SPECIAL RELATIVITY AND SPACETIME GEOMETRY

Routes to Special Relativity

At the turn of the $19^{\text{th}}/20^{\text{th}}$ centuries, several physicists worked on what eventually became special relativity theory from rather different approaches. Historically, Albert Einstein was the first in 1905. However, he always said that if not for himself, then Hendrick Lorentz or Henri Poincare would have reached the same results within a year or two. So let me briefly summarize their approaches:

1. Hendrick Lorentz — and independently George FitzGerald — started by explaining the negative result of the Michelson–Morley experiment by arguing that the aether wind causes length contraction in the direction of the wind. They did not have a good mechanist for this contraction but simply worked by analogy with the ellipsoid deformation of the Coulomb electric field of a rapidly moving charge. But later they argued that IF all the inter-atomic forces are of electromagnetic origin OR behave like the electromagnetic forces THEN the macroscopic bodies would indeed be contracted in the direction of their motion (relative to the aether).

   Next, Lorentz and Joseph Larmor found that the electron's motion in a rapidly moving body slows down by the factor $1/\gamma = \sqrt{1-(v/c)^2}$, so presuming that all macroscopic processes are related to the electrons' motion they argued for the time dilation. Consequently, a Michelson–Morley-like experiment with unequal-length arms (which was not performed until 1931) should also see no aether wind.

   Given the length contraction and the time dilation, Lorentz *et al* worked out the spacetime coordinate transformations between the frame of the aether and the frame of a moving body. Next came the coordinate transforms between two different inertial frames, and miraculously, the aether frame completely dropped out from this transform, which made the aether unnecessary and irrelevant. Also, the Maxwell equations turned out to be invariant under the Lorentz transforms.

2. Henri Poincare started with Lorentz transforms and saw that they form a group related by the analytic continuation from a 4-dimensional space to the 4-dimensional *spacetime*.

In the process, Poincare worked out the relativistic velocity addition formula

$$v_{o2} \;=\; \frac{v_{01}+u_{12}}{1+\frac{v_{01}u_{12}}{c^2}} \tag{1}$$

and corrected the Lorentz's transformation rules for the charge density and the current density. Thanks to these corrections, the Maxwell equations became Lorentz-invariant to all orders in v/c rather than just to the second order.

Continuing Poincare's (and Einstein's) work, Hermann Minkowski worked out the *pseudo–Euclidean* geometry of the spacetime. I shall explain this geometry is some detail later in these notes.

3. Albert Einstein did not even know about the Michelson–Morley experiment before 1905. Instead, he simply wanted to reconcile the electrodynamics with the relativity principle, so he postulated that the relativity principle must be universal but also the speed of light in vacuum should also be universal. To resolve the obvious paradox between the two postulates, he had to give up the universality of the time; instead the time runs at different rates in different frames of reference. Gives these presumptions, a few *thought experiments* lead him to the Lorentz contraction and time dilation and hence to the Lorentz transforms.

EINSTEIN'S POSTULATES

Albert Einstein developed his special relativity theory starting with two basic postulates, although later he added a third postulate of causality.

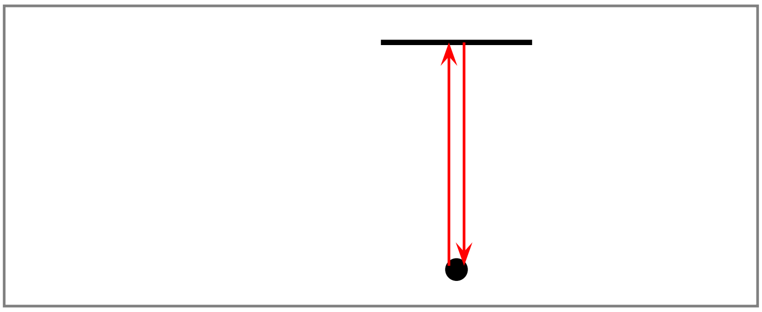
- **Postulate I:** *relativity of motion*: All laws of nature must be equally valid in all inertial frames of reference.
- **Postulate II:** *universality of* c: In any inertial frame, the light — any any other kind of an EM wave — propagates through the vacuum at the same speed c in all direction, regardless of the light source's velocity.
- **Postulate III:** *relativistic causality*: No material body nor any signal may travel faster than light in the vacuum.

Einstein's *thought experiments* involved an observer riding an ultra-fast train — so fast its speed u is comparable to the speed of light,[⋆] — and another observer standing on the ground, and comparing what the two observers see.

- ⋆ Thought experiment#1: observing the train itself. Both observers — on the train and on the ground — see that the train's wheels stay on the rails, so the wheel base must be equal to the railway's gauge. However, one observer sees the train in motion and the rails at rest while the other sees the train at rest while the rails run back, so if there were any effect of the motion on the dimensions $\perp$ to the velocity, one observer would see the changed gauge while the other would see the changed wheel base. Either way, the train would not stay on the rails. Except it does, which both observers can confirm. Consequently, *the dimension $\perp$to the velocity are not affected by the motion.* Thus, for the train moving in x direction,

$$y_{\text{train}} = y_{\text{ground}} \quad \text{and} \quad z_{\text{train}} = z_{\text{ground}} \,. \tag{2}$$

- ⋆ Thought experiment#2: A light beam bounces off the mirror back to its source across the train. Schematically, in the frame of the train we have



(3)

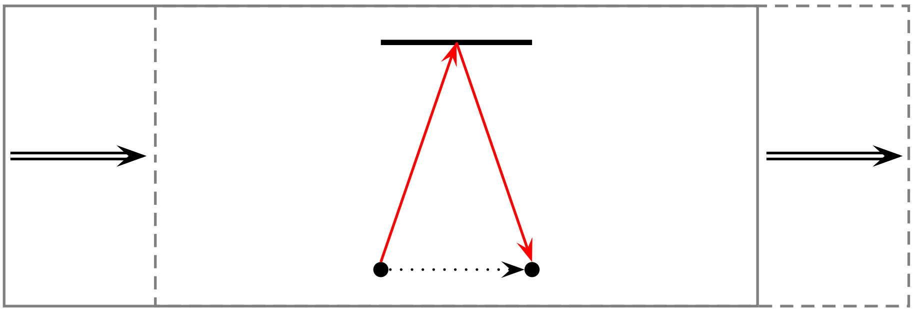
Consequently, the light travels the path of length $2L$ and takes time

$$t_{\text{train}} = \frac{2L}{c} \tag{4}$$

Now consider the same experiment in the ground frame. As the light bounces back

---

[⋆] Nowadays we would use a starship rather than a train, but Einstein wrote in 1905 and forth in the y direction, the train keeps moving in the x direction, thus



(5)

The distance in the y direction travelled by the light is the same $2L$ as in the train frame — *cf.* thought experiment#1, but in the ground frame the light has to travel in both y and x directions. Specifically, to get back to its source, the x component of of the light's velocity should be the same as the train velocity u, $v_x = u$. At the same time, the net speed of the light should be c (by the second postulate), thus

$$v_x^2 \;+\; v_y^2 \;=\; c^2 \tag{6}$$

and hence

$$|v_y| \;=\; \sqrt{c^2 - v_x^2} \;=\; \sqrt{c^2 - u^2}\,. \tag{7}$$

Consequently, in the ground frame the time it takes the light to travel back and forth is

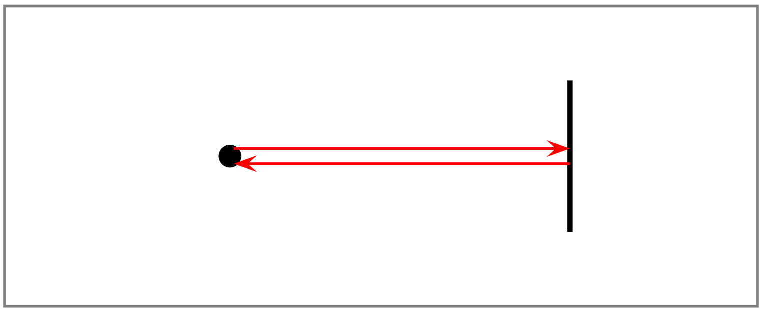
$$t_{\text{ground}} \;=\; \frac{2L}{|v_y|} \;=\; \frac{2L}{\sqrt{c^2 - u^2}}\,. \tag{8}$$

Comparing the times between the same events (emission and detection of light) in the two frames of reference, we see that

$$\frac{\mathrm{t}_{\text{ground}}}{\mathrm{t}_{\text{train}}} \;=\; \frac{c}{\sqrt{c^2 - u^2}} \;\overset{\text{def}}{=}\; \gamma \;>\; 1. \tag{9}$$

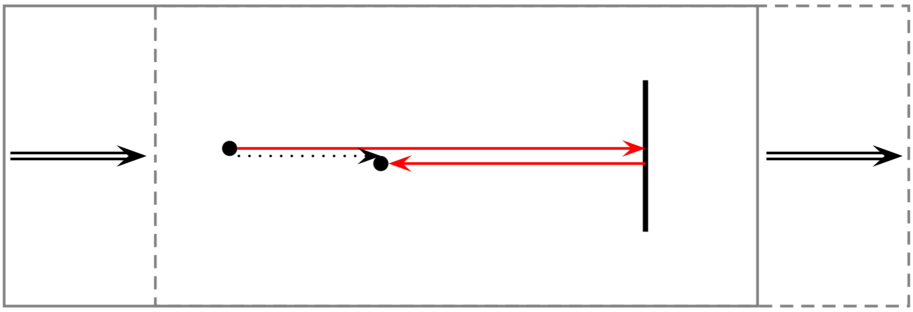
This is precisely the Lorentz's time dilation factor.

