Vector analysis in special relativity

2.1 Definition of a vector

For the moment we will use the notion of a vector that we carry over from Euclidean geometry, that a vector is something whose components transform as do the coordinates under a coordinate transformation. Later on we shall define vectors in a more satisfactory manner.

The typical vector is the displacement vector, which points from one event to another and has components equal to the coordinate differences:

\[ \Delta \vec{x} \rightarrow (\Delta t, \Delta x, \Delta y, \Delta z). \] (2.1)

In this line we have introduced several new notations: an arrow over a symbol denotes a vector (so that \( \vec{x} \) is a vector having nothing particular to do with the coordinate \( x \)), the arrow after \( \Delta \vec{x} \) means ‘has components’, and the \( O \) underneath it means ‘in the frame \( O \)’; the components will always be in the order \( t, x, y, z \) (equivalently, indices in the order 0, 1, 2, 3). The notation \( \rightarrow_{O} \) is used in order to emphasize the distinction between the vector and its components. The vector \( \Delta \vec{x} \) is an arrow between two events, while the collection of components is a set of four coordinate-dependent numbers. We shall always emphasize the notion of a vector (and, later, any tensor) as a geometrical object: something which can be defined and (sometimes) visualized without referring to a specific coordinate system. Another important notation is

\[ \Delta \vec{x} \rightarrow_{O} \{\Delta x^{\alpha}\}, \] (2.2)

where by \( \{\Delta x^{\alpha}\} \) we mean all of \( \Delta x^{0}, \Delta x^{1}, \Delta x^{2}, \Delta x^{3} \). If we ask for this vector’s components in another coordinate system, say the frame \( \bar{O} \), we write

\[ \Delta \vec{x} \rightarrow_{\bar{O}} \{\Delta x^{\tilde{\alpha}}\}. \]

That is, we put a bar over the index to denote the new coordinates. The vector \( \Delta \vec{x} \) is the same, and no new notation is needed for it when the frame is changed. Only the
components of it change. What are the new components $\Delta \vec{x}$? We get them from the Lorentz transformation:

$$\Delta \vec{x} = \frac{\Delta x^0}{\sqrt{1 - v^2}} - \frac{v \Delta x^1}{\sqrt{1 - v^2}}$$

Since this is a linear transformation, it can be written

$$\Delta \vec{x} = \sum_{\beta=0}^{3} \Lambda^\vec{0}_\beta \Delta x^\beta,$$

where $\{\Lambda^\vec{0}_\beta\}$ are four numbers, one for each value of $\beta$. In this case

$$\Lambda^\vec{0}_0 = 1/\sqrt{(1 - v^2)}, \quad \Lambda^\vec{0}_1 = -v/\sqrt{(1 - v^2)},$$

$$\Lambda^\vec{0}_2 = \Lambda^\vec{0}_3 = 0.$$

A similar equation holds for $\Delta \vec{x}^1$, and so in general we write

$$\Delta \vec{x}^\vec{\alpha} = \sum_{\beta=0}^{3} \Lambda^\vec{\alpha}_\beta \Delta x^\beta, \text{ for arbitrary } \vec{\alpha}. \quad (2.3)$$

Now $\{\Lambda^\vec{\alpha}_\beta\}$ is a collection of 16 numbers, which constitutes the Lorentz transformation matrix. The reason we have written one index up and the other down will become clear when we study differential geometry. For now, it enables us to introduce the final bit of notation, the Einstein summation convention: whenever an expression contains one index as a superscript and the same index as a subscript, a summation is implied over all values that index can take. That is,

$$A_\alpha B^\alpha \text{ and } T^\gamma E_{\gamma \alpha}$$

are shorthand for the summations

$$\sum_{\alpha=0}^{3} A_\alpha B^\alpha \text{ and } \sum_{\gamma=0}^{3} T^\gamma E_{\gamma \alpha},$$

while

$$A_\alpha B^\beta, T^\gamma E_{\beta \alpha}, \text{ and } A_\beta A_\beta$$

do not represent sums on any index. The Lorentz transformation, Eq. (2.3), can now be abbreviated to

$$\Delta \vec{x}^\vec{\alpha} = \Lambda^\vec{\alpha}_\beta \Delta x^\beta, \quad (2.4)$$

saving some messy writing.

1 This is what some books on linear algebra call a ‘passive’ transformation: the coordinates change, but the vector does not.
Notice that Eq. (2.4) is identically equal to
\[ \Delta x^{\bar{\alpha}} = \Lambda^{\bar{\alpha}}_{\gamma} \Delta x^{\gamma}. \]
Since the repeated index (\(\beta\) in one case, \(\gamma\) in the other) merely denotes a summation from 0 to 3, it doesn’t matter what letter is used. Such a summed index is called a dummy index, and relabeling a dummy index (as we have done, replacing \(\beta\) by \(\gamma\)) is often a useful tool in tensor algebra. There is only one thing we should not replace the dummy index \(\beta\) with: a Latin index. The reason is that Latin indices can (by our convention) only take the values 1, 2, 3, whereas \(\beta\) must be able to equal zero as well. Thus, the expressions
\[ \Lambda^{\bar{\alpha}}_{\beta} \Delta x^{\beta} \text{ and } \Lambda^{\bar{\alpha}}_{\bar{i}} \Delta x^{i} \]
are not the same; in fact we have
\[ \Lambda^{\bar{\alpha}}_{\beta} \Delta x^{\beta} = \Lambda^{\bar{\alpha}}_{0} \Delta x^{0} + \Lambda^{\bar{\alpha}}_{i} \Delta x^{i}. \]  
(2.5)

Eq. (2.4) is really four different equations, one for each value that \(\bar{\alpha}\) can assume. An index like \(\bar{\alpha}\), on which no sum is performed, is called a free index. Whenever an equation is written down with one or more free indices, it is valid if and only if it is true for all possible values the free indices can assume. As with a dummy index, the name given to a free index is largely arbitrary. Thus, Eq. (2.4) can be rewritten as
\[ \Delta x^{\bar{\gamma}} = \Lambda^{\bar{\gamma}}_{\beta} \Delta x^{\beta}. \]
This is equivalent to Eq. (2.4) because \(\bar{\gamma}\) can assume the same four values that \(\bar{\alpha}\) could assume. If a free index is renamed, it must be renamed everywhere. For example, the following modification of Eq. (2.4),
\[ \Delta x^{\bar{\gamma}} = \Lambda^{\bar{\alpha}}_{\beta} \Delta x^{\beta}, \]
makes no sense and should never be written. The difference between these last two expressions is that the first guarantees that, whatever value \(\bar{\gamma}\) assumes, both \(\Delta x^{\bar{\gamma}}\) on the left and \(\Lambda^{\bar{\gamma}}_{\beta}\) on the right will have the same free index. The second expression does not link the indices in this way, so it is not equivalent to Eq. (2.4).

