_2^0 - x_1| < |x_2 - x_1| \tag{2.13.2}$$

However, this is impossible, because particle $B$ was assumed to travel from $x_1$ to $x_2$, and (2.13.2) would require its speed to be greater than unity, that is, than the speed of light. To put it another way, the temporal order of events at $x_1$ and $x_2$ is affected by Lorentz transformations only if $x_1 - x_2$ is spacelike, that is,

$$\eta_{\alpha\beta}(x_1 - x_2)^\alpha(x_1 - x_2)^\beta > 0$$

whereas a particle can travel from $x_1$ to $x_2$ only if $x_1 - x_2$ is timelike, that is,

$$\eta_{\alpha\beta}(x_1 - x_2)^\alpha(x_1 - x_2)^\beta < 0$$

Although the relativity of temporal order raises no problems for classical physics, it plays a profound role in quantum theories. The uncertainty principle tells us that when we specify that a particle is at position $x_1$ at time $t_1$, we cannot also define its velocity precisely. In consequence there is a certain chance of a particle getting from $x_1$ to $x_2$ even if $x_1 - x_2$ is spacelike, that is, $|x_1^0 - x_2| > |x_1^0 - x_2^0|$. To be more precise, the probability of a particle reaching $x_2$ if it starts at $x_1$ is nonnegligible as long as

$$(x_1^0 - x_2^0)^2 - (x_1 - x_2)^2 < \frac{\hbar^2}{m^2}$$

where $\hbar$ is Planck’s constant (divided by $2\pi$) and $m$ is the particle mass. (Such space-time intervals are very small even for elementary particle masses; for instance, if $m$ is the mass of a proton then $\hbar/m = 2 \times 10^{-14}$ cm or in time units $6 \times 10^{-23}$ sec. Recall that in our units 1 sec = $3 \times 10^{10}$ cm.) We are thus faced again with our paradox; if one observer sees a particle emitted at $x_1$, and absorbed at $x_2$, and if $(x_1^0 - x_2^0)^2 - (x_1 - x_2)^2$ is positive (but less than $\hbar^2/m^2$), then a second observer may see the particle absorbed at $x_2$ at a time $t_2$ before the time $t_1$ it is emitted at $x_1$.

There is only one known way out of this paradox. The second observer must see a particle emitted at $x_2$ and absorbed at $x_1$. But in general the particle seen by the second observer will then necessarily be different from that seen by the first. For instance, if the first observer sees a proton turn into a neutron and a positive
pi-meson at \( x_1 \) and then sees the pi-meson and some other neutron turn into a proton at \( x_2 \), then the second observer must see the neutron at \( x_2 \) turn into a proton and a particle of negative charge, which is then absorbed by a proton at \( x_1 \) that turns into a neutron. Since mass is a Lorentz invariant, the mass of the negative particle seen by the second observer will be equal to that of the positive pi-meson seen by the first observer. There is such a particle, called a negative pi-meson, and it does indeed have the same mass as the positive pi-meson. This reasoning leads us to the conclusion that for every type of charged particle there is an oppositely charged particle of equal mass, called its antiparticle. Note that this conclusion does not obtain in nonrelativistic quantum mechanics or in relativistic classical mechanics; it is only in relativistic quantum mechanics that antiparticles are a necessity.\(^{13}\) And it is the existence of antiparticles that leads to the characteristic feature of relativistic quantum dynamics, that given enough energy we can create arbitrary numbers of particles and their antiparticles.

2 BIBLIOGRAPHY

Special Relativity

For a more comprehensive treatment of special relativity, see any one of the following books:


Relativistic Hydrodynamics


Representations of the Lorentz Group

2 REFERENCES

11. See, for example, L. I. Schiff, Quantum Mechanics (3rd ed., McGraw-Hill, New York, 1968), Section 27.