- ⋆ Thought experiment#3: Again, bounce light beam of a mirror, but this time along the train's motion rather than across it. In the train's frame



$$t_{\text{train}} = \frac{2L_{\text{train}}}{c}, \tag{10}$$

while in the ground frame



(11)

hence

$$t_{\text{ground}} = \frac{L_{\text{ground}}}{c-u} + \frac{L_{\text{ground}}}{c+u} = \frac{2cL_{\text{ground}}}{c^2-u^2} = \gamma^2 \times \frac{2L_{\text{ground}}}{c} \tag{12}$$

where the length L_{ground} in the ground frame may be different from the length L_{train} in the train frame. Indeed, comparing eqs. (10) and (12), we get

$$\frac{t_{\text{ground}}}{t_{\text{train}}} = \gamma^2 \times \frac{L_{\text{ground}}}{L_{\text{train}}}, \tag{13}$$

while according to the thought experiment#2

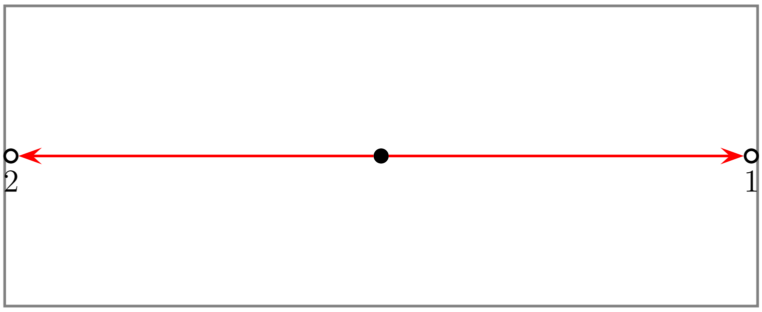
$$\frac{\mathrm{t}_{\text{ground}}}{\mathrm{t}_{\text{train}}} = \gamma \qquad \text{rather than } \gamma^2, \tag{14}$$

hence

$$\frac{L_{\text{ground}}}{L_{\text{train}}} = \frac{1}{\gamma} < 1. \tag{15}$$

Note: this is precisely that same length contraction factor as in Lorentz–FitzGerald theory, but without any reference to the aether.

- ⋆ Thought experiment#4: Put a flashlight in the middle of the train care and two light detectors at the two ends of the car, one on the locomotive side and the other on the caboose side. In the frame of the train

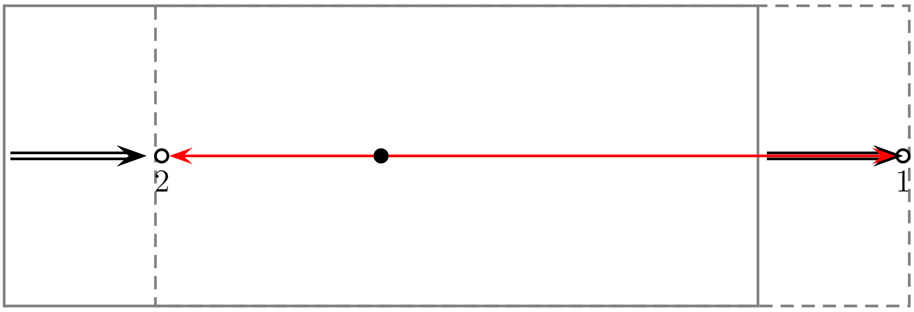


(16)

the light reaches the two detectors at the same time,

$$t_1 \;=\; t_2 \;=\; t_0 \;+\; \frac{L}{c}\,. \tag{17}$$

But in the ground frame



(18)

the back detector 2 moves towards the light while the front detector moves away from the light, so the front detector is lit after the back detector:

$$t_1 \;=\; t_0 \;+\; \frac{L}{c-v} \;>\; t_0 \;+\; \frac{L}{c+v} \;=\; t_2. \tag{19}$$

Thus, *simultaneity is relative to the frame of reference*: Two events may appear simultaneous in one frame, while in another frame one event may happen before the other.

Lorentz Transforms

The two Einstein postulates — the universality of Physics Laws in all inertial frames, and the universality of light speed in the vacuum in all inertial frames — are inconsistent with the universality of Time. Instead, the time must run at different rates in different inertial frames. Moreover, even the simultaneity of two events is relative to a particular frame, so the relation between times t and t' in some frames K and K' cannot be a relation between the two times alone, $t' = f(t)$, but must also involve the space coordinates, thus $t' = f(t, x, y, z)$, because otherwise the simultaneity would be absolute. Therefore, changing a frame of reference mixes the time and the space coordinates with each other, just like a rotation mixes up the space coordinates with each other. Mathematically, this means that the 3D space and the time combine into the 4D *spacetime*, and changing the frame of reference amounts to a 4D coordinate transform in that spacetime.

Specifically, for the inertial frames K and K' moving at velocity $\mathbf{u}$ relative to each other in x direction, the spacetime coordinate transform — called the *Lorentz transform* or the *Lorentz boost* — works according to

$$x' = \gamma(x - ut), \quad y' = y, \quad z' = z, \quad t' = \gamma\left(t - \frac{u}{c^2}\,x\right) \tag{20}$$

in one direction, and

$$x = \gamma(x' + ut'), \quad y = y', \quad z = z', \quad t = \gamma\left(t' + \frac{u}{c^2}\,x'\right) \tag{21}$$

in the opposite direction. For both directions,

$$\gamma = \frac{1}{\sqrt{1 - (u/c)^2}}\,. \tag{22}$$

The direct and the reverse Lorentz transforms look completely similar (except for the sign of u) — in particular, they both have $\partial t'/\partial t = \gamma > 1$ and $\partial t/\partial t' = \gamma > 1$, — but once we apply them to a physical body which is at rest in one frame but moves in the other frame, we break the symmetry between the direct and the reverse transform. Consequently, we find that *in the rest frame of a body, its time runs slower than in any other frame.* Indeed, in the rest frame $x_{\rm rest}\equiv 0$, hence in the lab frame where the body is moving at velocity v,

$$t_{\rm lab}\;=\;\gamma_v\left(t_{\rm rest}\,+\,\frac{v}{c^2}\times x_{\rm rest}\right)\;=\;\gamma_v\times t_{\rm rest}\;=\;\frac{t_{\rm rest}}{\sqrt{1-(v/c)^2}}\,. \tag{23}$$

This is the famous *relativistic time dilation.*

Note: the reverse transform from the lab frame to the rest frame looks different because in the lab frame $x_{\rm lab}\;=\;v\times t_{\rm lab}$, hence

$$t_{\rm rest}\;=\;\gamma_v\left(t_{\rm lab}\,-\,\frac{v}{c^2}\times(x_{\rm lab}=vt_{\rm lab})\right)\;=\;\gamma_v\times\left(1\,-\,\frac{v^2}{c^2}\right)\times t_{\rm lab}\;=\;\sqrt{1-(v/c)^2}\times t_{\rm lab}\,. \tag{24}$$

Thus, both directions of the Lorentz transform produce the same result: the time in the rest frame runs slower than the time in the lab frame.

Similar arguments apply to the *Lorentz contraction of length: a body viewed in the frame where it moves looks shorter than it is in its rest frame.* To see how this works, note that the length of a body is the distance between its ends at the same instance of time. Thus, in the lab frame where the body moves at velocity v, its two ends are at

$$x^{(1)}_{\rm lab}\;=\;v\times t_{\rm lab}\,,\quad x^{(2)}_{\rm lab}\;=\;v\times t_{\rm lab}\;+\;L_{\rm lab}\,, \tag{25}$$

so that

$$x^{(2)}_{\rm lab}(t_{\rm lab})\;-\;x^{(1)}_{\rm lab}(t_{\rm lab})\;=\;L_{\rm lab}\ \text{at all}\ t_{\rm lab}\,. \tag{26}$$

In the rest frame, this translates to

$$x^{(1)}_{\rm rest}\;=\;\gamma_v(x^{(1)}_{\rm lab}\;-\;vt_{\rm lab})\;=\;0,\quad x^{(2)}_{\rm rest}\;=\;\gamma_v(x^{(2)}_{\rm lab}\;-\;vt_{\rm lab})\;=\;\gamma_v\times L_{\rm lab} \tag{27}$$

hence $L_{\rm rest}=\gamma_v\times L_{\rm lab}$, or equivalently the Lorentz contraction

$$L_{\rm lab}\;=\;\frac{L_{\rm rest}}{\gamma_v}\;=\;\sqrt{1-(v/c)^2}\times L_{\rm rest}\;<\;L_{\rm rest}\,. \tag{28}$$

The reverse transform yields the same result — the length is shorter in the lab frame — but to see it we need to pay attention to the simultaneity of the two ends. Indeed, given

$x^{(1)}_{\rm rest} \equiv 0$ and $x^{(2)}_{\rm rest} \equiv L_{\rm rest}$, we get

$$x^{(1)}_{\rm lab} \;=\; \gamma_v v \times t^{(1)}_{\rm rest}\,, \quad x^{(2)}_{\rm lab} \;=\; \gamma_v \times L_{\rm rest} \;+\; \gamma_v v \times t^{(2)}_{\rm rest}\,, \tag{29}$$

where the times $t^{(1)}_{\rm rest}$ and $t^{(2)}_{\rm rest}$ must be chosen such that the corresponding lab-frame times

$$t^{(1)}_{\rm lab} \;=\; \gamma_v \times t^{(1)}_{\rm rest} \quad\text{and}\quad t^{(2)}_{\rm lab} \;=\; \gamma_v \times t^{(2)}_{\rm rest} \;+\; \frac{\gamma_v v}{c^2} \times L_{\rm rest} \tag{30}$$

are equal to each other. Thus,

$$t^{(2)}_{\rm rest} \;-\; t^{(1)}_{\rm rest} \;=\; -\frac{v}{c^2} \times L_{\rm rest} \tag{31}$$

and hence

$$\begin{aligned} L_{\rm lab} \;&=\; x^{(2)}_{\rm lab} \;-\; x^{(1)}_{\rm lab} \;=\; \gamma_v \times L_{\rm rest} \;+\; \gamma_v v \bigl(t^{(2)}_{\rm rest} - t^{(1)}_{\rm rest}\bigr)\\ &=\; \gamma_v \times \left(1 \;-\; \frac{v^2}{c^2}\right) \times L_{\rm rest} \;=\; \sqrt{1-(v/c)^2} \times L_{\rm rest} \;<\; L_{\rm rest}\,. \end{aligned} \tag{32}$$

Relativistic velocity addition rule follows from the Lorentz transforms between the frames. Suppose a body moves at velocity $\mathbf{v}'$ relative to the frame K', which in turn moves at velocity $\mathbf{u}$ relative to the frame K. Then the body's velocity $\mathbf{v}$ relative to the frame K is not $\mathbf{u}+\mathbf{v}'$ but given by a more complicated formula. In this section, I'll write this formula for $\mathbf{u}$ and $\mathbf{v}'$ being in the same direction x (or in the opposite directions for $v' < 0$); I'll come back to the more general case in the later section about the *proper velocity* 4–vector.