The general vector\(^2\) is defined by a collection of numbers (its components in some frame, say \(\mathcal{O}\))
\[ \bar{\mathbf{A}} \rightarrow (A^{0}, A^{1}, A^{2}, A^{3}) = \{A^{\alpha}\}, \]  
(2.6)
and by the rule that its components in a frame \(\mathcal{O}\) are
\[ A^{\bar{\alpha}} = \Lambda^{\bar{\alpha}}_{\beta} A^{\beta}. \]  
(2.7)

\(^2\) Such a vector, with four components, is sometimes called a four-vector to distinguish it from the three-component vectors we are used to in elementary physics, which we shall call three-vectors. Unless we say otherwise, a ‘vector’ is always a four-vector. We denote four-vectors by arrows, e.g. \(\vec{A}\), and three-vectors by boldface, e.g. \(\mathbf{A}\).
That is, its components transform the same way the coordinates do. Remember that a vector can be defined by giving four numbers (e.g. \((10^8, -10^{-16}, 5.8368, \pi)\)) in some frame; then its components in all other frames are uniquely determined. Vectors in spacetime obey the usual rules: if \(\vec{A}\) and \(\vec{B}\) are vectors and \(\mu\) is a number, then \(\vec{A} + \vec{B}\) and \(\mu \vec{A}\) are also vectors, with components

\[
\vec{A} + \vec{B} \rightarrow (A^0 + B^0, A^1 + B^1, A^2 + B^2, A^3 + B^3), \quad \mu \vec{A} \rightarrow (\mu A^0, \mu A^1, \mu A^2, \mu A^3).
\]

(2.8)

Thus, vectors add by the usual parallelogram rule. Notice that we can give any four numbers to make a vector, except that if the numbers are not dimensionless, they must all have the same dimensions, since under a transformation they will be added together.

### 2.2 Vector algebra

#### Basis vectors

In any frame \(\mathcal{O}\) there are four special vectors, defined by giving their components:

\[
\begin{align*}
\vec{e}_0 &\rightarrow (1, 0, 0, 0), \\
\vec{e}_1 &\rightarrow (0, 1, 0, 0), \\
\vec{e}_2 &\rightarrow (0, 0, 1, 0), \\
\vec{e}_3 &\rightarrow (0, 0, 0, 1).
\end{align*}
\]

(2.9)

These definitions define the basis vectors of the frame \(\mathcal{O}\). Similarly, \(\mathcal{O}'\) has basis vectors \(\vec{e}_0 \rightarrow (1, 0, 0, 0)\), etc.

Generally, \(\vec{e}_0 \neq \vec{e}_0\), since they are defined in different frames. The reader should verify that the definition of the basis vectors is equivalent to

\[
(e_\alpha)^\beta = \delta_\alpha^\beta.
\]

(2.10)

That is, the \(\beta\) component of \(\vec{e}_\alpha\) is the Kroncker delta: 1 if \(\beta = \alpha\) and 0 if \(\beta \neq \alpha\).

Any vector can be expressed in terms of the basis vectors. If

\[
\vec{A} \rightarrow (A^0, A^1, A^2, A^3),
\]

then

\[
\begin{align*}
\vec{A} &= A^0\vec{e}_0 + A^1\vec{e}_1 + A^2\vec{e}_2 + A^3\vec{e}_3, \\
\vec{A} &= A^\alpha\vec{e}_\alpha.
\end{align*}
\]

(2.11)
In the last line we use the summation convention (remember always to write the index on $\vec{e}$ as a subscript in order to employ the convention in this manner). The meaning of Eq. (2.11) is that $\vec{A}$ is the linear sum of four vectors $A^0\vec{e}_0$, $A^1\vec{e}_1$, etc.

**Transformation of basis vectors**

The discussion leading up to Eq. (2.11) could have been applied to any frame, so it is equally true in $\bar{O}$:

$$\vec{A} = A^{\bar{a}}\vec{e}_{\bar{a}}.$$  

This says that $\vec{A}$ is also the sum of the four vectors $A^{\bar{0}}\vec{e}_{\bar{0}}, A^{\bar{1}}\vec{e}_{\bar{1}}$, etc. These are not the same four vectors as in Eq. (2.11), since they are parallel to the basis vectors of $\bar{O}$ and not of $O$, but they add up to the same vector $\vec{A}$. It is important to understand that the expressions $A^{\alpha}\vec{e}_{\alpha}$ and $A^{\bar{a}}\vec{e}_{\bar{a}}$ are not obtained from one another merely by relabeling dummy indices. Barred and unbarred indices cannot be interchanged, since they have different meanings. Thus, $\{A^{\bar{a}}\}$ is a different set of numbers from $\{A^{\alpha}\}$, just as the set of vectors $\{\vec{e}_{\bar{a}}\}$ is different from $\{\vec{e}_{\alpha}\}$. But, by definition, the two sums are the same:

$$A^{\alpha}\vec{e}_{\alpha} = A^{\bar{a}}\vec{e}_{\bar{a}}, \quad (2.12)$$

and this has an important consequence: from it we deduce the transformation law for the basis vectors, i.e. the relation between $\{\vec{e}_{\alpha}\}$ and $\{\vec{e}_{\bar{a}}\}$. Using Eq. (2.7) for $A^{\bar{a}}$, we write Eq. (2.12) as

$$\Lambda^{\bar{a}}_{\beta A^{\alpha}}\vec{e}_{\bar{a}} = A^{\alpha}\vec{e}_{\alpha}.$$  

On the left we have two sums. Since they are finite sums their order doesn’t matter. Since the numbers $\Lambda^{\alpha}_{\beta A}$ and $A^{\beta}$ are just numbers, their order doesn’t matter, and we can write

$$A^{\beta}\Lambda^{\bar{a}}_{\beta A^{\alpha}}\vec{e}_{\bar{a}} = A^{\alpha}\vec{e}_{\alpha}.$$  

Now we use the fact that $\beta$ and $\bar{a}$ are dummy indices: we change $\beta$ to $\alpha$ and $\bar{a}$ to $\bar{\beta}$,

$$A^{\alpha}\Lambda^{\bar{\beta}}_{\alpha A^{\alpha}}\vec{e}_{\bar{\beta}} = A^{\alpha}\vec{e}_{\alpha}.$$  

This equation must be true for all sets $\{A^{\alpha}\}$, since $\vec{A}$ is an arbitrary vector. Writing it as

$$A^{\alpha}(\Lambda^{\bar{\beta}}_{\alpha A^{\alpha}}\vec{e}_{\bar{\beta}} - \vec{e}_{\alpha}) = 0$$

we deduce

$$\Lambda^{\bar{\beta}}_{\alpha A^{\alpha}}\vec{e}_{\bar{\beta}} - \vec{e}_{\alpha} = 0 \text{ for every value of } \alpha,$$

or

$$\vec{e}_{\alpha} = \Lambda^{\bar{\beta}}_{\alpha A^{\alpha}}\vec{e}_{\bar{\beta}}. \quad (2.13)$$
This gives the law by which basis vectors change. It is not a component transformation: it gives the basis $\{\vec{e}_\alpha\}$ of $\mathcal{O}$ as a linear sum over the basis $\{\vec{e}_{\bar{\alpha}}\}$ of $\mathcal{O}$. Comparing this to the law for components, Eq. (2.7),

$$A^{\bar{\beta}} = \Lambda^{\bar{\beta}}_{\alpha}A^\alpha,$$

we see that it is different indeed.

The above discussion introduced many new techniques, so study it carefully. Notice that the omission of the summation signs keeps things neat. Notice also that a step of key importance was relabeling the dummy indices: this allowed us to isolated the arbitrary $A^\alpha$ from the rest of the things in the equation.

**An example**

Let $\mathcal{O}$ move with velocity $v$ in the $x$ direction relative to $\mathcal{O}$. Then the matrix $[\Lambda^{\bar{\beta}}_{\alpha}]$ is

$$
\begin{pmatrix}
\gamma & -v\gamma & 0 & 0 \\
-v\gamma & \gamma & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix},
$$

where we use the standard notation

$$\gamma := 1/\sqrt{1 - v^2}.$$

Then, if $\vec{A} \rightarrow (5, 0, 0, 2)$, we find its components in $\mathcal{O}$ by

$$A^\bar{0} = \Lambda^\bar{0}_0 A^0 + \Lambda^\bar{0}_1 A^1 + \cdots = \gamma \cdot 5 + (-v\gamma) \cdot 0 + 0 \cdot 0 + 0 \cdot 2 = 5\gamma.$$

Similarly,

$$A^\bar{1} = -5v\gamma,$$

$$A^\bar{2} = 0,$$

$$A^\bar{3} = 2.$$

Therefore, $\vec{A} \rightarrow (5\gamma, -5v\gamma, 0, 2)$.

