SPACETIME GEOMETRY

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, the relation between the times in some frames K and K' cannot be a relations between the two times alone, $t' = f(t)$, but must also involve the space coordinates, thus $t' = f(t, x, y, z)$: Otherwise, the symmetry between the two frames K and K' would force $t' = t$. 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

$$x' = \gamma(x - ut), \quad y' = y, \quad z' = z, \quad t' = \gamma\left(t - \frac{u}{c^2}x\right) \tag{1}$$

in one direction, and

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

in the opposite direction. For both directions,

$$\gamma = \frac{1}{\sqrt{1 - (u/c)^2}}\,. \tag{3}$$

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 $x_{\text{rest}} \equiv 0$, hence in the lab frame where the body is moving at velocity v,

$$t_{\text{lab}} = \gamma_v \left(t_{\text{rest}} + \frac{v}{c^2} \times x_{\text{rest}} \right) = \gamma_v \times t_{\text{rest}} = \frac{t_{\text{rest}}}{\sqrt{1 - (v/c)^2}} . \tag{4}$$

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 $x_{\text{lab}} = v \times t_{\text{lab}}$, hence

$$t_{\text{rest}} = \gamma_v \left(t_{\text{lab}} - \frac{v}{c^2} \times (x_{\text{lab}} = vt_{\text{lab}}) \right) = \gamma_v \times \left(1 - \frac{v^2}{c^2} \right) \times t_{\text{lab}} = \sqrt{1 - (v/c)^2} \times t_{\text{lab}} . \tag{5}$$

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,

$$x^{(1)}_{\text{lab}} = v \times t_{\text{lab}} , \quad x^{(2)}_{\text{lab}} = v \times t_{\text{lab}} + L_{\text{lab}} , \tag{6}$$

so that

$$x^{(2)}_{\text{lab}}(t_{\text{lab}}) - x^{(1)}_{\text{lab}}(t_{\text{lab}}) = L_{\text{lab}} \ \forall t_{\text{lab}} . \tag{7}$$

In the rest frame, this translates to

$$x^{(1)}_{\text{rest}} = \gamma_v (x^{(1)}_{\text{lab}} - vt_{\text{lab}}) = 0, \quad x^{(2)}_{\text{rest}} = \gamma_v (x^{(2)}_{\text{lab}} - vt_{\text{lab}}) = \gamma_v \times L_{\text{lab}} \tag{8}$$

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

$$L_{\text{lab}} = \frac{L_{\text{rest}}}{\gamma_v} = \sqrt{1 - (v/c)^2} \times L_{\text{rest}} < L_{\text{rest}} . \tag{9}$$

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

$$x^{(1)}_{\rm lab} \;=\; \gamma_v v\times t^{(1)}_{\rm rest}\,,\quad x^{(2)}_{\rm lab} \;=\; \gamma_v\times L_{\rm rest} \;+\; \gamma_v v\times t^{(2)}_{\rm rest}\,, \tag{10}$$

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

$$t^{(1)}_{\rm lab} \;=\; \gamma_v v\times t^{(1)}_{\rm rest}\quad\text{and}\quad t^{(2)}_{\rm lab} \;=\; \gamma_v v\times t^{(2)}_{\rm rest} \;+\; \frac{\gamma_v v}{c^2}\times L_{\rm rest} \tag{11}$$

are equal to each other. Thus,

$$t^{(2)}_{\rm rest} \;-\; t^{(1)}_{\rm rest} \;=\; -\frac{v}{c^2}\times L_{\rm rest} \tag{12}$$

and hence

$$\begin{aligned} L_{\rm lab} \;&=\; x^{(2)}_{\rm lab} \;-\; x^{(1)}_{\rm lab} \;=\; \gamma_v\times L_{\rm rest} \;+\; \gamma_v v\big(t^{(2)}_{\rm rest}-t^{(1)}_{\rm rest}\big)\\ &=\; \gamma_v\times\left(1\;-\;\frac{v^2}{c^2}\right)\times L_{\rm rest} \;=\; \sqrt{1-(v/c)^2}\times L_{\rm rest} \;<\; L_{\rm rest}\,. \end{aligned} \tag{13}$$

Next consider the relativistic velocity addition. 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 these notes, I'll write this formula for $\mathbf{u}$ and $\mathbf{v}'$ being in the same direction x (or in the opposite directions for $v'<0$); the more general case is a part of your next homework.