In the K' frame $x' = v' \times t'$. Translating these spacetime coordinates to the K frame, we get

$$\begin{aligned} x \;&=\; \gamma_u (x' + u \times t') \;=\; \gamma_u \times (v' + u) \times t',\\ t \;&=\; \gamma_u \left(t' \;+\; \frac{u}{c^2} \times x'\right) \;=\; \gamma_u \times \left(1 \;+\; \frac{uv'}{c^2}\right) \times t', \end{aligned} \tag{33}$$

and therefore

$$x \;=\; \frac{u+v'}{1 \;+\; \frac{uv'}{c^2}} \times t. \tag{34}$$

This gives us the body's velocity v relative to the K frame as

$$v \;=\; \frac{u+v'}{1+(uv'/c^2)}\,. \tag{35}$$

This velocity addition formula may look strange, but that's what it takes to implement Einstein's second postulate. Indeed, for the light which moves at velocity $v'=\pm c$ relative to the frame K', it's velocity v relative to the K frame is also

$$v \;=\; \frac{u\pm c}{1\,\pm\,(u/c)} \;=\; \pm c. \tag{36}$$

Also, for the light moving through some transparent material at a reduced speed c/n relative to that material, eq. (35) leads to the Fresnel–Fizeau formula for the speed of light relative to lab frame in which the transparent material moves at velocity $u\ll c$. Indeed,

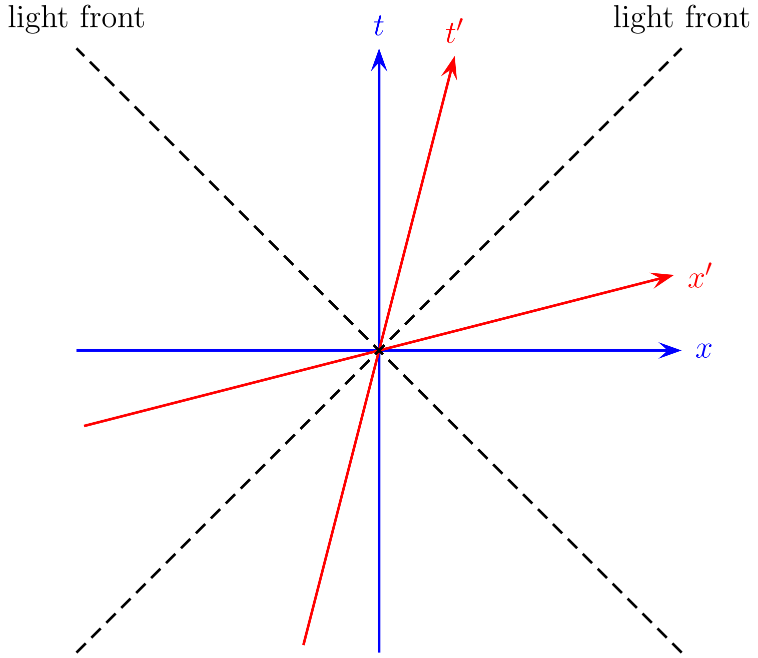
$$\begin{aligned} v \;&=\; \frac{\frac{c}{n}\,+\,u}{1\,+\,\frac{u}{nc}} \\ &=\; \left(\frac{c}{n}\,+\,u\right)\times\left(1\,-\,\frac{u}{nc}\,+\,O(u^2/c^2)\right) \\ &=\; \frac{c}{n}\,+\,u\,-\,\frac{u}{n^2}\,+\,O(u^2/c^2) \end{aligned} \tag{37}$$

hence

$$v\,-\,\frac{c}{n} \;=\; \frac{n^2-1}{n^2}\times u\,+\,O(u^2/c^2) \;\approx\; \frac{n^2-1}{n^2}\times u. \tag{38}$$

Spacetime Geometry

Geometrically, the Lorentz transform between the frames K and K' becomes a *pseudo–Euclidean rotation* of the x and t coordinate axes:



Note that the transform $(x,t) \to (x',t')$ tilts the two coordinate axes towards the same diagonal rather that rotates both of them in the same direction — that's why we call it *pseudo*–Euclidean.

The reason we compare a Lorentz transform to a rotation is that they both preserve a quadratic invariant: A Euclidean rotation in the (x,y) plane

$$x' \;=\; \cos\phi\times x \;-\; \sin\phi\times y, \quad y' \;=\; \cos\phi\times y \;+\; \sin\phi\times x \tag{39}$$

preserves the radius^2,

$$r^2 \;=\; x^2 \;+\; y^2 \;=\; x'^2 \;+\; y'^2, \tag{40}$$

while the Lorentz boost in the x direction preserves the so-called interval^2,

$$I^2\ =\ c^2t^2\ -\ x^2\ =\ c^2t'^2\ -\ x'^2. \tag{41}$$

Indeed,

$$\begin{aligned} I'^2\ &=\ c^2t'^2\ -\ x'^2\\ &=\ c^2\gamma^2\left(t-\frac{v}{c^2}\times x\right)^2\ -\ \gamma^2(x-vt)^2\\ &=\ \gamma^2\left[\left(c^2\times t^2\ -\ 2t\times vx\ +\ \frac{v^2}{c^2}\times x^2\right)\ -\ \left(x^2\ -\ 2x\times vt\ +\ v^2t^2\right)\right]\\ &=\ \gamma^2(c^2-v^2)\times t^2\ +\ 0\times xt\ +\ \gamma^2\left(\frac{v^2}{c^2}\ -\ 1\right)\times x^2\\ &=\ \gamma^2\left(1\ -\ \frac{v^2}{c^2}\right)\times(c^2t^2-x^2)\\ &=\ 1\times(c^2t^2-x^2)\ =\ I^2. \end{aligned} \tag{42}$$

But the minus sign between $(ct)^2$ and x^2 in the definition of the invariant interval makes the spacetime geometry pseudo–Euclidean — also called Minkowski — rather than Euclidean.

Before we go any further with the Minkowski geometry, let's include all 3 dimensions of space and hence all 4 dimensions of spacetime. In 3–vector terms, the Lorentz boost between 2 inertial frames moving at velocity $\mathbf{u}$ relative to each other becomes

$$\begin{aligned} \mathbf{r}'\ &=\ \mathbf{r}_\perp\ +\ \gamma_u\mathbf{r}_\parallel\ -\ \gamma_u\mathbf{u}\,t\\ &=\ \mathbf{r}\ +\ \frac{\gamma_u-1}{u^2}\,(\mathbf{r}\cdot\mathbf{u})\mathbf{u}\ -\ \gamma_u\mathbf{u}\,t\,,\\ t'\ &=\ \gamma_u\left(t\ -\ \frac{\mathbf{u}\cdot\mathbf{r}}{c^2}\right), \end{aligned} \tag{43}$$

and all such boosts preserve the $(3+1)$–dimensional interval^2

$$I^2\ =\ (ct)^2\ -\ \mathbf{r}^2\ =\ (ct)^2\ -\ x^2\ -\ y^2\ -\ z^2. \tag{44}$$

Actually, a better definition of the interval involves a pair of *events*, *i.e.*, spacetime points — one at time t_1 and location $\mathbf{r}_1$ and the other at time t_2 and location $\mathbf{r}_2$. *The interval* I_{12}

between two such events is defined as

$$I_{12}^2 \;=\; c^2(t_2-t_1)^2 \;-\; (\mathbf{r}_2-\mathbf{r}_2)^2 \;=\; c^2(t_2-t_1)^2 \;-(x_2-x_1)^2 \;-\; (y_2-y_1)^2 \;-\; (z_2-z_1)^2, \tag{45}$$

and just like in eq. (44), this interval is the same in all reference frames. Mathematically, this is similar to the distance^2 $= (x_2-x_1)^2+(y_2-y_1)^2+(z_2-z_1)^2$ between two space points being the same in all coordinate systems.

But unlike the Euclidean distance, the interval^2 is not positive-definite, so there are 3 kinds of intervals according to the sign of I^2:

- timelike intervals $I_{12}^2 > 0$ with $c|t_2-t_1| > |\mathbf{r}_2-\mathbf{r}_1|$;
- spacelike intervals $I_{12}^2 < 0$, with $c|t_2-t_1| < |\mathbf{r}_2-\mathbf{r}_1|$;
- lightlike intervals $I_{12}^2 = 0$, with $c|t_2-t_1| = |\mathbf{r}_2-\mathbf{r}_1|$.

Theorem: *For two events with a timelike or lightlike interval between them, their time order — which is later than which — is the same in all frame of reference. But for a spacelike interval between two events, their time order depends on the frame of reference, $t_2 > t_1$ in one frame but $t'_2 < t'_1$ in another frame.*

Proof: Let $\Delta t = t_2 - t_1$ and $\Delta\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_2$ in some frame of reference K. Then in another frame K' moving at velocity $\mathbf{u}$ relative to K,

$$\Delta t' \;=\; \gamma_u\left(\Delta t \;-\; \frac{\mathbf{u}}{c^2}\cdot\Delta\mathbf{r}\right). \tag{46}$$

Suppose the interval $(\Delta t, \Delta\mathbf{r})$ is timelike or lightlike, $c^2\Delta t^2 \geq \Delta\mathbf{r}^2$. Then for any speed u slower than light

$$\left|\frac{\mathbf{u}}{c^2}\cdot\Delta\mathbf{r}\right| \;\leq\; \frac{u}{c}\times\frac{|\Delta\mathbf{r}|}{c} \;<\; 1\times|\Delta t|, \tag{47}$$

hence

$$\mathrm{sign}(\Delta t') \;=\; \mathrm{sign}\left(\Delta t \;-\; \frac{\mathbf{u}}{c^2}\cdot\Delta\mathbf{r}\right) \;=\; \mathrm{sign}(\Delta t). \tag{48}$$

Thus, for a timelike (or lightlike) interval between two events, their time order — either $\Delta t > 0$ and (2) is later than (1) or else $\Delta t < 0$ and (2) is earlier tan (1) — is the same in all frames of reference. Such frame-independent time order is called *absolute*.