The basis vectors are expressible as

$$\vec{e}_\alpha = \Lambda^{\bar{\beta}}_{\alpha} \vec{e}_{\bar{\beta}}$$

or

$$\vec{e}_0 = \Lambda^{\bar{0}}_{00} \vec{e}_0 + \Lambda^{\bar{1}}_{00} \vec{e}_1 + \cdots = \gamma \vec{e}_0 - v\gamma \vec{e}_1.$$
Similarly,

\[ \vec{e}_1 = -v\gamma \vec{e}_0 + \gamma \vec{e}_1, \]

\[ \vec{e}_2 = \vec{e}_2, \]

\[ \vec{e}_3 = \vec{e}_3. \]

This gives \( O \)'s basis in terms of \( \bar{O} \)'s, so let us draw the picture (Fig. 2.1) in \( \bar{O} \)'s frame: This transformation is of course exactly what is needed to keep the basis vectors pointing along the axes of their respective frames. Compare this with Fig. 1.5(b).

**Inverse transformations**

The only thing the Lorentz transformation \( \Lambda^{\beta\alpha} \) depends on is the relative velocity of the two frames. Let us for the moment show this explicitly by writing

\[ \Lambda^{\beta\alpha} = \Lambda_{\beta\alpha}(v). \]

Then

\[ \tilde{e}_\alpha = \Lambda_{\beta\alpha}(v)\tilde{e}_\beta. \]  

(2.14)

If the basis of \( O \) is obtained from that of \( \bar{O} \) by the transformation with velocity \( v \), then the reverse must be true if we use \(-v\). Thus we must have

\[ \tilde{e}_{\tilde{\mu}} = \Lambda_{\nu\tilde{\mu}}(-v)\tilde{e}_\nu. \]

(2.15)

In this equation I have used \( \tilde{\mu} \) and \( \nu \) as indices to avoid confusion with the previous formula. The bars still refer, of course, to the frame \( \bar{O} \). The matrix \( [\Lambda_\nu^{\nu\tilde{\mu}}] \) is exactly the matrix \( [\Lambda^{\beta\alpha}] \) except with \( v \) changed to \(-v\). The bars on the indices only serve to indicate the names of the observers involved: they affect the entries in the matrix \( [\Lambda] \) only in that the matrix is always constructed using the velocity of the upper-index frame relative to the
lower-index frame. This is made explicit in Eqs. (2.14) and (2.15). Since \( v \) is the velocity of \( \bar{O} \) (the upper-index frame in Eq. (2.14)) relative to \( O \), then \(-v\) is the velocity of \( O \) (the upper-index frame in Eq. (2.15)) relative to \( \bar{O} \). Exer. 11, § 2.9, will help you understand this point.

We can rewrite the last expression as

\[
\vec{e}_\beta = \Lambda^v_\beta (-v) \vec{e}_v.
\]

Here we have just changed \( \bar{\mu} \) to \( \bar{\beta} \). This doesn’t change anything: it is still the same four equations, one for each value of \( \bar{\beta} \). In this form we can put it into the expression for \( \vec{e}_\alpha \), Eq. (2.14):

\[
\vec{e}_\alpha = \Lambda^{\bar{\beta}}_\alpha (v) \vec{e}_\bar{\beta} = \Lambda^{\bar{\beta}}_\alpha (v) \Lambda^v_\bar{\beta} (-v) \vec{e}_v. \tag{2.16}
\]

In this equation only the basis of \( O \) appears. It must therefore be an identity for all \( v \). On the right there are two sums, one on \( \bar{\beta} \) and one on \( v \). If we imagine performing the \( \bar{\beta} \) sum first, then the right is a sum over the basis \( \{ \vec{e}_v \} \) in which each basis vector \( \vec{e}_v \) has coefficient

\[
\sum_{\bar{\beta}} \Lambda^{\bar{\beta}}_\alpha (v) \Lambda^v_\bar{\beta} (-v). \tag{2.17}
\]

Imagine evaluating Eq. (2.16) for some fixed value of the index \( \alpha \). If the right side of Eq. (2.16) is equal to the left, the coefficient of \( \vec{e}_\alpha \) on the right must be 1 and all other coefficients must vanish. The mathematical way of saying this is

\[
\Lambda^{\bar{\beta}}_\alpha (v) \Lambda^v_\bar{\beta} (-v) = \delta^v_\alpha,
\]

where \( \delta^v_\alpha \) is the Kronecker delta again. This would imply

\[
\vec{e}_\alpha = \delta^v_\alpha \vec{e}_v,
\]

which is an identity.

Let us change the order of multiplication above and write down the key formula

\[
\Lambda^v_\bar{\beta} (-v) \Lambda^\bar{\beta} _\alpha (v) = \delta^v_\alpha. \tag{2.18}
\]

This expresses the fact that the matrix \( [\Lambda^v_\bar{\beta} (-v)] \) is the inverse of \( [\Lambda^{\bar{\beta}}_\alpha (v)] \), because the sum on \( \bar{\beta} \) is exactly the operation we perform when we multiply two matrices. The matrix \( (\delta^v_\alpha) \) is, of course, the identity matrix.

The expression for the change of a vector’s components,

\[
A^{\bar{\beta}} = \Lambda^{\bar{\beta}}_\alpha (v) A^\alpha,
\]

also has its inverse. Let us multiply both sides by \( \Lambda^v_\bar{\beta} (-v) \) and sum on \( \bar{\beta} \). We get

\[
\Lambda^v_\bar{\beta} (-v) A^{\bar{\beta}} = \Lambda^v_\bar{\beta} (-v) \Lambda^{\bar{\beta}}_\alpha (v) A^\alpha = \delta^v_\alpha A^\alpha = A^v.
\]
This says that the components of $\vec{A}$ in $\mathcal{O}$ are obtained from those in $\tilde{\mathcal{O}}$ by the transformation with $-v$, which is, of course, correct.

The operations we have performed should be familiar to you in concept from vector algebra in Euclidean space. The new element we have introduced here is the index notation, which will be a permanent and powerful tool in the rest of the book. Make sure that you understand the geometrical meaning of all our results as well as their algebraic justification.

### 2.3 The four-velocity

A particularly important vector is the four-velocity of a world line. In the three-geometry of Galileo, the velocity was a vector tangent to a particle’s path. In our four-geometry we define the four-velocity $\vec{U}$ to be a vector tangent to the world line of the particle, and of such a length that it stretches one unit of time in that particle’s frame. For a uniformly moving particle, let us look at this definition in the inertial frame in which it is at rest. Then the four-velocity points parallel to the time axis and is one unit of time long. That is, it is identical with $\vec{e}_0$ of that frame. Thus we could also use as our definition of the four-velocity of a uniformly moving particle that it is the vector $\vec{e}_0$ in its inertial rest frame. The word ‘velocity’ is justified by the fact that the spatial components of $\vec{U}$ are closely related to the particle’s ordinary velocity $v$, which is called the three-velocity. This will be demonstrated in the example below, Eq. (2.21).