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

$$\begin{aligned} x \;&=\; \gamma_u(x'+u\times t') \;=\; \gamma_u\times(v'+u)\times t',\\ t \;&=\; \gamma_u\left(t'\;+\;\frac{u}{c^2}\times x'\right) \;=\; \gamma_u\times\left(1\;+\;\frac{uv'}{c^2}\right)\times t', \end{aligned} \tag{14}$$

and therefore

$$x \;=\; \frac{u+v'}{1\;+\;\frac{uv'}{c^2}}\times t. \tag{15}$$

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

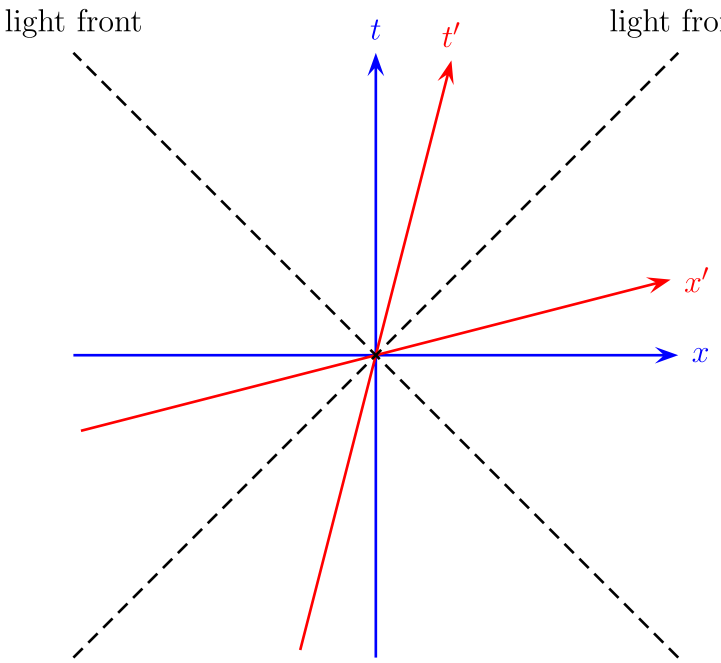
$$v \;=\; \frac{u+v'}{1+(uv'/c^2)}\,. \tag{16}$$

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 ta velocity $v' = \pm c$ relative to the frame K', it's velocity v relative to the K frame is also

$$v = \frac{u \pm c}{1 \pm (u/c)} = \pm c. \tag{17}$$

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

$$x' = \cos\phi \times x - \sin\phi \times y, \quad y' = \cos\phi \times y + \sin\phi \times x \tag{18}$$

preserves the radius^2,

$$r^2 \;=\; x^2 \;+\; y^2 \;=\; x'^2 \;+\; y'^2, \tag{19}$$

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

$$I^2 \;=\; c^2t^2 \;-\; x^2 \;=\; c^2t'^2 \;-\; x'^2. \tag{20}$$

Indeed,

$$\begin{aligned}
I'^2 \;&=\; c^2t'^2 \;-\; x'^2 \\
&=\; c^2\gamma^2\left(t - \frac{v}{c^2}\times x\right)^2 \;-\; \gamma^2(x-vt)^2 \\
&=\; \gamma^2\left[\left(c^2\times t^2 \;-\; 2t\times vx \;+\; \frac{v^2}{c^2}\times x^2\right) \;-\; \left(x^2 \;-\; 2x\times vt \;+\; v^2t^2\right)\right] \\
&=\; \gamma^2(c^2-v^2)\times t^2 \;+\; 0\times xt \;+\; \gamma^2\left(\frac{v^2}{c^2} - 1\right)\times x^2 \\
&=\; \gamma^2\left(1 - \frac{v^2}{c^2}\right)\times(c^2t^2-x^2) \\
&=\; 1\times(c^2t^2-x^2) \;=\; I^2.
\end{aligned} \tag{21}$$

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

$$\begin{aligned}
\mathbf{x}' \;&=\; \mathbf{x}_\perp \;+\; \gamma_u\mathbf{x}_\parallel \;-\; \gamma_u\mathbf{u}\,t \\
&=\; \mathbf{x} \;+\; \frac{\gamma_u - 1}{u^2}\,(\mathbf{x}\cdot\mathbf{u})\mathbf{u} \;-\; \gamma_u\mathbf{u}\,t\,, \\
t' \;&=\; \gamma_u\left(t - \frac{\mathbf{u}\cdot\mathbf{x}}{c^2}\right),
\end{aligned} \tag{22}$$

