In other words, contravariant and covariant vectors of the same type differ in their signs on the space-like components. For the case of the position four-vector we have explicitly

$$x^{\mu=0} = ct \tag{1.83}$$

$$x^{\mu=1} = x \tag{1.84}$$

$$x_{\mu=0} = ct \tag{1.85}$$

$$x_{\mu=1} = -x \tag{1.86}$$

Choosing the vector components with upper indices to correspond to non-relativistic notation is a plain matter of convention. This means that the previous formulation of the position four-vector in the usual non-covariant notation will now be changed. We choose the same four-vector to be defined by its contravariant components (upper indices), as

$$x := \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \qquad \text{With components } \{x^\mu\} \tag{1.87}$$

As a known example, consider the scalar product of the position vector with itself. Obviously,

$$\begin{aligned} x'_\mu \, x^{\mu'} = x^{\mu'} \, x'_\mu &= x^\nu \, x_\nu = x_\nu \, x^\nu \\ (ct')^2 - (x')^2 - (y')^2 - (z')^2 &= (ct)^2 - x^2 - y^2 - z^2 \end{aligned} \tag{1.88}$$

which is just identical to Eq. (1.58), confirming the logical consistency of the formalism.

The generalization of these concepts to rank-n four-tensors is straightforward and will be useful for the reformulation of electrodynamics and in the framework of general relativity, as well as in many other fields of physics.

1.2.3 SpaceTime Diagrams

The phenomenological consequences of the Lorentz tranformation can be conveniently visualized through a technique already introduced by Minkowski: SpaceTime diagrams. A reduction to two dimensions – time and one spatial dimension – is sufficient for many purposes.

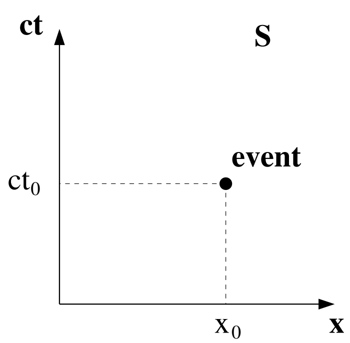


Figure 1.8:

An **event** is represented by a point in SpaceTime with coordinates $\{ct_0, x_0\}$.

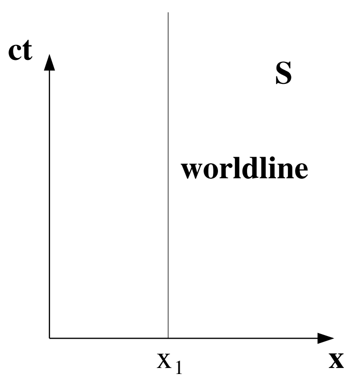


Figure 1.9:

An observer/object at rest at position x_1 is represented by a **worldline** in the SpaceTime diagram.

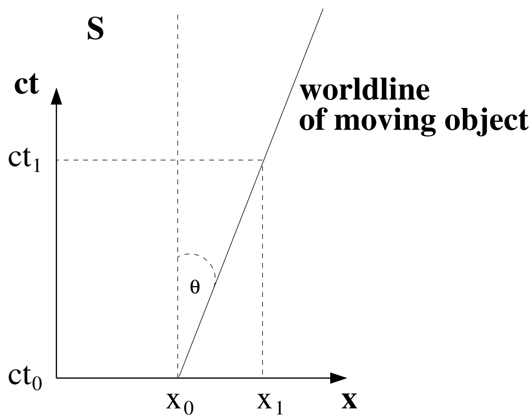
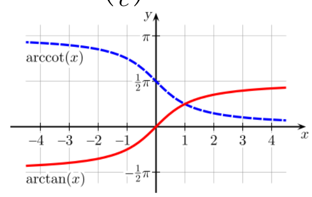


Figure 1.10:

An object moving in x-direction with velocity v in coordinates of S. This velocity follows from graphical observation: $\tan\vartheta = \frac{x_1 - x_0}{ct_1 - ct_0} = \frac{v(t_1-t_0)}{c(t_1-t_0)} = \frac{v}{c}$. So $\vartheta = \arctan\left(\frac{v}{c}\right)$.



Since $1 \geq \frac{v}{c} \geq 0$ we have $\max\left(\arctan\left(\frac{v}{c}\right)\right) = \max(\vartheta) = \frac{\pi}{4}$.