An accelerated particle has no inertial frame in which it is always at rest. However, there is an inertial frame which momentarily has the same velocity as the particle, but which a moment later is of course no longer comoving with it. This frame is the momentarily comoving reference frame (MCRF), and is an important concept. (Actually, there are an infinity of MCRFs for a given accelerated particle at a given event; they all have the same velocity, but their spatial axes are obtained from one another by rotations. This ambiguity

![Figure 2.2](image-url)  

**Figure 2.2** The four-velocity and MCRF basis vectors of the world line at $\mathcal{A}$.
Vector analysis in special relativity

will usually not be important.) The four-velocity of an accelerated particle is defined as the \( \vec{e}_0 \) basis vector of its MCRF at that event. This vector is tangent to the (curved) world line of the particle. In Fig. 2.2 the particle at event \( A \) has MCRF \( \bar{O} \), the basis vectors of which are shown. The vector \( \vec{e}_0 \) is identical to \( \bar{U} \) there.

2.4 The four-momentum

The four-momentum \( \bar{p} \) is defined as

\[
\bar{p} = m\bar{U},
\]

where \( m \) is the rest mass of the particle, which is its mass as measured in its rest frame. In some frame \( O \) it has components conventionally denoted by

\[
\bar{p} \rightarrow O (E, p^1, p^2, p^3).
\]

We call \( p^0 \) the energy \( E \) of the particle in the frame \( O \). The other components are its spatial momentum \( p^i \).

An example

A particle of rest mass \( m \) moves with velocity \( v \) in the \( x \) direction of frame \( O \). What are the components of the four-velocity and four-momentum? Its rest frame \( \bar{O} \) has time basis vector \( \vec{e}_0 \), so, by definition of \( \bar{p} \) and \( \bar{U} \), we have

\[
\bar{U} = \vec{e}_0, \quad \bar{p} = m\bar{U},
\]

\[
U^\alpha = \Lambda^\alpha_\beta (\vec{e}_0 )^\beta = \Lambda^\alpha_0, \quad p^\alpha = m\Lambda^\alpha_0.
\]

Therefore we have

\[
U^0 = (1 - v^2)^{-1/2}, \quad p^0 = m(1 - v^2)^{-1/2},
\]

\[
U^1 = v(1 - v^2)^{-1/2}, \quad p^1 = mv(1 - v^2)^{-1/2},
\]

\[
U^2 = 0, \quad p^2 = 0,
\]

\[
U^3 = 0, \quad p^3 = 0.
\]

For small \( v \), the spatial components of \( \bar{U} \) are \((v, 0, 0)\), which justifies calling it the four-velocity, while the spatial components of \( \bar{p} \) are \((mv, 0, 0)\), justifying its name. For small \( v \), the energy is

\[
E := p^0 = m(1 - v^2)^{-1/2} \simeq m + \frac{1}{2}mv^2.
\]

This is the rest-mass energy plus the Galilean kinetic energy.
Conservation of four-momentum

The interactions of particles in Galilean physics are governed by the laws of conservation of energy and of momentum. Since the components of $\vec{p}$ reduce in the nonrelativistic limit to the familiar Galilean energy and momentum, it is natural to postulate that the correct relativistic law is that the four-vector $\vec{p}$ is conserved. That is, if several particles interact, then

$$\vec{p} := \sum_{\text{all particles}} \vec{p}_{(i)},$$

(2.22)

where $\vec{p}_{(i)}$ is the $i$th particle’s momentum, is the same before and after each interaction.

This law has the status of an extra postulate, since it is only one of many where the nonrelativistic limit is correct. However, like the two fundamental postulates of SR, this one is amply verified by experiment. Not the least of its new predictions is that the energy conservation law must include rest mass: rest mass can be decreased and the difference turned into kinetic energy and hence into heat. This happens every day in nuclear power stations.

There is an important point glossed over in the above statement of the conservation of four-momentum. What is meant by ‘before’ and ‘after’ a collision? Suppose there are two collisions, involving different particles, which occur at spacelike separated events, as below. When adding up the total four-momentum, should we take them as they are on the line of constant time $t$ or on the line of constant $\bar{t}$? As measured by $O$, event $A$ in Fig. 2.3 occurs before $t = 0$ and $B$ after, so the total momentum at $t = 0$ is the sum of the momenta after $A$ and before $B$. On the other hand, to $\bar{O}$ they both occur before $\bar{t} = 0$ and so the total momentum at $\bar{t} = 0$ is the sum of the momenta after $A$ and after $B$. There is even a

![Figure 2.3](image-url)

When several collisions are involved, the individual four-momentum vectors contributing to the total four-momentum at any particular time may depend upon the frame, but the total four-momentum is the same four-vector in all frames; its components transform from frame to frame by the Lorentz transformation.
frame in which $B$ is earlier than $A$ and the adding-up may be the reverse of $O$’s. There is really no problem here, though. Since each collision conserves momentum, the sum of the momenta before $A$ is the same as that after $A$, and likewise for $B$. So every inertial observer will get the same total four-momentum vector $\vec{p}$. (Its components will still be different in different frames, but it will be the same vector.) This is an important point: any observer can define his line of constant time (this is actually a three-space of constant time, which is called a hypersurface of constant time in the four-dimensional spacetime), at that time add up all the momenta, and get the same vector as any other observer does. It is important to understand this, because such conservation laws will appear again.

Center of momentum (CM) frame

The center of momentum frame is the inertial frame where the total momentum vanishes:

$$\sum_i \vec{p}_{(i)} \rightarrow \text{CM} (E_{\text{TOTAL}}, 0, 0, 0).$$

(2.23)

As with MCRFs, any other frame at rest relative to a CM frame is also a CM frame.

2.5 Scalar product

Magnitude of a vector

By analogy with the interval we define

$$\vec{A}^2 = -(A^0)^2 + (A^1)^2 + (A^2)^2 + (A^3)^2$$

(2.24)

to be the magnitude of the vector $\vec{A}$. Because we defined the components to transform under a Lorentz transformation in the same manner as $(\Delta t, \Delta x, \Delta y, \Delta z)$, we are guaranteed that

$$-(A^0)^2 + (A^1)^2 + (A^2)^2 + (A^3)^2 = -(A^0)^2 + (A^1)^2 + (A^2)^2 + (A^3)^2.$$  

(2.25)

The magnitude so defined is a frame-independent number, i.e. a scalar under Lorentz transformations.

This magnitude doesn’t have to be positive, of course. As with intervals we adopt the following names: if $\vec{A}^2$ is positive, $\vec{A}$ is a spacelike vector; if zero, a null vector; and if negative, a timelike vector. Thus, spatially pointing vectors have positive magnitude, as is usual in Euclidean space. It is particularly important to understand that a null vector is not a zero vector. That is, a null vector has $\vec{A}^2 = 0$, but not all $A^\alpha$ vanish; a zero vector is defined as one, where all of the components vanish. Only in a space where $\vec{A}^2$ is positive-definite does $\vec{A}^2 = 0$ require $A^\alpha = 0$ for all $\alpha$. 

We define the scalar product of $\vec{A}$ and $\vec{B}$ to be

$$\vec{A} \cdot \vec{B} = -A^0 B^0 + A^1 B^1 + A^2 B^2 + A^3 B^3$$  \hspace{1cm} (2.26)$$

in some frame $\mathcal{O}$. We now prove that this is the same number in all other frames. We note first that $\vec{A} \cdot \vec{A}$ is just $A^2$, which we know is invariant. Therefore $(\vec{A} + \vec{B}) \cdot (\vec{A} + \vec{B})$, which is the magnitude of $\vec{A} + \vec{B}$, is also invariant. But from Eqs. (2.24) and (2.26) it follows that

$$(\vec{A} + \vec{B}) \cdot (\vec{A} + \vec{B}) = A^2 + B^2 + 2\vec{A} \cdot \vec{B}.$$  

Since the left-hand side is the same in all frames, as are first two terms on the right, then the last term on the right must be as well. This proves the frame invariance of the scalar product.