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

$$I^2 \;=\; (ct)^2 \;-\; \mathbf{x}^2 \;=\; (ct)^2 \;-\; x^2 \;-\; y^2 \;-\; z^2. \tag{23}$$

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

between two such events is defined as

$$I_{12}^2 \;=\; c^2(t_2-t_1)^2 \;-\; (\mathbf{x}_2-\mathbf{x}_2)^2 \;=\; c^2(t_2-t_1)^2 \;-(x_2-x_1)^2 \;-\; (y_2-y_1)^2 \;-\; (z_2-z_1)^2, \tag{24}$$

and just like in eq. (23), 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 $c|t_2 - t_1| > |\mathbf{x}_2 - \mathbf{x}_1|$;
- spacelike intervals $I_{12}^2 < 0$, with $c|t_2 - t_1| < |\mathbf{x}_2 - \mathbf{x}_1|$;
- lightlike intervals $I_{12}^2 = 0$, with $c|t_2 - t_1| = |\mathbf{x}_2 - \mathbf{x}_1|$.

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 $\Delta\mathbf{x} = \mathbf{x}_2 - \mathbf{x}_2$ in some frame of reference K. Then in another frame K' moving at velocity $\mathbf{u}$ relative to K,

$$\Delta t' \;=\; \gamma_u\left(\Delta t \;-\; \frac{\mathbf{u}}{c^2}\cdot\Delta\mathbf{x}\right). \tag{25}$$

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

$$\left|\frac{\mathbf{u}}{c^2}\cdot\Delta\mathbf{x}\right| \;\leq\; \frac{u}{c}\times\frac{|\Delta\mathbf{x}|}{c} \;<\; 1\times|\Delta t|, \tag{26}$$

hence

$$\mathrm{sign}(\Delta t') \;=\; \mathrm{sign}\left(\Delta t \;-\; \frac{\mathbf{u}}{c^2}\cdot\Delta\mathbf{x}\right) \;=\; \mathrm{sign}(\Delta t). \tag{27}$$

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 $(\Delta t, \Delta\mathbf{x})$ with $c||\Delta t| < |\Delta\mathbf{x}|$ in some frame K. Let

$$u_0 \;=\; \frac{c^2|\Delta t|}{|\Delta\mathbf{x}|}\,. \tag{28}$$

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 $\operatorname{sign}(\Delta t)\Delta\mathbf{x}$, then

$$\frac{\mathbf{u}\cdot\mathbf{x}}{c^2} \;=\; \operatorname{sign}(\Delta t)\times\frac{u\,|\Delta\mathbf{x}|}{c^2} \;=\; \operatorname{sign}(\Delta t)\times\frac{u}{u_0}\times\left(\frac{u_0|\Delta\mathbf{x}}{c^2} \;=\; |\Delta t|\right) \;=\; \frac{u}{u_0}\times\Delta t. \tag{29}$$

Consequently, in the K' frame the time difference

$$\Delta t' \;=\; \gamma_u\left(\Delta t \;-\; \frac{\mathbf{u}}{c^2}\cdot\Delta\mathbf{x}\right) \tag{30}$$

becomes

$$\Delta t' \;=\; \gamma_u\left(1 \;-\; \frac{u}{u_0}\right)\times\Delta t\,, \tag{31}$$

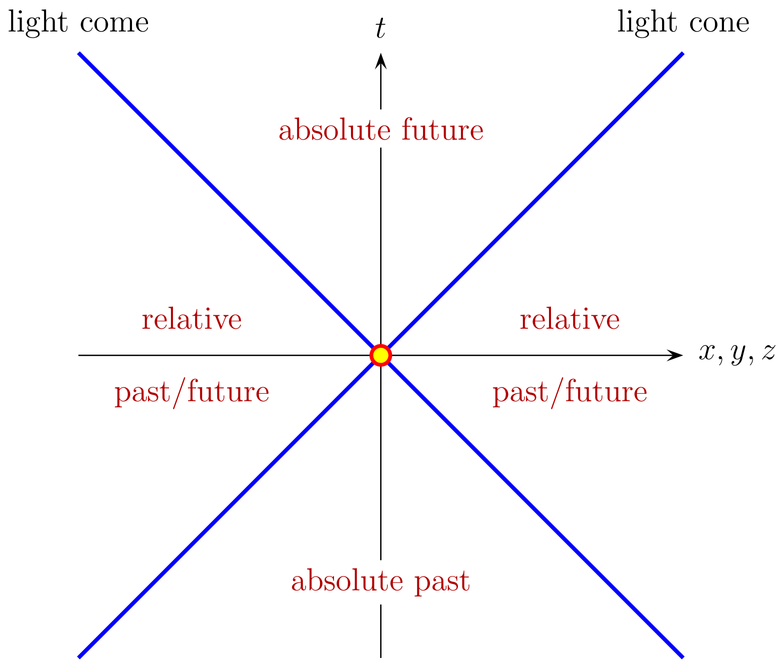
and since $u > u_0$, this $\Delta t'$ has the opposite sign from the Δt! Thus, if in the original frame K the event (2) happens later than the event (1), $\Delta t > 0$, then in the K' frame $\Delta t' < 0$ and the event (2) happens before the event (1). Likewise, if in the K frame (2) happens before (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 $(t_0, \mathbf{x}_0)$ form a double cone in spacetime