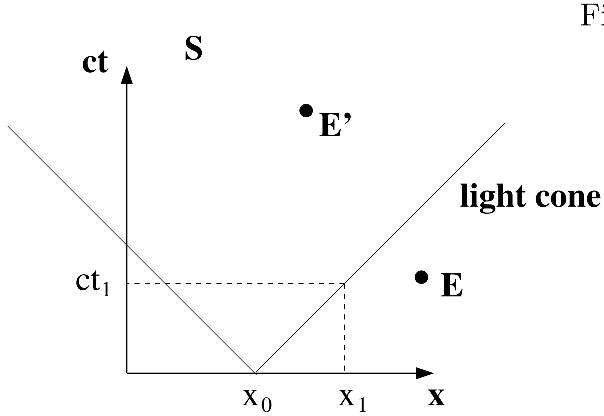


Figure 1.11:

Emission of a light pulse as event at $\{ct_0, x_0\}$ and propagation of the light pulse in frame S. Event E cannot be **causally connected** with an event at $\{ct_0, x_0\}$, because information cannot propagate faster than with c. Event E' can be **causally connected** with an event at $\{ct_0, x_0\}$ because it is inside the light cone.

As a simple illustration for the scenario with event E, imagine a distant observer appearing within $\delta t = \varepsilon t$ around t_0 at $x_E > x_1$. The observer cannot see the light pulse because he has "disappeared" before the light pulse can reach his position.

1.2.4 Space-, Light-, and Time-Like Four-Vectors

With the establishments of the preceding subsections we can revisit four-vectors in the context of SpaceTime diagrams and come to a couple of interesting conclusions.

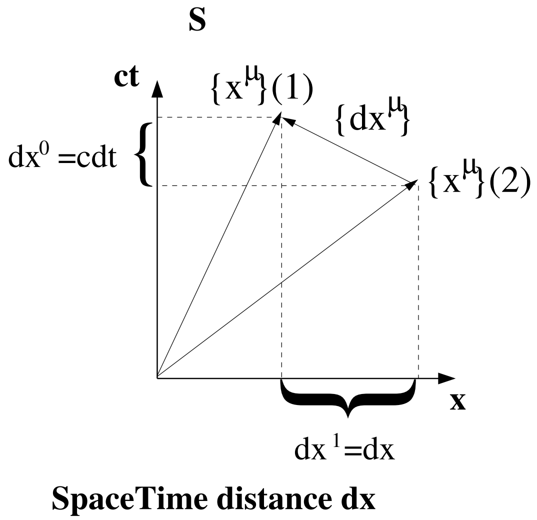


Figure 1.12:

Two events (1) and (2) and their SpaceTime distance represented by vectors in a two-dimensional vector space.

From Fig. (1.12) we infer that

$$\begin{aligned} dx^0 &= x^0(1) - x^0(2) = c(t_1 - t_2) = cdt \\ dx^1 &= x^1(1) - x^1(2) = x_1 - x_2 = dx. \end{aligned}$$

[20] The scalar product of the four-vector dx with itself is then

$$\begin{aligned} D &:= (ds)^2 = dx^\mu dx_\mu = dx^\mu \, g_{\mu\nu} \, dx^\nu \\ &= \begin{pmatrix} c(t_1 - t_2) & x_1 - x_2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} c(t_1 - t_2) \\ (x_1 - x_2) \end{pmatrix} \\ &= \begin{pmatrix} c(t_1 - t_2) & x_1 - x_2 \end{pmatrix} \begin{pmatrix} c(t_1 - t_2) \\ -(x_1 - x_2) \end{pmatrix} \qquad (1.89) \\ &= c^2(t_1 - t_2)^2 - (x_1 - x_2)^2. \qquad (1.90) \end{aligned}$$

D is a Lorentz scalar (it is the scalar product of a contra- and a co-variant four-vector), i.e., all of our following conclusions are Lorentz invariant. We distinguish three general cases:

$D = 0$ From this it follows that $c^2(t_1 - t_2)^2 = (x_1 - x_2)^2$ which means that ds is on a light cone.

Four-vectors v with $v^\mu v_\mu = 0$ are called "**light-like**" four-vectors.

[20] The notation dx^0 here means $dx^{\mu=0}$.

$D < 0$ This would correspond to the case shown in Fig. (1.12) if we would assume the events to be the points chosen for illustration. Then, if we suppose that for the Lorentz-transformed spatial components $x_1^{k'} - x_2^{k'} = 0 \quad \forall k \in \{1, \ldots, 3\}$, then $\Rightarrow D = c^2(t_1' - t_2')^2 \geq 0$ which is in contradiction with the assumption. This means that given $D < 0$ there is no reference frame K' in which the two events occur at the same position and chronologically.