Two vectors $\vec{A}$ and $\vec{B}$ are said to be orthogonal if $\vec{A} \cdot \vec{B} = 0$. The minus sign in the definition of the scalar product means that two vectors orthogonal to one another are not necessarily at right angles in the spacetime diagram (see examples below). An extreme example is the null vector, which is orthogonal to itself! Such a phenomenon is not encountered in spaces where the scalar product is positive-definite.

**Example**

The basis vectors of a frame $\mathcal{O}$ satisfy:

$$\vec{e}_0 \cdot \vec{e}_0 = -1,$$

$$\vec{e}_1 \cdot \vec{e}_1 = \vec{e}_2 \cdot \vec{e}_2 = \vec{e}_3 \cdot \vec{e}_3 = +1,$$

$$\vec{e}_\alpha \cdot \vec{e}_\beta = 0 \quad \text{if} \quad \alpha \neq \beta.$$  

They thus make up a tetrad of mutually orthogonal vectors: an orthonormal tetrad, which means orthogonal and normalized to unit magnitude. (A timelike vector has ‘unit magnitude’ if its magnitude is $-1$.) The relations above can be summarized as

$$\vec{e}_\alpha \cdot \vec{e}_\beta = \eta_{\alpha\beta},$$  \hspace{1cm} (2.27)$$

where $\eta_{\alpha\beta}$ is similar to a Kronecker delta in that it is zero when $\alpha \neq \beta$, but it differs in that $\eta_{00} = -1$, while $\eta_{11} = \eta_{22} = \eta_{33} = +1$. We will see later that $\eta_{\alpha\beta}$ is in fact of central importance: it is the metric tensor. But for now we treat it as a generalized Kronecker delta.

**Example**

The basis vectors of $\bar{\mathcal{O}}$ also satisfy

$$\vec{e}_\bar{\alpha} \cdot \vec{e}_\bar{\beta} = \eta_{\bar{\alpha}\bar{\beta}},$$

where $\eta_{\bar{\alpha}\bar{\beta}}$ is similar to $\eta_{\alpha\beta}$ but differs in that $\eta_{\bar{0}\bar{0}} = -1$, while $\eta_{\bar{1}\bar{1}} = \eta_{\bar{2}\bar{2}} = \eta_{\bar{3}\bar{3}} = +1$. We will see later that $\eta_{\bar{\alpha}\bar{\beta}}$ is in fact of central importance: it is the metric tensor. But for now we treat it as a generalized Kronecker delta.
so that, in particular, \( \vec{e}_0 \cdot \vec{e}_1 = 0 \). Look at this in the spacetime diagram of \( \mathcal{O} \), Fig. 2.4: The two vectors certainly are not perpendicular in the picture. Nevertheless, their scalar product is zero. The rule is that two vectors are orthogonal if they make equal angles with the 45° line representing the path of a light ray. Thus, a vector tangent to the light ray is orthogonal to itself. This is just another way in which SR cannot be ‘visualized’ in terms of notions we have developed in Euclidean space.

**Example**

The four-velocity \( \vec{U} \) of a particle is just the time basis vector of its MCRF, so from Eq. (2.27) we have

\[
\vec{U} \cdot \vec{U} = -1.
\]  

(2.28)

### 2.6 Applications

**Four-velocity and acceleration as derivatives**

Suppose a particle makes an infinitesimal displacement \( \vec{d} \mathbf{x} \), whose components in \( \mathcal{O} \) are \((dr, dx, dy, dz)\). The magnitude of this displacement is, by Eq. (2.24), just \(-dr^2 + dx^2 + dy^2 + dz^2\). Comparing this with Eq. (1.1), we see that this is just the interval, \( ds^2 \):

\[
ds^2 = d\mathbf{x} \cdot d\mathbf{x}.
\]  

(2.29)

Since the world line is timelike, this is negative. This led us (Eq. (1.9)) to define the proper time \( d\tau \) by

\[
(d\tau)^2 = -d\mathbf{x} \cdot d\mathbf{x}.
\]  

(2.30)
Now consider the vector \( \frac{d\vec{x}}{d\tau} \), where \( d\tau \) is the square root of Eq. (2.30) (Fig. 2.5). This vector is tangent to the world line since it is a multiple of \( d\vec{x} \). Its magnitude is
\[
\frac{d\vec{x}}{d\tau} \cdot \frac{d\vec{x}}{d\tau} = \frac{d\vec{x} \cdot d\vec{x}}{(d\tau)^2} = 1.
\]
It is therefore a timelike vector of unit magnitude tangent to the world line. In an MCRF,
\[
d\vec{x} \rightarrow \begin{pmatrix} dt \\ 0 \\ 0 \\ 0 \end{pmatrix},
\]
so that
\[
d\vec{x} \frac{d\vec{x}}{d\tau} \rightarrow \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}
\]
or
\[
d\vec{x} \frac{d\vec{x}}{d\tau} = (\vec{e}_0)_{\text{MCRF}}.
\]
This was the definition of the four-velocity. So we have the useful expression
\[
\vec{U} = \frac{d\vec{x}}{d\tau}.
\]  
(2.31)
Moreover, let us examine
\[
\frac{d\vec{U}}{d\tau} = \frac{d^2\vec{x}}{d\tau^2},
\]
which is some sort of four-acceleration. First we differentiate Eq. (2.28) and use Eq. (2.26):
\[
\frac{d}{d\tau} (\vec{U} \cdot \vec{U}) = 2\vec{U} \cdot \frac{d\vec{U}}{d\tau}.
\]
But since \( \vec{U} \cdot \vec{U} = 1 \) is a constant we have
\[
\vec{U} \cdot \frac{d\vec{U}}{d\tau} = 0.
\]

![Figure 2.5](image)

The infinitesimal displacement vector \( d\vec{x} \) tangent to a world line.
Since, in the MCRF, $\vec{U}$ has only a zero component, this orthogonality means that
\[
\frac{d\vec{U}}{d\tau} \rightarrow_{\text{MCRF}} (0, a^1, a^2, a^3).
\]
This vector is defined as the *acceleration* four-vector $\vec{a}$:

\[
\vec{a} = \frac{d\vec{U}}{d\tau}, \quad \vec{U} \cdot \vec{a} = 0. \tag{2.32}
\]

Exer. 19, § 2.9, justifies the name ‘acceleration’.

**Energy and momentum**

Consider a particle whose momentum is $\vec{p}$. Then
\[
\vec{p} \cdot \vec{p} = m^2 \vec{U} \cdot \vec{U} = -m^2. \tag{2.33}
\]
But
\[
\vec{p} \cdot \vec{p} = -E^2 + (p^1)^2 + (p^2)^2 + (p^3)^2.
\]
Therefore,
\[
E^2 = m^2 + \sum_{i=1}^{3} (p^i)^2. \tag{2.34}
\]
This is the familiar expression for the total energy of a particle.