Now consider a spacelike interval $(\Delta t, \Delta\mathbf{r})$ with $c|\Delta t| < |\Delta\mathbf{r}|$ in some frame K. Let

$$u_0 \;=\; \frac{c^2|\Delta t|}{|\Delta\mathbf{r}|}\,. \tag{49}$$

For a spacelike interval $u_0 < c$, so another frame K' may move faster than u_0 relative to K. Specifically, let's pick the K' frame which moves at speed $u > u_0$ in the direction of $\operatorname{sign}(\Delta t)\Delta\mathbf{r}$, then

$$\frac{\mathbf{u}\cdot\mathbf{r}}{c^2} \;=\; \operatorname{sign}(\Delta t)\times\frac{u\,|\Delta\mathbf{r}|}{c^2} \;=\; \operatorname{sign}(\Delta t)\times\frac{u}{u_0}\times\left(\frac{u_0|\Delta\mathbf{r}|}{c^2}\;=\;|\Delta t|\right) \;=\; \frac{u}{u_0}\times\Delta t. \tag{50}$$

Consequently, in the K' frame the time difference

$$\Delta t' \;=\; \gamma_u\left(\Delta t \,-\, \frac{\mathbf{u}}{c^2}\cdot\Delta\mathbf{r}\right) \tag{51}$$

becomes

$$\Delta t' \;=\; \gamma_u\left(1 \,-\, \frac{u}{u_0}\right)\times\Delta t\,, \tag{52}$$

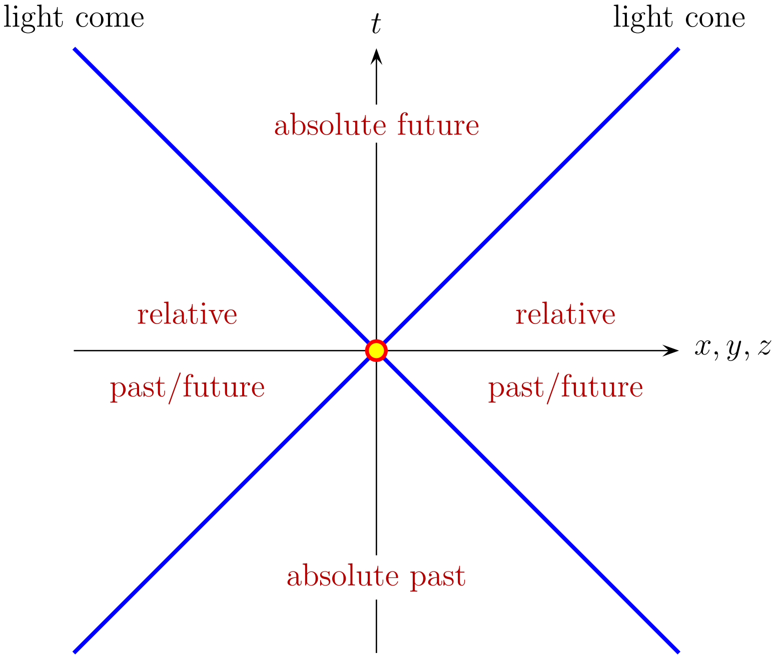
and since $u > u_0$, this $\Delta t'$ has the opposite sign from the Δt! Thus, if in the original frame K the event (2) happens later than the event (1), $\Delta t > 0$, then in the K' frame $\Delta t' < 0$ and the event (2) happens before the event (1). Likewise, if in the K frame (2) happens before (1) then in the K' frame (2) happens after (1). Either way, the time order of the events (1) and (2) is different in different reference frames; such frame-dependent time order is called *relative.*

Events at lightlike intervals from a given event $(t_0, \mathbf{r}_0)$ form a double cone in spacetime

$$t \;=\; t_0 \;\pm\; \frac{|\mathbf{r}-\mathbf{r}_0|}{c} \tag{53}$$

called the *light cone.* Physically, the light cone is spanned by all the light rays beginning or ending at point $\mathbf{r}_0$ at time t_0, hence the name. The light cone is invariant under all

Lorentz transforms (this is the second Einstein postulate), and it divides the spacetime into 3 causally distinct regions:



- The absolute future region
$$t \;>\; t_0 \;+\; \frac{|\mathbf{r}-\mathbf{r}_0|}{c} \tag{54}$$
comprises events which are later than $(t_0, \mathbf{r}_0)$ in all frames of reference.
- Likewise, the absolute past region
$$t \;<\; t_0 \;-\; \frac{|\mathbf{r}-\mathbf{r}_0|}{c} \tag{55}$$
comprises events which are earlier than $(t_0, \mathbf{r}_0)$ in all frames of reference.
- Finally, the relative past/future region
$$t_0 \;-\; \frac{|\mathbf{r}-\mathbf{r}_0|}{c} \;<\; t \;<\; t_0 \;+\; \frac{|\mathbf{r}-\mathbf{r}_0|}{c} \tag{56}$$
which are can be earlier or later than $(t_0, \mathbf{r}_0)$ depending on a reference frame.

- ∗ For example, consider an event which happens in α Centauri system in the year 2000 by the Earthly calendar. Relative to that event, Earth history prior to 1996 is absolute past, Earth history after 2004 is absolute future, but the 8 year period between 1996 and 2004 is relative past/future. For example, a 1999 event on Earth happens earlier than that α Centauri event in the frame of the Earth (or of the α Centauri), but in the frame of a spaceship flying from Earth to α Centauri at speed $u > \frac{1}{4}c$ the 1999 event on Earth would happen later than the 2000 event on α Centauri.

- ∘ Here is a better picture of the light cones.

Causality means that an event in the past can cause or influence an event in the future but not the other way around: the future cannot influence the past. Relativistically, causality has to work in all reference frames, so if the time order of two events is frame-dependent, then neither event can cause influence the other. Thus, *relativistic causality* means that *an event may cause or influence other events only in its absolute future.* Likewise, *an event can be caused or influenced only by events in its absolute past.* In terms of signals communicating between events, *no signal can travel faster than light in vacuum*, because the time order of sending and receiving such a signal would be frame-dependent:

$$\text{for } \mathbf{r}_2 - \mathbf{r}_1 \;=\; \mathbf{v}(t_2 - t_1) \text{ and } |\mathbf{v}| > c, \quad I_{12}^2 < 0, \tag{57}$$

and a spacelike interval between (1) and (2) means frame-dependent time order. Moreover, since any material body can act as a signal, *no material body can travel faster than light in vacuum.*

The relativistic causality is often called the third Einstein postulate, usually stated as *no material body nor any information can travel faster than the light in vacuum in any reference frame.*

RELATIVISTIC PARADOXES

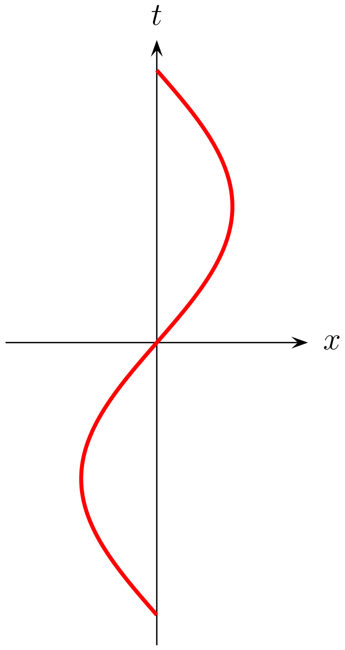
Missing sections:

- ⋆ **Barn and ladder paradox** about a ladder of (rest) length L flying into a barn of the same (rest) length L.

- ⋆ **Twin paradox** about the relative ages of an astronaut and his two who stays on Earth.

Worldlines and Proper Time

In spacetime terms, a moving particle spans a continuous family of events $(t, \mathbf{r}(t))$ parametrized by the time coordinate t; geometrically, this family is a line in 4D spacetime called the *worldline*. For a particle moving at constant velocity $\mathbf{v}$ the worldline is straight, while for an accelerating particle the worldline is curved. Here is an example worldline for a particle oscillating in x direction:



For massive particles or any macroscopic bodies, the velocity $d\mathbf{r}/dt$ is always slower than the speed of light, so the infinitesimal intervals $(dt, d\mathbf{r})$ along the worldline are always timelike,

$$dI^2 \;=\; c^2dt^2 \;-\; \mathbf{r}^2 \;=\; (c^2-\mathbf{v}^2)dt^2 \;>\; 0. \tag{58}$$

In the frame which happens to move at the same velocity $\mathbf{v}(t_0)$ as the particle at the moment t_0, the interval^2 is simply $c^2dt'^2$, so up to the overall factor c, the infinitesimal interval dI is the infinitesimal time in the particle's rest frame. If we replace the particle with a macroscopic body equipped with its own clock, then this clock would measure time

$$d\tau \;=\; \frac{dI}{c} \;=\; \sqrt{1-(v/c)^2}\times dt. \tag{59}$$

This time τ is called *the proper time* of the moving body/particle, and for a body moving at a variable velocity, the proper time obtains as an integral

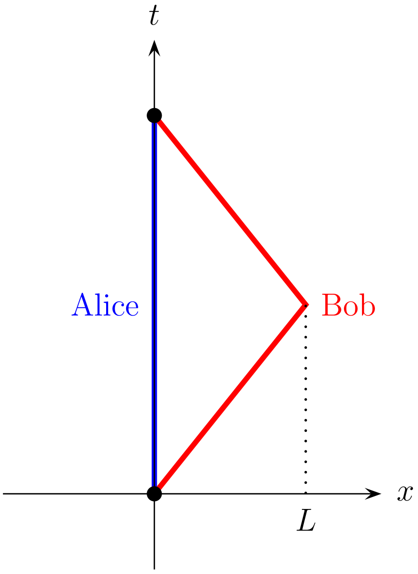
$$\tau \;=\; \int dt\,\sqrt{1\,-\,v^2(t)/c^2}\,, \tag{60}$$

or in worldline terms,

$$\tau \;=\; \int\limits_{\text{worldline}} \sqrt{(dt)^2\,-\,\frac{1}{c^2}(d\mathbf{r})^2}\,. \tag{61}$$

Note that although eq. (60) expresses the proper time of a moving body in terms of its motion in some reference frame, the actual proper time along the body's worldline is the same in all frames of reference.