On the other hand, if $t_1' - t_2' = 0 \Rightarrow D = -(x_1' - x_2')^2 \leq 0$. Therefore, a reference frame K' exists such that the two events occur **simultaneously** but at **different positions**. D is ensuingly a "**space-like**" interval[21]. Note that in this case event 1 is outside the light cone based in event 2.

Four-vectors v with $v^\mu v_\mu < 0$ are called "**space-like**" four-vectors.

$D > 0$ In this case, if $t_1' - t_2' = 0$, *i.e.*, in a frame S' the two events occur simultaneously, then $\Rightarrow D = -(x_1' - x_2')^2 \leq 0$ contradicts the assumption and the two events cannot occur simultaneously in any K'.

However, if $x_1^{k'} - x_2^{k'} = 0 \rightarrow D = c^2(t_1' - t_2')^2 > 0$ is possible which means that there exists a frame K' in which the two events occur at the **same point in space** and **chronologically**. They can be causally connected (inside the respective light cone).

Four-vectors v with $v^\mu v_\mu > 0$ are called "**time-like**" four-vectors.

[21] to be distinguished from the space-like *components* of a four-vector which is quite a different thing

1.3 Relativistic Mechanics

The new (Einsteinian) view of space and time, or rather SpaceTime, entails that most of the physics that had been invented thus far had to be re-written in order to be in accord with the principles of special relativity. The new relativistic theories should be constructed such that their non-relativistic limit gives rise to the "old", classical version of pre-Einsteinian physical theory.

We will begin with mechanics. Newton's mechanics was known to be Galilei, but not Lorentz invariant. The evident program is thus to formulate a Lorentz invariant rendering of classical[22] mechanics. This shall be attempted by retaining the structure of the equations of motion and by replacing the ordinary three-vectors by four-vectors in SpaceTime.

1.3.1 Proper Time

In non-relativistic mechanics velocity is written as

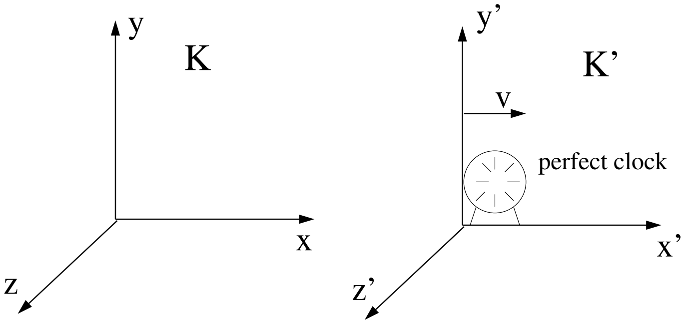
$$v_x = \frac{dx}{dt}. \tag{1.91}$$

We can guess that the relativistic version of velocity might involve the four-vector x and a four-vector v which should have the same transformation properties as x. But we know from Eq. (1.28) that time is obviously not a Lorentz scalar, $t' \neq t$ for a simple boost. In order to make v a contravariant vector, as we would expect it to be, we will have to formulate a Lorentz invariant quantity "dt".

We have seen as a result of the Michelson-Morley experiment using the Lorentz boost transformation that a time interval $\Delta t'$ in a clock's frame (time interval measured by the experimenters on Earth at a fixed

[22]I use the "international meaning" of 'classical' which is 'non-quantum' and not 'non-relativistic non-quantum'. In other words, we are going to develop relativistic classical mechanics.

position in K') relates to the time interval Δt for the same two events in an observer frame



A perfect (unaffected by any violent acceleration and undestroyable) clock in its rest frame K', moving with velocity v relative to a laboratory observer in K.

as to

$$\Delta t' = \frac{1}{\gamma(v)}\,\Delta t. \tag{1.92}$$

If we define $v(t)$ as the relative velocity in the infinitesimal time interval dt, we can generalize to time-dependent velocities and, therefore, accelerations. From now on we will call

$$d\tau := dt' = \frac{1}{\gamma(v)}\,dt \tag{1.93}$$

the **proper time** differential in the rest frame of the clock. Let's take a closer look at the properties of $d\tau$.