Suppose an observer $\bar{O}$ moves with four-velocity $\vec{U}_{\text{obs}}$ not necessarily equal to the particle’s four-velocity. Then
\[
\vec{p} \cdot \vec{U}_{\text{obs}} = \vec{p} \cdot \bar{e}_0,
\]
where $\bar{e}_0$ is the basis vector of the frame of the observer. In that frame the four-momentum has components
\[
\vec{p} \rightarrow_{\bar{O}} (\bar{E}, p^1, p^2, p^3).
\]
Therefore, we obtain, from Eq. (2.26),
\[
-\vec{p} \cdot \vec{U}_{\text{obs}} = \bar{E}. \tag{2.35}
\]
This is an important equation. It says that the energy of the particle relative to the observer, $\bar{E}$, can be computed by anyone in any frame by taking the scalar product $\vec{p} \cdot \vec{U}_{\text{obs}}$. This is called a ‘frame-invariant’ expression for the energy relative to the observer. It is almost always helpful in calculations to use such expressions.
2.7 Photons

No four-velocity

Photons move on null lines, so, for a photon path,
\[ \mathbf{d}\vec{x} \cdot \mathbf{d}\vec{x} = 0. \] (2.36)

Therefore \( d\tau \) is zero and Eq. (2.31) shows that the four-velocity cannot be defined. Another way of saying the same thing is to note that there is no frame in which light is at rest (the second postulate of SR), so there is no MCRF for a photon. Thus, no \( \vec{e}_0 \) in any frame will be tangent to a photon’s world line.

Note carefully that it is still possible to find vectors tangent to a photon’s path (which, being a straight line, has the same tangent everywhere): \( d\vec{x} \) is one. The problem is finding a tangent of unit magnitude, since they all have vanishing magnitude.

Four-momentum

The four-momentum of a particle is not a unit vector. Instead, it is a vector where the components in some frame give the particle energy and momentum relative to that frame. If a photon carries energy \( E \) in some frame, then in that frame \( p^0 = E \). If it moves in the \( x \) direction, then \( p^y = p^z = 0 \), and in order for the four-momentum to be parallel to its world line (hence be null) we must have \( p^x = E \). This ensures that
\[ \vec{p} \cdot \vec{p} = -E^2 + E^2 = 0. \] (2.37)

So we conclude that photons have spatial momentum equal to their energy.

We know from quantum mechanics that a photon has energy
\[ E = h\nu, \] (2.38)

where \( \nu \) is its frequency and \( h \) is Planck’s constant, \( h = 6.6256 \times 10^{-34} \text{ J s} \).

This relation and the Lorentz transformation of the four-momentum immediately give us the Doppler-shift formula for photons. Suppose, for instance, that in frame \( O \) a photon has frequency \( \nu \) and moves in the \( x \) direction. Then, in \( \bar{O} \), which has velocity \( v \) in the \( x \) direction relative to \( O \), the photon’s energy is
\[
\bar{E} = E/\sqrt{(1 - v^2)} - p^xv/\sqrt{(1 - v^2)} \\
= h\nu/\sqrt{(1 - v^2)} - h\nu v/\sqrt{(1 - v^2)}. 
\]

Setting this equal to \( h\bar{\nu} \) gives \( \bar{\nu} \), the frequency in \( \bar{O} \):
\[ \bar{\nu}/\nu = (1 - v)/\sqrt{(1 - v^2)} = \sqrt{(1 - v)/(1 + v)}. \] (2.39)

This is generalized in Exer. 25, § 2.9.
The rest mass of a photon must be zero, since

\[ m^2 = -\vec{p} \cdot \vec{p} = 0. \]  

(2.40)

Any particle whose four-momentum is null must have rest mass zero, and conversely. The only known zero rest-mass particle is the photon. Neutrinos are very light, but not massless. (Sometimes the ‘graviton’ is added to this list, since gravitational waves also travel at the speed of light, as we shall see later. But ‘photon’ and ‘graviton’ are concepts that come from quantum mechanics, and there is as yet no satisfactory quantized theory of gravity, so that ‘graviton’ is not really a well-defined notion yet.) The idea that only particles with zero rest mass can travel at the speed of light is reinforced by the fact that no particle of finite rest mass can be accelerated to the speed of light, since then its energy would be infinite. Put another way, a particle traveling at the speed of light (in, say, the \( x \) direction) has 

\[ \frac{p_1}{p_0} = 1, \]  

while a particle of rest mass \( m \) moving in the \( x \) direction has, from the equation 

\[ \vec{p} \cdot \vec{p} = -m^2, \]  

\[ \frac{p_1}{p_0} = \left[ 1 - \frac{m^2}{(p_0)^2} \right]^{1/2}, \]  

which is always less than one, no matter how much energy the particle is given. Although it may seem to get close to the speed of light, there is an important distinction: the particle with \( m \neq 0 \) always has an MCRF, a Lorentz frame in which it is at rest, the velocity \( v \) of which is \( \frac{p_1}{p_0} \) relative to the old frame. A photon has no rest frame.

### 2.8 Further reading

We have only scratched the surface of relativistic kinematics and particle dynamics. These are particularly important in particle physics, which in turn provides the most stringent tests of SR. See Hagedorn (1963) or Wiedemann (2007).

### 2.9 Exercises

1. Given the numbers \( \{A^0 = 5, A^1 = 0, A^2 = -1, A^3 = -6\}, \) \( \{B_0 = 0, B_1 = -2, B_2 = 4, B_3 = 0\}, \) \( \{C_{00} = 1, C_{01} = 0, C_{02} = 2, C_{03} = 3, C_{30} = -1, C_{10} = 5, C_{11} = -2, C_{12} = -2, C_{13} = 0, C_{21} = 5, C_{22} = 2, C_{23} = -2, C_{20} = 4, C_{31} = -1, C_{32} = -3, C_{33} = 0\} \), find:
   - (a) \( A^\alpha B_\alpha \)
   - (b) \( A^\alpha C_{\alpha\beta} \) for all \( \beta \)
   - (c) \( A^\nu C_{\gamma\nu} \) for all \( \sigma \)
   - (d) \( A^\nu C_{\mu\nu} \) for all \( \mu \)
   - (e) \( A^\alpha B_\beta \) for all \( \alpha, \beta \)
   - (f) \( A^i B_i \)
   - (g) \( A^i B_j \) for all \( i, j \)
2. Identify the free and dummy indices in the following equations and change them into equivalent expressions with different indices. How many different equations does each expression represent?
   - (a) \( A^\alpha B_\alpha = 5 \)
   - (b) \( A^\mu = \Lambda^\mu_\nu A^\nu \)
   - (c) \( T^{\alpha\mu\lambda} A_\mu C_\lambda \gamma = D^{\gamma\alpha} \)
   - (d) \( R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = G_{\mu\nu} \)
3. Prove Eq. (2.5).
4. Given the vectors \( \vec{A} \to \mathcal{O} \) \( (5, -1, 0, 1) \) and \( \vec{B} \to \mathcal{O} \) \( (-2, 1, 1, -6) \), find the components in \( \mathcal{O} \) of
   - (a) \( -6 \vec{A} \)
   - (b) \( 3 \vec{A} + \vec{B} \)
   - (c) \( -6 \vec{A} + 3 \vec{B} \)
5 A collection of vectors \( \{ \vec{a}, \vec{b}, \vec{c}, \vec{d} \} \) is said to be linearly independent if no linear combination of them is zero except the trivial one, \( 0\vec{a} + 0\vec{b} + 0\vec{c} + 0\vec{d} = 0 \).

(a) Show that the basis vectors in Eq. (2.9) are linearly independent.

(b) Is the following set linearly independent? \( \{ \vec{a}, \vec{b}, \vec{c}, 5\vec{a} + 3\vec{b} - 2\vec{c} \} \).