Macroscopically, all processes on board a spaceship — from the clocks in shipboard computers to the astronauts' aging — happen according to the proper time of the ship. And that's how we resolve the twin paradox: Each twin ages by his/her proper time along his/her proper time, and those proper times do not depend on the frame of reference. For example, in the simplified example of astronaut's instant acceleration from zero to $+u$, then to $-u$ and finally to zero, we have (in the Earth frame)



$$\begin{aligned}
\tau_A &= \frac{L}{u}\,,\\
\tau_B &= \frac{L}{u}\sqrt{1-(u/c)^2}\,,\\
\tau_B &= \frac{\tau_A}{\gamma} \;<\; \tau_A\,.
\end{aligned} \tag{62}$$

For today's non-relativistic spacecraft, the difference between the on-board proper time and the Earth time is very small, but it's important to the GPS satellites and some other kind of space probes. Indeed, to act as navigation aid to people on Earth, the GPS satellites need very precise clocks synchronized with Earth clocks and with each other: An error of just a microsecond would lead to a navigation error of few thousands feet. To keep such precise time, the GPS satellites use atomic clocks bases on fixed and well-known frequencies of atomic transitions. But these clocks measure the proper time on board the satellite rather than the Earth time, so to keep them synchronized the on-board computers must correct for the relativistic time dilation. (And also for the gravitational corrections of general relativity, but that's a subject for a different class.)

The elementary particles move much faster that the satellites, so for them there is much bigger difference between the particle's proper time and the lab time. In particular, the average lifetimes of unstable particles are always in terms of the proper time of the moving particle. For example, the average lifetime of a muon is 2 microseconds of its proper time, but for a muon traveling at speed $v = 0.9998\,c$, $d\tau \approx 0.02\,dt$, so 2 μs of proper time stretch to 100 μs of lab time, long enough for the muon to travel 30 km from the stratosphere to the ground.

MOTION AT CONSTANT ACCELERATION

If we ever build a manned starship, we would want the astronauts on board to feel a constant artificial gravity similar to g for the duration of the flight. This means the ship should keep accelerating (or decelerating) at a constant rate $a \approx g$, so let's figure out what a constant acceleration means in relativistic context.

Suppose at time t_0 the ship moves at some velocity $\mathbf{v}_0$ relative to Earth, and consider its motion in the inertial frame K' that also moves at the velocity $\mathbf{v}_0$ relative to the Earth. In that frame, $\mathbf{v}'(t'_0) = 0$, so at times t' not too different from the t'_0 motion at constant acceleration $\mathbf{a}$ means

$$\mathbf{v}'(t') \approx (t' - t'_0)\mathbf{a} \quad \text{as long as} \quad |\mathbf{v}'| \ll c. \tag{63}$$

Note: it is in this frame that the astronauts feel the acceleration as artificial gravity, so it's in this frame that we should keep $|\mathbf{a}| \approx g$. Also, the time t' in this frame is precisely the proper time τ of the ship, and everything aboard the spaceship proceeds according to this proper time. In particular, the astronauts age according to the proper time τ.

Now, suppose the spaceship flies along a straight line — which we take to be the x axis, — so we may use the 1-dimensional formula for addition of velocities,

$$v = \frac{v_0 + \delta v'}{1 + \frac{v_0 \delta v'}{c^2}} \tag{64}$$

For

$$\delta v' = a\delta\tau \ll c, \tag{65}$$

eq. (64) becomes

$$v = v_0 + \delta v' - \frac{v_0^2}{c^2} \times \delta v' + O\big(\delta v'^2/c^2\big) \tag{66}$$

and hence

$$\delta v \approx \left(1 - \frac{v_0^2}{c^2}\right) \times a\delta\tau. \tag{67}$$

Consequently, *the ship's velocity v relative to the Earth as a function of the ship's proper time τ obtains as a solution of the differential equation*

$$\frac{dv}{d\tau} = a\left(1 - \frac{v^2}{c^2}\right). \tag{68}$$

To solve this equation, we rewrite it as

$$\frac{d\beta}{1 - \beta^2} = \frac{a}{c}\, d\tau \tag{69}$$

and integrate the two sides. Integrating the RHS side here is trivial,

$$\int \frac{a}{c}\, d\tau = \frac{a}{c} \times \tau, \tag{70}$$

while on the LHS

$$\int \frac{d\beta}{1 - \beta^2} = \frac{1}{2}\int\left(\frac{d\beta}{1+\beta} + \frac{d\beta}{1-\beta}\right) = \frac{1}{2}\Big(\ln(1+\beta) - \ln(1-\beta)\Big) = \operatorname{ar\,tanh}(\beta). \tag{71}$$

Thus, the solution of eq. (68) is

$$\operatorname{ar}\tanh(\beta) \;=\; \frac{a}{c}\times\tau \tag{72}$$

and therefore

$$v(\tau) \;=\; c\times\tanh\frac{a\tau}{c}. \tag{73}$$

Next, consider the ship's velocity as a function of the Earth time t. Infinitesimally, the Earth time and the proper time are related as

$$dt \;=\; \gamma\times d\tau \tag{74}$$

where

$$\begin{aligned}\gamma \;&=\; \frac{1}{\sqrt{1-\beta^2}}\,, \qquad \langle\!\langle\,\text{which for } v \text{ as in eq. (73) becomes}\,\rangle\!\rangle\\ &\longrightarrow\; \frac{1}{\sqrt{1-\tanh^2(a\tau/c)}} \;=\; \cosh(a\tau/c).\end{aligned} \tag{75}$$

Consequently

$$dt \;=\; \cosh\frac{a\tau}{c}\times d\tau \;=\; d\left(\frac{c}{a}\sinh\frac{a\tau}{c}\right), \tag{76}$$

hence

$$\frac{at}{c} \;=\; \sinh\frac{a\tau}{c}\,, \tag{77}$$

and therefore

$$\begin{aligned}\frac{v(t)}{c} \;&=\; \tanh(a\tau/c)\\ &=\; \frac{\sinh(a\tau/c)}{\sqrt{1+\sinh^2(a\tau/c)}} \;=\; \frac{(at/c)}{\sqrt{1+(at/c)^2}}\\ &=\; \frac{at}{\sqrt{c^2+(at)^2}}\,.\end{aligned} \tag{78}$$

In other words,

$$v(t) \;=\; \frac{at\times c}{\sqrt{c^2+(at)^2}}\,. \tag{79}$$

Finally, the distance travelled by the ship in a given Earth time t is

$$\Delta x = \int v(t)\,dt = \frac{c}{a}\int\frac{a^2t\,dt}{\sqrt{c^2+(at)^2}} = \frac{c}{a}\int d\Big(\sqrt{c^2+(at)^2}\Big) = \frac{c}{a}\Delta\Big(\sqrt{c^2+(at)^2}\Big). \tag{80}$$

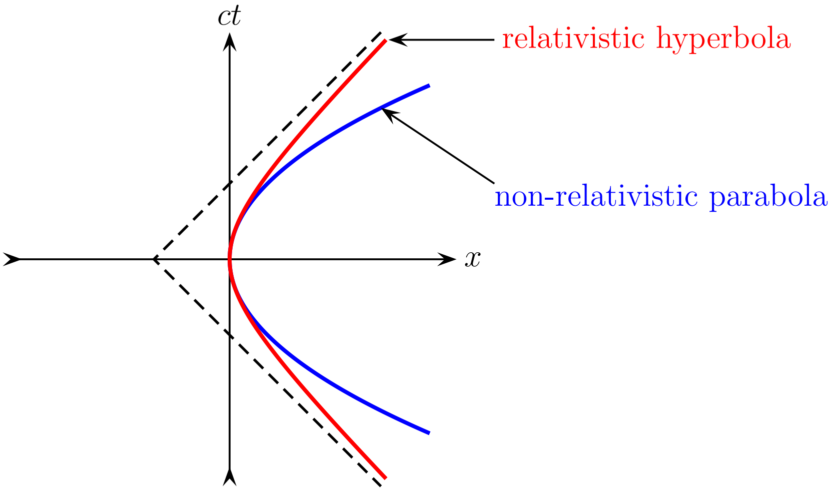
Thus, counting x from the Solar system and t from launch in the Solar system at zero initial velocity,

$$x(t) = \frac{c}{a}\Big(\sqrt{c^2+(at)^2}-c\Big) = \sqrt{(ct)^2+(c^2/a)^2} - (c^2/a). \tag{81}$$

Or in terms of the shipboard time τ,

$$x(\tau) = \frac{c^2}{a}\left(\cosh\frac{a\tau}{c}-1\right). \tag{82}$$

The worldline of the ship moving according to eq. (81) is a hyperbola so the relativistic motion at constant acceleration is sometimes called the *hyperbolic motion*, just like the non-relativistic motion at constant acceleration is called the *parabolic motion*. The diagram below shows the hyperbolic motion in red and the parabolic motion in blue:



(83)

To illustrate all of the above formulae for motion at constant acceleration, consider a starship flying to the Epsilon Eridani system, about 10.5 light-years away from the Solar system. Through the first half of this distance, the ship accelerates at $a = g = 9.8$ m/s^2, and then through the second half it decelerates at the same rate, so eventually it reaches the target system at zero velocity. (Relative to the Sun and also relative to the Epsilon Eridani). Thus, at half-point of the outbound flight we have

$$\frac{L}{2} \times \frac{g}{c^2} \;=\; \cosh\frac{g\tau}{c} \;-\; 1. \tag{84}$$

Numerically,

$$\frac{g}{c} \;\approx\; 1.03\ \mathrm{yr}^{-1}, \tag{85}$$

so half-way to the Epsilon Eridani

$$\frac{L}{2} \times \frac{g}{c^2} \;=\; \frac{10.5\ \mathrm{ly}}{2\mathrm{c}} \times \frac{1.03}{\mathrm{yr}} \;\approx\; 5.4, \tag{86}$$

hence

$$\frac{g\tau}{c} \;=\; \mathrm{ar}\cosh(5.4+1) \;\approx\; 2.54 \tag{87}$$

and therefore

$$\tau \;\approx\; 2.54 \times \frac{1\ \mathrm{yr}}{1.03}\ 2.47\ \mathrm{yr}. \tag{88}$$

This is the half-point time by the astronauts' clock. But the half-point time by the Earth clock is substantially longer:

$$\frac{gt}{c} \;=\; \sinh\frac{g\tau}{c} \;=\; \sinh(2.54) \;=\; 6.32, \tag{89}$$

hence

$$t \;=\; 6.32 \times \frac{1\ \mathrm{yr}}{1.03} \;=\; 6.14\ \mathrm{yr}. \tag{90}$$

And these are times to travel through a $\frac{1}{2}$ of the one-way journey, so the time to get to the Epsilon Eridani is twice as long by either clock: 4.94 years by the astronauts' clock and 12.28 years by the Earth clock.