$$t \;=\; t_0 \;\pm\; \frac{|\mathbf{x}-\mathbf{x}_0|}{c} \tag{32}$$

called the *light cone*. Physically, the light cone is spanned by all the light rays beginning or ending at point $\mathbf{x}_0$ 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

$$t \;>\; t_0 \;+\; \frac{|\mathbf{x}-\mathbf{x}_0|}{c} \tag{33}$$

comprises events which are later than $(t_0, \mathbf{x}_0)$ in all frames of reference.

- Likewise, the absolute past region

$$t \;<\; t_0 \;-\; \frac{|\mathbf{x}-\mathbf{x}_0|}{c} \tag{34}$$

comprises events which are earlier than $(t_0, \mathbf{x}_0)$ in all frames of reference.

- Finally, the relative past/future region

$$t_0 \;-\; \frac{|\mathbf{x}-\mathbf{x}_0|}{c} \;<\; t \;<\; t_0 \;+\; \frac{|\mathbf{x}-\mathbf{x}_0|}{c} \tag{35}$$

which are can be earlier or later than $(t_0, \mathbf{x}_0)$ 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.

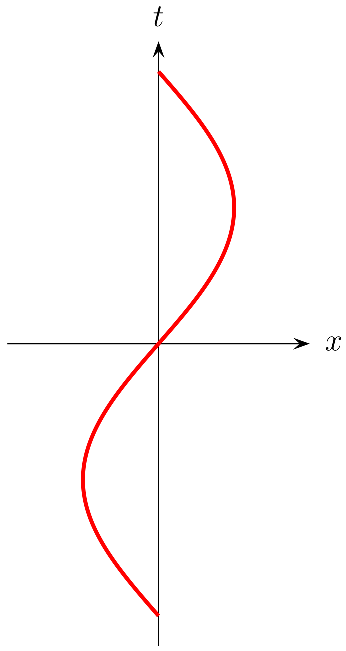
Causality means that an even in the past can cause or influence an event in the future but not the other way around: the future cannot influence the past. Relativistically, causality has to work in all reference frames, so if the time order of two events is frame-dependent, then neither even can cause influence the other. Thus, *relativistic causality* means that *an event may cause or influence other events only in its absolute future.* Likewise, *and 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:

$$\text{for } \mathbf{x}_2 - \mathbf{x}_1 \;=\; \mathbf{v}(t_2 - t_1) \text{ and } |\mathbf{v}| > c, \quad I_{12}^2 < 0, \tag{36}$$

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.*

Worldlines: In spacetime terms, a moving particle spans a continuous family of events $(t, \mathbf{x}(t))$ 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 $d\mathbf{x}/dt$ is always slow than the speed of light, so the infinitesimal intervals $(dt, d\mathbf{x})$ along the worldline are always timelike,

$$dI^2 \;=\; c^2 dt^2 \;-\; \mathbf{x}^2 \;=\; (c^2 - \mathbf{v}^2)dt^2 \;>\; 0. \tag{37}$$

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 $c^2 dt'^2$, 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

$$d\tau \;=\; \frac{dI}{c} \;=\; \sqrt{1 - (v/c)^2} \times dt. \tag{38}$$

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

$$\tau \;=\; \int dt\, \sqrt{1 \,-\, v^2(t)/c^2}\,, \tag{39}$$

or in worldline terms,

$$\tau \;=\; \int\limits_{\text{worldline}} \sqrt{(dt)^2\;-\;\frac{1}{c^2}(d\mathbf{x})^2}\,. \tag{40}$$

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. Even with today's non-relativistic technology, the clocks on GPS satellites run according the proper time, and the clock rate has to be corrected to keep those clocks synchronized with the ground-based clocks. Microscopically, the atomic transitions have definite frequencies WRT to the proper time of a moving atom rather than the lab-frame time. Likewise, the lifetimes of unstable particles are 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, time enough to travel 30 km from the stratosphere to the ground.