6 In the \( t - x \) spacetime diagram of \( O \), draw the basis vectors \( \vec{e}_0 \) and \( \vec{e}_1 \). Draw the corresponding basis vectors of the frame \( \tilde{O} \) that moves with speed 0.6 in the positive \( x \) direction relative to \( O \). Draw the corresponding basis vectors of \( \bar{O} \) a frame that moves with speed 0.6 in the positive \( x \) direction relative to \( \tilde{O} \).

7 (a) Verify Eq. (2.10) for all \( \alpha, \beta \).

(b) Prove Eq. (2.11) from Eq. (2.9).

8 (a) Prove that the zero vector \( (0, 0, 0, 0) \) has these same components in all reference frames.

(b) Use (a) to prove that if two vectors have equal components in one frame, they have equal components in all frames.

9 Prove, by writing out all the terms, that

\[
\sum_{\alpha=0}^{3} \left( \sum_{\beta=0}^{3} \Lambda_{\alpha}^\beta A^\beta \vec{e}_\alpha \right) = \sum_{\beta=0}^{3} \left( \sum_{\alpha=0}^{3} \Lambda_{\alpha}^\beta A^\alpha \vec{e}_\beta \right)
\]

Since the order of summation doesn’t matter, we are justified in using the Einstein summation convention to write simply \( \Lambda_{\alpha}^\beta A^\alpha \vec{e}_\beta \), which doesn’t specify the order of summation.

10 Prove Eq. (2.13) from the equation \( A^\alpha (\Lambda_{\alpha}^\beta \vec{e}_\beta - \vec{e}_\alpha) = 0 \) by making specific choices for the components of the arbitrary vector \( \vec{A} \).

11 Let \( \Lambda_{\alpha}^\beta \) be the matrix of the Lorentz transformation from \( O \) to \( \tilde{O} \), given in Eq. (1.12).

Let \( \vec{A} \) be an arbitrary vector with components \( (A^0, A^1, A^2, A^3) \) in frame \( O \).

(a) Write down the matrix of \( \Lambda^\nu_{\mu}(-v) \).

(b) Find \( A^{\tilde{\alpha}} \) for all \( \tilde{\alpha} \).

(c) Verify Eq. (2.18) by performing the indicated sum for all values of \( v \) and \( \alpha \).

(d) Write down the Lorentz transformation matrix from \( \tilde{O} \) to \( O \), justifying each entry.

(e) Use (d) to find \( A^\beta \) from \( A^{\tilde{\alpha}} \). How is this related to Eq. (2.18)?

(f) Verify, in the same manner as (c), that

\[
\Lambda^\nu_{\tilde{\beta}}(v)\Lambda^{\tilde{\alpha}}_{\nu}(-v) = \delta^{\tilde{\alpha}}_{\tilde{\beta}}.
\]

(g) Establish that

\[
\vec{e}_\alpha = \delta^\nu_{\alpha} \vec{e}_\nu
\]

and

\[
A^{\tilde{\beta}} = \delta^\tilde{\beta}_{\tilde{\mu}} A^{\tilde{\mu}}.
\]

12 Given \( \vec{A} \rightarrow O \ (0, -2, 3, 5) \), find:

(a) the components of \( \vec{A} \) in \( \tilde{O} \), which moves at speed 0.8 relative to \( O \) in the positive \( x \) direction;

(b) the components of \( \vec{A} \) in \( \bar{O} \), which moves at speed 0.6 relative to \( \tilde{O} \) in the positive \( x \) direction;

(c) the magnitude of \( \vec{A} \) from its components in \( O \);
13. Let $\bar{O}$ move with velocity $v$ relative to $O$, and let $\bar{O}'$ move with velocity $v'$ relative to $\bar{O}$.
(a) Show that the Lorentz transformation from $O$ to $\bar{O}'$ is
\[
\Lambda^{\bar{a}}_{\mu} = \Lambda^{\bar{a}}_{\bar{\gamma}}(v') \Lambda^{\bar{\gamma}}_{\mu}(v).
\] (2.41)
(b) Show that Eq. (2.41) is just the matrix product of the matrices of the individual Lorentz transformations.
(c) Let $v = 0.6\bar{e}_x$, $v' = 0.8\bar{e}_y$. Find $\Lambda^{\bar{a}}_{\mu}$ for all $\mu$ and $\bar{a}$.
(d) Verify that the transformation found in (c) is indeed a Lorentz transformation by showing explicitly that $\Delta s^2 = \Delta \bar{s}^2$ for any $(\Delta t, \Delta x, \Delta y, \Delta z)$.
(e) Compute
\[
\Lambda^{\bar{a}}_{\mu}(v) \Lambda^{\bar{\gamma}}_{\bar{\gamma}}(v')
\]
for $v$ and $v'$, as given in (c), and show that the result does not equal that of (c). Interpret this physically.

14. The following matrix gives a Lorentz transformation from $O$ to $\bar{O}$:
\[
\begin{pmatrix}
1.25 & 0 & 0 & .75 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
.75 & 0 & 0 & 1.25
\end{pmatrix}
\]
(a) What is the velocity (speed and direction) of $\bar{O}$ relative to $O$?
(b) What is the inverse matrix to the given one?
(c) Find the components in $O$ of a vector $\vec{A} \rightarrow \bar{O}$ (1, 2, 0, 0).

15. (a) Compute the four-velocity components in $O$ of a particle whose speed in $O$ is $v$ in the positive $x$ direction, by using the Lorentz transformation from the rest frame of the particle.
(b) Generalize this result to find the four-velocity components when the particle has arbitrary velocity $v$, with $|v| < 1$.
(c) Use your result in (b) to express $v$ in terms of the components $\{U^\alpha\}$.
(d) Find the three-velocity $v$ of a particle whose four-velocity components are (2, 1, 1, 1).

16. Derive the Einstein velocity-addition formula by performing a Lorentz transformation with velocity $v$ on the four-velocity of a particle whose speed in the original frame was $W$.

17. (a) Prove that any timelike vector $\vec{U}$ for which $U^0 > 0$ and $\vec{U} \cdot \vec{U} = -1$ is the four-velocity of some world line.
(b) Use this to prove that for any timelike vector $\vec{V}$ there is a Lorentz frame in which $\vec{V}$ has zero spatial components.

18. (a) Show that the sum of any two orthogonal spacelike vectors is spacelike.
(b) Show that a timelike vector and a null vector cannot be orthogonal.

19. A body is said to be uniformly accelerated if its acceleration four-vector $\vec{a}$ has constant spatial direction and magnitude, say $\vec{a} \cdot \vec{a} = a^2 \geq 0$.  

(a) Show that this implies that \( \ddot{a} \) always has the same components in the body’s MCRF, and that these components are what one would call ‘acceleration’ in Galilean terms. (This would be the physical situation for a rocket whose engine always gave the same acceleration.)

(b) Suppose a body is uniformly accelerated with \( a = 10 \, \text{m s}^{-2} \) (about the acceleration of gravity on Earth). If the body starts from rest, find its speed after time \( t \). (Be sure to use the correct units.) How far has it traveled in this time? How long does it take to reach \( v = 0.999 \) ?

(c) Find the elapsed proper time for the body in (b), as a function of \( t \). (Integrate \( \text{d}t \) along its world line.) How much proper time has elapsed by the time its speed is \( v = 0.999 \)? How much would a person accelerated as in (b) age on a trip from Earth to the center of our Galaxy, a distance of about \( 2 \times 10^{20} \, \text{m} \)?