Finally, the round trip journey — including 1 year spend in the Epsilon Eridani system — lasts $2 \times 4.94 + 1.00 = 10.88$ years by the astronauts' clock and $2 \times 12.28 + 1.00 = 25.56$ years by the Earth clock. Thus, if an astronaut was 30 years old when the journey started he would be about 41 years old when he comes back, but his twin brother (assuming he has one) would be 55.5 years old.

4–VECTORS

The 4 spacetime coordinates of an event can be combined into a 4–vector

$$x^\mu \;=\; (x^0, x^1, x^2, x^3) \;=\; (ct, x, y, z). \tag{91}$$

Other combinations of a 3–scalar and a 3–vector which transform similarly under Lorentz transformations also form 4–vectors; we shall see quite a few examples in the remaining weeks of this class. But before we go there, let me fix the notation conventions.

- The components of a 4–vector like x^μ are indexed by lower-case Greek letters, usually from the middle of the alphabet — μ and ν labels are particularly common. The Latin indices i, j, k, ℓ are used for components of 3–vectors rather than 4–vectors.
- The 4–vector indices $\mu, \nu, \ldots$ take values $0, 1, 2, 3$. The (A^1, A^2, A^3) components of a 4–vector A^μ comprise a 3–vector $\mathbf{A}$, and under rotations of the 3D space they indeed behave as components of a 3–vector. The A^0 component is invariant under space rotations, so it's a 3–scalar.
- A 4–vector index can be upper or lower, and it makes a difference, $A_\mu \neq A^\mu$. It's possible to trade an upper index for a lower index or vice verse — and in a moment I'll explain how, — but this changes the components and not just the typography!

⋆ Einstein summation convention: If in a product of 4–vectors or tensors the same index appears twice — once upstairs and once downstairs — than there is implicit summation over that index. For example,

$$A^\mu B_\mu \;=\; \sum_{\mu=0,1,2,3} A^\mu B_\mu \,. \tag{92}$$

However, there is no implicit summation over two similar upper indices, or two similar lower indices, or indices appearing more then twice. In all such cases, you should explicitly indicate whether you want that index to be summed over or not. Although most commonly, such malformed indices are simply typos and need to be corrected.

In 3D vector notations, a rotation of the coordinate system can be summarized as $x'_i = R_{ij}x_j$ for some orthogonal 3×3 matrix R_{ij}. Likewise, a Lorentz transform of the 4 spacetime coordinates can be summarized in 4–vector notations as $x'^{\mu} = L^{\mu}{}_{\nu}x^{\nu}$ for some *pseudo-orthogonal* 4×4 matrix L^{μ}_{ν}. For example, for a Lorentz boost of velocity v in the x^1 direction

$$\begin{pmatrix} x'^0 \\ x'^1 \\ x'^2 \\ x'^3 \end{pmatrix} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \tag{93}$$

where

$$\beta = \frac{v}{c} \quad \text{and} \quad \gamma = \frac{1}{\sqrt{1-\beta^2}}. \tag{94}$$

or in terms of explicit matrix elements,

$$L^0{}_0 = L^1{}_1 = \gamma, \quad L^0{}_1 = L^1{}_0 = -\beta\gamma, \quad L^2{}_2 = L^3{}_3 = 1, \quad \text{other } L^{\mu}{}_{\nu} = 0. \tag{95}$$

The components of all other 4–vectors — also called Lorentz vectors — must transform exactly like the coordinates $x^{\mu} = (ct, x^1, x^2, x^3)$, namely $A'^{\mu} = L^{\mu}{}_{\nu}A^{\nu}$ for exactly the same $L^{\mu}{}_{\nu}$ matrix as the coordinates. Otherwise, it would not be a Lorentz vector but just an array of 4 numbers.

Let's make another parallel between 3–vectors and 4–vectors. Rotations of 3-space leave invariant the length^2 of any vector, or more generally a dot product of any two vectors,

$$\mathbf{a}\cdot\mathbf{b} = a_1b_1 + a_2b_2 + a_3b_3 = a'_1b'_1 + a'_2b'_2 + a'_3b'_3\,. \tag{96}$$

Likewise, the Lorentz transforms leave invariant the interval^2, or similar quadratic combination of other 4–vectors like

$$(A)^2 = (A^0)^2 - (A^1)^2 - (A^2)^2 - (A^3)^2 = (A'^0)^2 - (A'^1)^2 - (A'^2)^2 - (A'^3)^2. \tag{97}$$

or more generally, the Lorentzian dot product or any two 4–vectors

$$(A\cdot B)\ =\ A^0B^0\ -\ A^1B^1\ -\ A^2B^2\ -\ A^3B^3\ =\ A'^0B'^0\ -\ A'^1B'^1\ -\ A'^2B'^2\ -\ A'^3B'^3. \tag{98}$$

In index notations,

$$(A\cdot B)\ =\ A^\mu g_{\mu\nu}B^\nu \tag{99}$$

where $g_{\mu\nu}$ is the *metric tensor*,

$$g_{00}\ =\ +1,\quad g_{11}\ =\ g_{22}\ =\ g_{33}\ =\ -1,\quad \text{and for } \mu\neq\nu\quad g_{\mu\nu}\ =\ 0, \tag{100}$$

or in matrix form,

$$g_{\mu\nu}\ =\ \begin{pmatrix} +1 & 0 & 0 & 0\\ 0 & -1 & 0 & 0\\ 0 & 0 & -1 & 0\\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{101}$$

Note: the metric tensor is invariant under Lorentz transforms, just like the Kronecker δ_{ij} tensor is invariant under space rotations.

As a matrix (101), the metric tensor squares to one, so its matrix inverse g^{-1} is the same as g. But in order to properly contract the Lorentz indices, the inverse metric tensor is written as

$$g^{\mu\nu}\ =\ \begin{pmatrix} +1 & 0 & 0 & 0\\ 0 & -1 & 0 & 0\\ 0 & 0 & -1 & 0\\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{102}$$

with upper rather than lower indices, so that we may express the matrix relation $gg = 1$ as

$$g^{\lambda\mu}g_{\mu\nu}\ =\ \delta^\lambda_\nu\,.$$

The metric tensor and its inverse allow us to raise and lower the Lorentz indices of 4–vectors and tensors according to

$$A_\mu\ =\ g_{\mu\nu}A^\nu,\quad A^\lambda\ =\ g^{\lambda\mu}A_\mu\,, \tag{103}$$

Since $g_{\mu\nu}$ and $g^{\mu\nu}$ are diagonal matrices with eigenvalues $(+1,-1,-1,-1)$, raising or lowering a Lorentz index amounts to a simple sign rule: keep the time component the same but change the signs of the space components,

$$A_0 = +A^0, \quad \text{but} \quad A_1 = -A^1, \quad A_2 = -A^2, \quad A_3 = -A^3, \tag{104}$$

for example,

$$x^\mu = (+ct, +x, +y, +z) \quad \text{but} \quad x_\mu = (+ct, -x, -y, -z). \tag{105}$$

<u>Sign convention</u>: whenever a 3–scalar A^0 and a 3–vector $\mathbf{A}$ combine into a 4–vector, we identify the (A^x, A^y, A^z) components of the 3-vector as the (A^1, A^2, A^3) components of the 4–vector A^μ with an upper rather than lower index, thus

$$A^\mu = (A^0, A^x, A^y, A^z) \quad \text{but} \quad A_\mu = (A^0, -A^x, -A^y, -A^z). \tag{106}$$

This sign convention works for most 4–vectors, except for the derivative 4-vector which combines the space derivative vector ∇ with the time derivative. For the derivative vector we let

$$\partial_\mu = \left(\frac{1}{c}\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \quad \text{while} \quad \partial^\mu = \left(\frac{1}{c}\frac{\partial}{\partial t}, -\frac{\partial}{\partial x}, -\frac{\partial}{\partial y}, -\frac{\partial}{\partial z}\right) \tag{107}$$

so that

$$\partial_\mu x^\nu = +\delta^\nu_\mu = \begin{cases} +1 & \text{for } \mu = \nu, \\ 0 & \text{for } \mu \neq \nu. \end{cases} \tag{108}$$

Raising and/or lowering the μ and ν indices here we also get

$$\partial^\mu x^\nu = g^{\mu\nu}, \quad \partial_\mu x_\nu = g_{\mu\nu}, \quad \text{and} \quad \partial^\mu x_\nu = \delta^\mu_\nu. \tag{109}$$

I'll come back to the derivative 4-vector in a moment, but first let me get to the whole point of raising or lowering the 4–vector indices: It allows writing the Lorentzian dot product in a more compact form

$$(A \cdot B) = A^\mu g_{\mu\nu} B^\nu = A^\mu B_\mu = A_\nu B^\nu. \tag{110}$$

(We may also write $(A \cdot B)$ as $A_\mu g^{\mu\nu} B_\nu$, although that would not be any more compact than $A^\mu g_{\mu\nu} B^\nu$.)