For this we calculate the scalar product of the four-distance ds in S with itself[23]:

$$\begin{aligned}
(ds)^2 &= dx^\mu\, dx_\mu = dx^\mu\, g_{\mu\nu}\, dx^\nu = (cdt)^2 - dx^2 - dy^2 - dz^2 \\
&= \left[c^2 - \frac{dx^2}{dt^2} - \frac{dy^2}{dt^2} - \frac{dz^2}{dt^2}\right] dt^2 \\
&= \left(c^2 - v^2\right)\, dt^2 = c^2\left(1 - \frac{v^2}{c^2}\right)\, dt^2 \\
&= c^2\, \frac{1}{\gamma(v)^2}\, dt^2 = c^2 d\tau^2
\end{aligned} \tag{1.94}$$

[23] Getting from the second to the third line can be seen by taking the scalar product of $\mathbf{v} = \frac{dx}{dt}\mathbf{e}_x + \frac{dy}{dt}\mathbf{e}_y + \frac{dz}{dt}\mathbf{e}_z +$ with itself. The resulting velocity is the speed of "something" in the coordinates of frame K, so this might be a clock at rest in K', and that is the reason why we can take this velocity as the one used in the Lorentz factor γ!

where in the last equality we have used Eq. (1.93).

However, we know that $ds^2 = ds'^2$ is a Lorentz scalar, see Eqs. (1.90) and (1.77). And since the speed of light, c, is the same in all reference frames, it follows that $d\tau^2$, and therefore also $d\tau$, have to be Lorentz scalars as well.

The physical interpretation of this finding is that the proper time interval (in the clock's frame) has to be the same for all observers. We can now proceed to building relativistic mechanics based on the proper time differential, $d\tau$.

1.3.2 Four-Velocity and Four-Acceleration

1.3.2.1 Four-Velocity

The obvious generalization of Eq. (1.91) in terms of four-vector quantities and the proper time differential is

$$\left\{\frac{dx^\mu}{d\tau}\right\} =: \{u^\mu\} \tag{1.95}$$

Now, by construction, the components u^μ transform like the components of a contravariant four-vector, x^μ, because $d\tau$ is Lorentz invariant. However, Eq. (1.95) is a sort of "mixed" expression, since τ refers to a clock's rest frame whereas x^μ is a coordinate of an arbitrary frame [24].

We derive the components of velocity in a general frame K as follows:

$$u^\mu = \frac{dx^\mu(t)}{d\tau} = \frac{dx^\mu(t)}{dt}\,\frac{dt}{d\tau} = \gamma(v)\,\frac{dx^\mu(t)}{dt} \tag{1.96}$$

where the chain rule has been used with t regarded as a function of τ.

Note that now we are considering the situation with a general three velocity $\mathbf{v} = \sum_{j=1}^{3} v_j \mathbf{e}_j$ in coordinates of K. The γ factor is then more precisely $\gamma(v := ||\mathbf{v}||)$.

[24] We may choose it to be the laboratory frame.

Summary for the velocity four-vector[25]:

$$\left\{ \begin{array}{c} u^0 \\ u^k \end{array} \right\} \equiv \left\{ \begin{array}{c} \gamma c \\ \gamma v_k \end{array} \right\} \qquad \forall k \in \{1, \ldots, 3\} \quad (1.97)$$

It is immediately obvious that in the non-relativistic limit the space-like components of velocity turn into the usual velocities with respect to reference frame K. The time-like component does not exist as component of a four vector in non-relativistic theory.

Let's have a look at the scalar product of the velocity four-vector with itself. We know from Eq. (1.77) that this product $u^2 = u^\mu\, u_\mu = u_\mu\, u^\mu$ has to be a Lorentz scalar. The direct calculation gives

$$u^\mu\, u_\mu = u^\mu\, g_{\mu\nu}\, u^\nu = \gamma(||\mathbf{v}||)^2(c^2 - v_x^2 - v_y^2 - v_z^2) = \frac{c^2}{c^2 - \mathbf{v}^2}(c^2 - \mathbf{v}^2) = c^2. \tag{1.98}$$

It is confirmed that u^2 is a Lorentz scalar. Furthermore, since $c^2 > 0$ it follows that u is a time-like four-vector.

u^0 is the time-like and u^k a space-like component of the velocity four-vector. Its slope in SpaceTime is, therefore, $\frac{\gamma c}{\gamma v_x} = \frac{c}{v_x}$, taking only one spatial direction for simplicity. However, from the discussion around Fig. (1.10) it is evident that this slope corresponds to the slope of the worldline of a moving particle in SpaceTime which is also $\frac{c(t_1 - t_0)}{v_x(t_1 - t_0)} = \frac{c}{v_x}$. We conclude that the velocity four-vector is tangent to the worldline of a moving particle in SpaceTime which explains why the four-vector u is time like.

[25] Bear in mind that we here have the contravariant component of a four-vector on the left-hand side and the component of usual velocity in non-relativistic notation on the right-hand side.