4–VECTORS

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

$$x^\mu \;=\; (x^0,x^1,x^2,x^3) \;=\; (ct,x,y,z). \tag{41}$$

Other combinations of a 3–scalar and a 3–vector which transform similarly under Lorentz transformations also form 4–vector; we shall see quite a few examples in the remaining 2 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,

$$A^\mu B_\mu \;=\; \sum_{\mu=0,1,2,3} A^\mu B_\mu\,. \tag{42}$$

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 where 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

$$\begin{pmatrix} x'^0 \\ x'^1 \\ x'^2 \\ x'^3 \end{pmatrix} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \tag{43}$$

where

$$\beta \;=\; \frac{v}{c} \quad\text{and}\quad \gamma \;=\; \frac{1}{\sqrt{1-\beta^2}}\,. \tag{44}$$

or in terms of explicit matrix elements,

$$L^0{}_0 \;=\; L^1{}_1 \;=\; \gamma, \quad L^0{}_1 \;=\; L^1{}_0 - = \;-\beta\gamma, \quad L^2{}_2 \;=\; L^3{}_3 \;=\; 1, \quad \text{other } L^\mu{}_\nu \;=\; 0. \tag{45}$$

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 vectors, or more generally a dot product of any two vectors,

$$\mathbf{a}\cdot\mathbf{b} \;=\; a_1b_1 \;+\; a_2b_2 \;+a_3b_3 \;=\; a'_1b'_1 \;+\; a'_2b'_2 \;+\; a'_3b'_3\,. \tag{46}$$

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

$$(A)^2 \;=\; (A^0)^2 - \;(A^1)^2 \;-\; (A^2)^2 \;-\; (A^3)^2 = \;(A'^0)^2 \;-\; (A'^1)^2 \;-\; (A'^2)^2 \;-\; (A'^3)^2. \tag{47}$$

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

$$(A\!\cdot\! B) \;=\; A^0B^0 \;-\; A^1B^1 \;-\; A^2B^2 \;-\; A^3B^3 \;=\; A'^0B'^0 \;-\; A'^1B'^1 \;-\; A'^2B'^2 \;-\; A'^3B'^3. \tag{48}$$

In index notations,

$$(A\cdot B) \;=\; A^{\mu}g_{\mu\nu}B^{\nu} \tag{49}$$

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

$$g_{00} \;=\; +1, \quad g_{11} \;=\; g_{22} \;=\; g_{33} \;=\; -1, \quad \text{and for } \mu\neq\nu \quad g_{\mu\nu} \;=\; 0, \tag{50}$$

or in matrix form,

$$g_{\mu\nu} \;=\; \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{51}$$

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

As a matrix (51), 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

$$g^{\mu\nu} = \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} . \tag{52}$$

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

$$A_{\mu} = g_{\mu\nu} A^{\nu}, \quad A^{\lambda} = g^{\lambda\mu} A_{\mu} , \tag{53}$$

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,

$$A_0 = +A^0, \quad \text{but} \quad A_1 = -A^1, \quad A_2 = -A^2, \quad A_3 = -A^3, \tag{54}$$

for example,

$$x^{\mu} = (+ct, +x, +y, +z) \quad \text{but} \quad x_{\mu} = (+ct, -x, -y, -z). \tag{55}$$

Sign convention: 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

$$A^{\mu} = (A^0, A_x, A_y, A_z) \quad \text{but} \quad A_{\mu} = (A^0, -A_x, -A_y, -A_z). \tag{56}$$

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

$$\partial_\mu = \left(\frac{1}{c}\frac{\partial}{\partial t},\frac{\partial}{\partial x},\frac{\partial}{\partial y},\frac{\partial}{\partial z}\right) \quad \text{while} \quad \partial^\mu = \left(\frac{1}{c}\frac{\partial}{\partial t},-\frac{\partial}{\partial x},-\frac{\partial}{\partial y},-\frac{\partial}{\partial z}\right) \tag{57}$$

so that

$$\partial_\mu x^\nu = +\delta^\nu_\mu = \begin{cases} +1 & \text{for } \mu=\nu, \\ 0 & \text{for } \mu\neq\nu. \end{cases} \tag{58}$$