20 The world line of a particle is described by the equations

\[
\begin{align*}
x(t) &= at + b \sin \omega t, \\
y(t) &= b \cos \omega t, \\
z(t) &= 0, \quad |b\omega| < 1,
\end{align*}
\]

in some inertial frame. Describe the motion and compute the components of the particle’s four-velocity and four-acceleration.

21 The world line of a particle is described by the parametric equations in some Lorentz frame

\[
\begin{align*}
t(\lambda) &= a \sinh \left( \frac{\lambda}{a} \right), \\
x(\lambda) &= a \cosh \left( \frac{\lambda}{a} \right),
\end{align*}
\]

where \( \lambda \) is the parameter and \( a \) is a constant. Describe the motion and compute the particle’s four-velocity and acceleration components. Show that \( \lambda \) is proper time along the world line and that the acceleration is uniform. Interpret \( a \).

22 (a) Find the energy, rest mass, and three-velocity \( \mathbf{v} \) of a particle whose four-momentum has the components \( (4, 1, 1, 0) \, \text{kg} \).

(b) The collision of two particles of four-momenta

\[
\vec{p}_1 \rightarrow (3, -1, 0, 0) \, \text{kg}, \quad \vec{p}_2 \rightarrow (2, 1, 1, 0) \, \text{kg}
\]

results in the destruction of the two particles and the production of three new ones, two of which have four-momenta

\[
\vec{p}_3 \rightarrow (1, 1, 0, 0) \, \text{kg}, \quad \vec{p}_4 \rightarrow (1, -\frac{1}{2}, 0, 0) \, \text{kg}.
\]

Find the four-momentum, energy, rest mass, and three-velocity of the third particle produced. Find the CM frame’s three-velocity.

23 A particle of rest mass \( m \) has three-velocity \( \mathbf{v} \). Find its energy correct to terms of order \( |\mathbf{v}|^4 \). At what speed \( |\mathbf{v}| \) does the absolute value of \( 0(|\mathbf{v}|^4) \) term equal \( \frac{1}{2} \) of the kinetic-energy term \( \frac{1}{2}m|\mathbf{v}|^2 \)?

24 Prove that conservation of four-momentum forbids a reaction in which an electron and positron annihilate and produce a single photon (\( \gamma \)-ray). Prove that the production of two photons is not forbidden.
25 (a) Let frame \( \tilde{O} \) move with speed \( v \) in the \( x \)-direction relative to \( O \). Let a photon have frequency \( \nu \) in \( O \) and move at an angle \( \theta \) with respect to \( O \)'s \( x \) axis. Show that its frequency in \( \tilde{O} \) is

\[
\tilde{\nu}/\nu = (1 - v \cos \theta)/\sqrt{1 - v^2}.
\]

(2.42)

(b) Even when the motion of the photon is perpendicular to the \( x \) axis \( (\theta = \pi/2) \) there is a frequency shift. This is called the transverse Doppler shift, and arises because of the time dilation. At what angle \( \theta \) does the photon have to move so that there is no Doppler shift between \( O \) and \( \tilde{O} \)?

(c) Use Eqs. (2.35) and (2.38) to calculate Eq. (2.42).

26 Calculate the energy that is required to accelerate a particle of rest mass \( m \neq 0 \) from speed \( v \) to speed \( v + \delta v \) \((\delta v \ll v)\), to first order in \( \delta v \). Show that it would take an infinite amount of energy to accelerate the particle to the speed of light.

27 Two identical bodies of mass 10 kg are at rest at the same temperature. One of them is heated by the addition of 100 J of heat. Both are then subjected to the same force. Which accelerates faster, and by how much?

28 Let \( \vec{A} \rightarrow O \) \((5, 1, -1, 0)\), \( \vec{B} \rightarrow O \) \((-2, 3, 1, 6)\), \( \vec{C} \rightarrow O \) \((2, -2, 0, 0)\). Let \( \tilde{O} \) be a frame moving at speed \( v = 0.6 \) in the positive \( x \) direction relative to \( O \), with its spatial axes oriented parallel to \( O \)'s.

(a) Find the components of \( \vec{A}, \vec{B}, \) and \( \vec{C} \) in \( \tilde{O} \).

(b) Form the dot products \( \vec{A} \cdot \vec{B}, \vec{B} \cdot \vec{C}, \vec{A} \cdot \vec{C}, \) and \( \vec{C} \cdot \vec{C} \) using the components in \( \tilde{O} \).

Verify the frame independence of these numbers.

(c) Classify \( \vec{A}, \vec{B}, \) and \( \vec{C} \) as timelike, spacelike, or null.

29 Prove, using the component expressions, Eqs. (2.24) and (2.26), that

\[
\frac{d}{d\tau}(\vec{U} \cdot \vec{U}) = 2\vec{U} \cdot \frac{d\vec{U}}{d\tau}.
\]

30 The four-velocity of a rocket ship is \( \vec{U} \rightarrow O \) \((2, 1, 1, 1)\). It encounters a high-velocity cosmic ray whose momentum is \( \vec{P} \rightarrow O \) \((300, 299, 0, 0) \times 10^{-27} \) kg. Compute the energy of the cosmic ray as measured by the rocket ship’s passengers, using each of the two following methods.

(a) Find the Lorentz transformations from \( O \) to the MCRF of the rocket ship, and use it to transform the components of \( \vec{P} \).

(b) Use Eq. (2.35).

(c) Which method is quicker? Why?

31 A photon of frequency \( \nu \) is reflected without change of frequency from a mirror, with an angle of incidence \( \theta \). Calculate the momentum transferred to the mirror. What momentum would be transferred if the photon were absorbed rather than reflected?

32 Let a particle of charge \( e \) and rest mass \( m \), initially at rest in the laboratory, scatter a photon of initial frequency \( \nu_i \). This is called Compton scattering. Suppose the scattered photon comes off at an angle \( \theta \) from the incident direction. Use conservation of four-momentum to deduce that the photon’s final frequency \( \nu_f \) is given by
\[
\frac{1}{v_f} = \frac{1}{v_i} + h \left( \frac{1 - \cos \theta}{m} \right). \tag{2.43}
\]

33 Space is filled with cosmic rays (high-energy protons) and the cosmic microwave background radiation. These can Compton scatter off one another. Suppose a photon of energy \( h \nu = 2 \times 10^{-4} \text{ eV} \) scatters off a proton of energy \( 10^9 m_p = 10^{18} \text{ eV} \), energies measured in the Sun’s rest frame. Use Eq. (2.43) in the proton’s initial rest frame to calculate the maximum final energy the photon can have in the solar rest frame after the scattering. What energy range is this (X-ray, visible, etc.)?

34 Show that, if \( \vec{A}, \vec{B}, \) and \( \vec{C} \) are any vectors and \( \alpha \) and \( \beta \) any real numbers,

\[
(a \vec{A}) \cdot \vec{B} = \alpha (\vec{A} \cdot \vec{B}),
\]

\[
\vec{A} \cdot (\beta \vec{B}) = \beta (\vec{A} \cdot \vec{B}),
\]

\[
\vec{A} \cdot (\vec{B} + \vec{C}) = \vec{A} \cdot \vec{B} + \vec{A} \cdot \vec{C},
\]

\[
(\vec{A} + \vec{B}) \cdot \vec{C} = \vec{A} \cdot \vec{C} + \vec{B} \cdot \vec{C}.
\]

35 Show that the vectors \( \{\vec{e}_\beta\} \) obtained from \( \{\vec{e}_\alpha\} \) by Eq. (2.15) satisfy \( \vec{e}_\alpha \cdot \vec{e}_\beta = \eta_{\alpha \beta} \) for all \( \alpha, \beta \).