I have already mentioned the Lorentz invariance of the metric tensor and hence of the dot product of two Lorentz vectors. This invariance serves as the very definition of the **Lorentz group** $O(3,1)$ (where $(3,1)$ stand for 3 dimensions of space and 1 of time). This symmetry group comprises all real 4×4 matrices $L^{\mu}{}_{\nu}$ — or equivalently, all linear transforms $x^{\mu}\to x'^{\mu}=L^{\mu}{}_{\nu}x^{\nu}$ — which preserve the Lorentzian dot product (99). In index notations, this means

$$(A'\cdot B')\;=\;A'^{\mu}g_{\mu\nu}B'^{\nu}\;=\;(L^{\mu}{}_{\alpha}A^{\alpha})g_{\mu\nu}(L^{\nu}{}_{\beta}B^{\beta})\;=\;A^{\alpha}(g_{\mu\nu}L^{\mu}{}_{\alpha}L^{\nu}{}_{\beta})B^{\beta}\tag{111}$$

should be equal to

$$(A\cdot B)\;=\;A^{\alpha}g_{\alpha\beta}B^{\beta}\tag{112}$$

for any two 4–vectors A^{α} and B^{β}, which obviously calls for

$$g_{\mu\nu}L^{\mu}{}_{\alpha}L^{\nu}{}_{\beta}\;=\;g_{\alpha\beta}\,.\tag{113}$$

Or in index-less matrix notations

$$L\;\in O(3,1)\quad\textbf{if and only if}\quad L^{\top}gL\;=\;g\tag{114}$$

where $L^{\top}$ is the transposed matrix L. By comparison, in Euclidean spaces the orthogonal $O(D)$ matrices R satisfy $R^{\top}R=1$, which may also be written in the form (114) for the Euclidean metric $g=1_{D\times D}$.

I do not have the class-time to go into the group theory of the Lorentz group defined by the matrix condition (114). Instead, let me simply state without proof that it is indeed a group — a matrix product of two Lorentz symmetries is itself a Lorentz symmetry, and so is the matrix inverse of any Lorentz symmetry, — and briefly describe its content. The Lorentz group $O(3,1)$ includes both continuous and discrete symmetries. The continuous subgroup $SO^{+}(3,1)$ (called the continuous Lorentz group) comprises:

1. All rotations of the 3D space (any angle, any axis).
2. All Lorentz boosts (any speed $v<c$, any direction).

3. All combinations of Lorentz boosts and space rotations.

The discrete Lorentz symmetries are the reversal of space P (the parity), the reversal of time T, and their combination PT. And of course, any of these discrete symmetries can be combined with a continuous Lorentz symmetry — a boost, a rotation, or a combination of both. For example, reflections off moving mirrors or a time reversals in moving frames are members of the $O(3,1)$ Lorentz group.

Whenever the time order of events is important but the left/right distinction is not, the relevant symmetry group is the *orthochronous Lorentz group* $O^+(3,1)$. It comprises the continuous Lorentz symmetries, the space reflection P, and all combinations thereof — but not the time reversal T. Under orthochronous Lorentz symmetries,

$$\mathrm{sign}(x'^0) \;=\; \mathrm{sign}(x^0) \quad \text{provided} \quad x^\mu x_\mu \;\geq\; 0. \tag{115}$$

In 3D, the scalars, the vectors, and the tensors are defined by their transformation properties under the rotation symmetries. Likewise, in 4D, the 4–scalars, the 4–vectors, and the 4–tensors are defined by their transformations under the continuous Lorentz symmetries (rotations, boosts, and their combinations).

- A genuine 4–scalar must be invariant under all the $SO^+(3,1)$ symmetries.
- The components of a genuine 4–vector must transform like the spacetime coordinates (ct,x,y,z),

$$A'^\mu \;=\; L^\mu_{\ \nu}A^\nu \quad \text{for the same } L^\mu_\nu \text{ as} \quad x'^\mu \;=\; L^\mu_{\ \nu}x^\nu. \tag{116}$$

- Every index of a genuine 4–tensor must transform like the index of a 4–vectors. For example, the components of a two-index 4–tensor $F^{\mu\nu}$ must transform according to

$$F'^{\mu\nu} \;=\; L^\mu_{\ \alpha}L^\nu_{\ \beta}F^{\alpha\beta}. \tag{117}$$

PROPER VELOCITY

The *proper velocity* is the spacetime displacement of some body per unit of its proper time τ,

$$u^\mu = \frac{dx^\mu}{d\tau}. \tag{118}$$

Under the Lorentz transforms, the dx^μ transforms like a proper 4–vector while $d\tau$ is Lorentz invariant, *i.e.* a Lorentz scalar, so their ratio (118) must be a proper 4–vector transforming as

$$u'^\mu = L^\mu_{\ \nu} u^\nu. \tag{119}$$

In components,

$$u^0 = \frac{d(ct)}{d\tau} = c\frac{dt}{d\tau} = c\gamma \tag{120}$$

while

$$\mathbf{u} = \frac{d\mathbf{r}}{d\tau} = \frac{d\mathbf{r}}{dt}\frac{dt}{d\tau} = \mathbf{v}\gamma, \tag{121}$$

thus

$$u^\mu = \bigl(\gamma c; \gamma v^x, \gamma v^y, \gamma v^z\bigr). \tag{122}$$

Consequently,

$$(u\cdot u) = u^\mu u_\mu = (u^0)^2 - \mathbf{u}^2 = \gamma^2(c^2 - \mathbf{v}^2) = c^2 \tag{123}$$

for any particle 3–velocity $\mathbf{v}$. The same result can be obtained in a manifest 4–vector form using the definition of the proper time,

$$(cd\tau)^2 = dI^2 = (dx\cdot dx), \tag{124}$$

hence

$$(u\cdot u) = \frac{(dx\cdot dx)}{(d\tau)^2} = c^2. \tag{125}$$

Yet another way to see that $(u\cdot u) = c^2$ is to note that this is true in the rest frame of the particle and hence must be true in any other frame since $(u\cdot u)$ is a Lorentz scalar.

The Lorentz transform eq. (119) is a good starting point for the relativistic velocity addition formula for the non-parallel velocities. Indeed, consider 2 inertial frames K_1 and K_2 moving relative to each other at velocity $\mathbf{v}_{12}$. (That is, K_1 moves at velocity $\mathbf{v}_{12}$ relative to the K_2.) Then, the Lorentz transform from K_1 to K_2 of any Lorentz 4–vector X^μ acts as

$$\begin{aligned} X_2^0 &= \gamma_{12}X_1^0 \;+\; \gamma_{12}\boldsymbol{\beta}_{12}\cdot\mathbf{X}_1, \\ \mathbf{X}_2 &= \gamma_{12}\boldsymbol{\beta}_{12}X_1^0 \;+\; \mathbf{X}_1 \;+\; \frac{\gamma_{12}-1}{\beta_{12}^2}(\boldsymbol{\beta}_{12}\cdot\mathbf{X}_1)\boldsymbol{\beta}_{12}\,. \end{aligned} \tag{126}$$

In particular, the proper velocity

$$u_1^\mu \;=\; (\gamma_1 c;\gamma_1\boldsymbol{\beta}_1 c) \tag{127}$$

of some body in the K_1 frame transforms to

$$\begin{aligned} u_2^0 &= \gamma_{12}u_1^0 \;+\; \gamma_{12}\boldsymbol{\beta}_{12}\cdot\mathbf{u}_1 \\ &= \gamma_{12}\gamma_1 c\Big(1 \;+\; \boldsymbol{\beta}_{12}\cdot\boldsymbol{\beta}_1\Big), \\ \mathbf{u}_2 &= \gamma_{12}\boldsymbol{\beta}_{12}u_1^0 \;+\; \mathbf{u}_1 \;+\; \frac{\gamma_{12}-1}{\beta_{12}^2}(\boldsymbol{\beta}_{12}\cdot\mathbf{u}_1)\boldsymbol{\beta}_{12} \\ &= \gamma_1 c\left(\gamma_{12}\boldsymbol{\beta}_{12} + \boldsymbol{\beta}_1 + \frac{\gamma_{12}-1}{\beta_{12}^2}(\boldsymbol{\beta}_{12}\cdot\boldsymbol{\beta}_1)\boldsymbol{\beta}_{12}\right). \end{aligned} \tag{128}$$

Consequently, the 3–velocity $\mathbf{v}_2$ of the same body in the K_2 frame — or rather $\boldsymbol{\beta}_2 = \mathbf{v}_2/c$ — is

$$\boldsymbol{\beta}_2 \;=\; \frac{\mathbf{u}_2}{\gamma_2 = u_2^0} \;=\; \frac{1}{1+\boldsymbol{\beta}_{12}\cdot\boldsymbol{\beta}_1}\left(\boldsymbol{\beta}_{12} + \frac{\boldsymbol{\beta}_1}{\gamma_{12}} + \frac{\gamma_{12}-1}{\gamma_{12}\beta_{12}^2}(\boldsymbol{\beta}_{12}\cdot\boldsymbol{\beta}_1)\boldsymbol{\beta}_{12}\right). \tag{129}$$

In particular, for the relative velocity $\mathbf{v}_{12}$ of the two frames being in the x direction, this formula becomes

$$\begin{aligned} \beta_2^x &= \frac{1}{1+\beta_{12}\beta_1^x}\left(\beta_{12} + \frac{\beta_1^x}{\gamma_{12}} + \frac{\gamma_{12}-1}{\gamma_{12}}\beta_1^x\right) \;=\; \frac{\beta_{12}+\beta_1^x}{1+\beta_{12}\beta_1^x}\,, \\ \beta_2^y &= \frac{1}{1+\beta_{12}\beta_1^x}\,\frac{\beta_1^y}{\gamma_{12}}\,, \\ \beta_2^z &= \frac{1}{1+\beta_{12}\beta_1^x}\,\frac{\beta_1^z}{\gamma_{12}}\,, \end{aligned} \tag{130}$$

or equivalently

$$\mathbf{v}_2 \;=\; \frac{c^2}{c^2+v_{12}v_1^x}\left((v_{12}+v_1^x)\hat{\mathbf{x}} \;+\; \frac{v_1^y\hat{\mathbf{y}}+v_2^z\hat{\mathbf{z}}}{\gamma_{12}}\right). \tag{131}$$

Finally, for a general direction of the relative velocity $\mathbf{v}_{12}$ of the two frames, let us split the object's velocity $\mathbf{v}_1$ into components $\mathbf{v}_1^{\|}$ that's $\|\ \mathbf{v}_{12}$ and $\mathbf{v}_1^{\perp}$ that's $\perp \mathbf{v}_{12}$. In terms of these components,

$$\mathbf{v}_2 \;=\; \frac{c^2}{c^2\,+\,\mathbf{v}_{12}\cdot\mathbf{v}_1}\left(\mathbf{v}_{12}\;+\;\mathbf{v}_1^{\|}\;+\;\frac{\mathbf{v}_1^{\perp}}{\gamma_{12}}\right). \tag{132}$$

For an example of using all these formulae, see problem 6 of your homework set#11.