1.3.2.2 Four-Acceleration

Following the same principles, it is a straightforward exercise to formulate four-acceleration.

$$\begin{aligned} \{b^\mu\} &= \left\{ \frac{du^\mu(t)}{d\tau} \right\} \\ b^\mu &= \frac{du^\mu(t)}{d\tau} = \frac{du^\mu(t)}{dt}\,\frac{dt}{d\tau} = \gamma(v)\,\frac{du^\mu(t)}{dt} \end{aligned} \tag{1.99}$$

Using the results for four-velocity we can calculate the individual components of four-acceleration.

$$\begin{aligned} b^0 &= \gamma(v(t))\,\frac{c\,d\gamma(v(t))}{dt} = \gamma(v(t))\,c\,\frac{d\gamma(v)}{dv}\,\frac{dv}{dt} \quad \text{with} \quad v = ||\mathbf{v}|| \\ & \frac{d||\mathbf{v}||}{dt} = \frac{1}{||\mathbf{v}||}\left(\mathbf{v}\cdot\frac{d\mathbf{v}}{dt}\right) \\ & \frac{d\gamma(v)}{dv} = \gamma^3\frac{||\mathbf{v}||}{c^2} \\ \Rightarrow b^0 &= \frac{\gamma^4}{c}\left(\mathbf{v}\cdot\frac{d\mathbf{v}}{dt}\right) \end{aligned}$$

Similarly,

$$\begin{aligned} b^k &= \gamma(v)\,\frac{du^k}{dt} \\ &= \gamma(v)\,\frac{d}{dt}\big(\gamma(v(t))\,v_k(t)\big) \\ &= \gamma(v)\,\left\{\left[\frac{||\mathbf{v}||}{c^2}\,\gamma^3\,v_k(t)\,\frac{1}{||\mathbf{v}||}\left(\mathbf{v}\cdot\frac{d\mathbf{v}}{dt}\right)\right] + \gamma a_k\right\} \\ &= \gamma(v)\,\left[\frac{\gamma^3}{c^2}\,v_k\left(\mathbf{v}\cdot\frac{d\mathbf{v}}{dt}\right) + \gamma a_k\right] \\ b^k &= \frac{\gamma^4}{c^2}\,v_k\left(\mathbf{v}\cdot\frac{d\mathbf{v}}{dt}\right) + \gamma^2 a_k. \end{aligned}$$

Summary for the acceleration four-vector:

$$\{b^\mu\} = \left\{ \begin{array}{c} \frac{\gamma^4}{c}\left(\mathbf{v}\cdot\frac{d\mathbf{v}}{dt}\right) \\ \frac{\gamma^4}{c^2}\, v_k\left(\mathbf{v}\cdot\frac{d\mathbf{v}}{dt}\right) + \gamma^2 a_k \end{array} \right\} \qquad \forall k \in \{1,\ldots,3\} \tag{1.100}$$

In the non-relativistic limit, we simply obtain:

$$\begin{aligned} \lim_{c\to\infty} b^0 &= 0 \\ \lim_{c\to\infty} b^k &\equiv a_k \end{aligned}$$

It can be shown that $b^\mu b_\mu < 0$ which means that four-acceleration is a space-like four-vector. However, it isn't really necessary to prove this since we can derive the result indirectly:

Using Eq. (1.98) we can also determine the Minkowski scalar product between four-velocity and four-acceleration. First it is noted that

$$\begin{aligned} \frac{d}{d\tau}\left(u^\mu u_\mu\right) &= \left(\frac{d}{d\tau}\,u^\mu\right) u_\mu + u^\mu\,\left(\frac{d}{d\tau}\,u_\mu\right) \\ &= \left(\frac{d}{d\tau}\,u^\mu\right) u_\mu + u^\mu\,\left(\frac{d}{d\tau}\,g_{\mu\nu}\,u^\nu\right) \\ &= \left(\frac{d}{d\tau}\,u^\mu\right) u_\mu + u^\nu\,g_{\nu\mu}\,\left(\frac{d}{d\tau}\,u^\mu\right) \\ &= \left(\frac{d}{d\tau}\,u^\mu\right) u_\mu + u_\mu\,\left(\frac{d}{d\tau}\,u^\mu\right) \\ &= 2b^\mu u_\mu \end{aligned}$$

$$b^\mu u_\mu = \frac{1}{2}\frac{d}{d\tau}\left(u^\mu u_\mu\right) = \frac{1}{2}\frac{d}{d\tau}c^2 = \frac{\gamma}{2}\,\frac{d}{dt}\,c^2 = 0 \tag{1.101}$$

This means that in Minkowski SpaceTime four-velocity and four-acceleration are always orthogonal four-vectors.