More Examples of 4–Vectors

Consider a plane wave $\psi(t,\mathbf{r})=\psi_0\exp(i\mathbf{k}\cdot\mathbf{r}-i\omega t)$. The the frequency ω and the wave vector $\mathbf{k}$ of this wave combine into a 4–vector

$$k^\mu \;=\; \left(\frac{\omega}{c},k_x,k_y,k_z\right). \tag{133}$$

In the 4-vector language, the phase of the plane wave becomes the Lorentzian dot product

$$\text{phase} \;=\; \mathbf{k}\cdot\mathbf{r}\;-\;\omega\times t\;=\;\mathbf{k}\cdot\mathbf{r}\;-\;\left(k^0=\frac{\omega}{c}\right)\times\left(x^0=ct\right)\;=\;-(k\cdot x)\;=\;-k_\mu x^\mu, \tag{134}$$

hence

$$\psi(x)\;=\;\psi_0\exp(-ik_\mu x^\mu). \tag{135}$$

Physically, the phase of a wave must be invariant under any Lorentz transform of the space-time coordinates, and the only way to achieve this invariance for all x is to make the k^μ transform like a Lorentz 4–vector,

$$k'^\mu \;=\; L^\mu{}_\nu k^\nu \quad\text{and}\quad x'^\mu \;=\; L^\mu{}_\nu x^\nu \quad\Longrightarrow\quad \text{phase}\;=\;-k_\mu x^\mu\;=\;-k'_\mu x'^\mu. \tag{136}$$

The *relativistic Doppler effect* follows from this Lorentz transformation formula for the wave 4–vector k^μ: If in one frame of reference a wave has frequency ω and wave vector $\mathbf{k}$, then in another frame moving at velocity $\mathbf{v}$ relative to the first frame, the wave's frequency is

$$\omega' = ck'^0 = c\gamma(k^0 - \vec{\beta}\cdot\mathbf{k}) = \gamma(\omega - \mathbf{v}\cdot\mathbf{k}). \tag{137}$$

In particular, for a light wave in vacuum $\mathbf{k} = (\omega/c)\mathbf{n}$ where $\mathbf{n}$ is a unit vector in the direction of the wave, hence

$$\omega' = \gamma(1 - \vec{\beta}\cdot\mathbf{n}) \times \omega. \tag{138}$$

Note: non-relativistically, there are two different formulae for the Doppler effect, one for the frequency change from a moving source to the medium through which the wave propagates, and the other for the change from the medium to the moving detector,

$$\omega_{\text{medium}} = \left(1 + \frac{\mathbf{v}_{\text{source}}\cdot\mathbf{n}}{u_{\text{wave}}}\right) \times \omega_{\text{source}}\,, \quad \omega_{\text{detector}} = \left(1 + \frac{\mathbf{v}_{\text{detector}}\cdot\mathbf{n}}{u_{\text{wave}}}\right)^{-1} \times \omega_{\text{medium}}\,. \tag{139}$$

But for a light wave, the “medium” does not matter, and all we need is the relative velocity $\mathbf{v}$ between the source and the detector, and we may use eq. (138) to go directly from the source frame to the detector frame. In particularly, for the source directly approaching or directly receding from the detector,

$$\frac{\omega_{\text{detector}}}{\omega_{\text{source}}} = \gamma(1 \mp \beta) = \frac{1 \mp \beta}{\sqrt{1-\beta^2}} = \sqrt{\frac{1 \mp \beta}{1 \pm \beta}}\,, \tag{140}$$

while for the relative motion ⊥ to the wave direction

$$\frac{\omega_{\text{detector}}}{\omega_{\text{source}}} = \gamma\,. \tag{141}$$

Another important example of a 4–vector is the derivative vector

$$\partial_\mu \;=\; \left(\frac{\partial}{\partial x^\mu}\right)_{\text{other}\,x^\nu} \;=\; \left(\frac{1}{c}\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right). \tag{142}$$

Let's prove that these derivatives indeed transform like a components of a Lorentz 4–vector. First, for any invertible linear transform $x'^\mu = L^\mu_{\ \nu} x^\nu$, the derivatives transform in the contragradient fashion,

$$\left(\frac{\partial}{\partial x'^\mu}\right)_{\text{other}\,x'^\alpha} \;=\; \left(\left(L^\top\right)^{-1}\right)_\mu^{\ \nu} \left(\frac{\partial}{\partial x^\nu}\right)_{\text{other}\,x^\beta}. \tag{143}$$

Second, if the transform matrix L belongs to the Lorentz symmetry group, then

$$L^\top g L \;=\; g \quad\Longrightarrow\quad gL \;=\; \left(L^\top\right)^{-1} g \quad\Longrightarrow\quad \left(L^\top\right)^{-1} \;=\; gLg^{-1}, \tag{144}$$

or in explicit index notations,

$$\left(\left(L^\top\right)^{-1}\right)_\mu^{\ \nu} \;=\; g_{\mu\alpha} L^\alpha_{\ \beta} g^{\beta\nu} \;=\; L_\mu^{\ \nu}, \tag{145}$$

where the second equality is simply raising and lowering of indices. Consequently, under any Lorentz transforms of the spacetime coordinates, the derivatives transform as

$$\partial'_\mu \;=\; L_\mu^{\ \nu}\partial_\nu \quad\Longrightarrow\quad \partial'^\mu \;=\; L^\mu_{\ \nu}\partial^\nu. \tag{146}$$

In other words, the 4 derivative operators ∂^μ indeed comprise a genuine Lorentz vector.

Corollary: The D'Alembert operator

$$\Box \;\stackrel{\text{def}}{=}\; \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2} \;=\; \partial_\mu \partial^\mu \tag{147}$$

is invariant under all Lorentz transforms. Consequently, a solution of the wave equation $\Box\psi(x) = 0$ in one frame of reference would also be a solution in any other reference frame.

Let me conclude these notes with a few more examples of Lorentz vectors and tensors which I shall discuss in detail in the following lectures.

- The energy and the momentum of a particle form a 4–vector

$$p^\mu = \left(\frac{E}{c}, p_x, p_y, p_z\right). \tag{148}$$

For a particle of rest mass m_0 moving at 4–velocity u^μ,

$$p^\mu = m_0 u^\mu \quad\Longrightarrow\quad E = \gamma m_0 c^2 \quad\text{and}\quad \mathbf{p} = \gamma m_0 \mathbf{v}. \tag{149}$$

- The electric charge density and the current density form a 4–vector

$$J^\mu = (c\rho, J_x, J_y, J_z). \tag{150}$$

- The scalar potential V and the vector potential $\mathbf{A}$ form a 4–vector. In the Gauss units

$$A^\mu = (V, A_x, A_y, A_z) \tag{151}$$

while in the MKSA units

$$A^\mu = \left(\tfrac{V}{c}, A_x, A_y, A_z\right). \tag{152}$$

- The electric field $\mathbf{E}$ and the magnetic field $\mathbf{B}$ form an antisymmetric Lorentz tensor $F^{\mu\nu} = -F^{\nu\mu}$. In the Gauss units

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ +E_x & 0 & -B_z & +B_y \\ +E_y & +B_z & 0 & -B_x \\ +E_z & -B_y & +B_x & 0 \end{pmatrix} \tag{153}$$

while in the MKSA units

$$F^{\mu\nu} = \begin{pmatrix} 0 & -\frac{1}{c}E_x & -\frac{1}{c}E_y & -\frac{1}{c}E_z \\ +\frac{1}{c}E_x & 0 & -B_z & +B_y \\ +\frac{1}{c}E_y & +B_z & 0 & -B_x \\ +\frac{1}{c}E_z & -B_y & +B_x & 0 \end{pmatrix}. \tag{154}$$

- Finally, the EM energy density U, the Poynting vector $\mathbf{S}$, and the Maxwell stress tensor

T_{ij} combine into a symmetric Lorentz tensor $T^{\mu\nu} = +T^{\nu\mu}$. In the Gauss units

$$T^{\mu\nu} \ = \begin{pmatrix} U & S_x & S_y & S_z \\ S_x & -T_{xx} & -T_{xy} & -T_{xz} \\ S_y & -T_{yx} & -T_{yy} & -T_{yz} \\ S_z & -T_{zx} & -T_{zy} & -T_{zz} \end{pmatrix} \tag{155}$$

while in the MKSA units

$$T^{\mu\nu} \ = \begin{pmatrix} U & \frac{1}{c}S_x & \frac{1}{c}S_y & \frac{1}{c}S_z \\ \frac{1}{c}S_x & -T_{xx} & -T_{xy} & -T_{xz} \\ \frac{1}{c}S_y & -T_{yx} & -T_{yy} & -T_{yz} \\ \frac{1}{c}S_z & -T_{zx} & -T_{zy} & -T_{zz} \end{pmatrix} \tag{156}$$