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

$$\partial^\mu x^\nu = g^{\mu\nu}, \quad \partial_\mu x_\nu = g_{\mu\nu}\,, \quad \text{and} \quad \partial^\mu x_\nu = \delta^\mu_\nu\,. \tag{59}$$

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

$$(A\cdot B) = A^\mu g_{\mu\nu} B^\nu = A^\mu B_\mu = A_\nu B^\nu\,. \tag{60}$$

(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 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 (49). In index notations, this means

$$(A'\cdot B') = A'^\mu g_{\mu\nu} B'^\nu = (L^\mu{}_\alpha A^\alpha) g_{\mu\nu} (L^\nu{}_\beta B^\beta) = A^\alpha (g_{\mu\nu} L^\mu{}_\alpha L^\nu{}_\beta) B^\alpha \tag{61}$$

should be equal to

$$(A\cdot B) = A^\alpha g_{\alpha\beta} B^\beta \tag{62}$$

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

$$g_{\mu\nu}L^\mu{}_\alpha L^\nu{}_\beta \;=\; g_{\alpha\beta}\,. \tag{63}$$

Or in indexless matrix notations

$$L\,\in O(3,1)\quad\textbf{if and only if}\quad L^\top gL \;=\; g \tag{64}$$

where $L^\top$ is the transposed matrix L.

I do not gave the class-time to go into the group theory of the Lorentz group defined by the matrix condition (64). 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,

$$\operatorname{sign}(x'^0) \;=\; \operatorname{sign}(x^0)\quad\text{provided}\quad x^\mu x_\mu \;\geq\; 0. \tag{65}$$

In 3D, scalars, vectors, and 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),

$$A'^{\mu} \;=\; L^{\mu}{}_{\nu}A^{\nu} \quad \text{for the same } L^{\mu}{}_{\nu} \text{ as} \quad x'^{\mu} \;=\; L^{\mu}{}_{\nu}x^{\nu}. \tag{66}$$

- 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

$$F'^{\mu\nu} \;=\; L^{\mu}{}_{\alpha}L^{\nu}{}_{\beta}F^{\alpha\beta}. \tag{67}$$

EXAMPLES OF 4–VECTORS

Along a worldline $x^{\mu}(\tau)$ of some particle, the spacetime coordinates x^{μ} comprise a 4–vector, while the proper time τ along the worldline is a 4–scalar. Consequently, the **4–velocity**

$$u^{\mu} \;=\; \frac{dx^{\mu}}{d\tau} \tag{68}$$

is a genuine 4–vector. In components,

$$u^0 \;=\; \gamma c, \quad u^1 \;=\; \gamma v_x, \quad u^2 \;=\; \gamma v_y, \quad u^3 \;=\; \gamma v_z, \tag{69}$$

hence

$$(u\cdot u) \;=\; u^{\mu}u_{\mu} \;=\; (\gamma c)^2 \;-\; (\gamma\mathbf{v})^2 \;=\; \gamma^2(c^2-\mathbf{v}^2) \;=\; c^2. \tag{70}$$

The same result can be obtained in a manifest 4–vector form using the definition of the proper time,

$$(cd\tau)^2 \;=\; dI^2 \;=\; (dx\cdot dx), \tag{71}$$

hence

$$(u\cdot u) \;=\; \frac{(dx\cdot dx)}{(d\tau)^2} \;=\; c^2. \tag{72}$$

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.

Another example of a 4–vector k^μ comprises the frequency ω and the wave vector $\mathbf{k}$ of a plane wave $\psi(t, \mathbf{x}) = \psi_0 \exp(i\mathbf{k} \cdot \mathbf{x} - i\omega t)$,

$$k^\mu \;=\; \left(\frac{\omega}{c}, k_x, k_y, k_z\right). \tag{73}$$

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

$$\text{phase} \;=\; \mathbf{k} \cdot \mathbf{x} \;-\; \omega \times t \;=\; \mathbf{k} \cdot \mathbf{x} \;-\; \left(k^0 = \frac{\omega}{c}\right) \times \left(x^0 = ct\right) \;=\; -(k \cdot x) \;=\; -k_\mu x^\mu, \tag{74}$$

hence

$$\psi(x) \;=\; \psi_0 \exp(-ik_\mu x^\mu). \tag{75}$$

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,

$$k'^\mu \;=\; L^\mu{}_\nu k^\nu \quad \text{and} \quad x'^\mu \;=\; L^\mu{}_\nu x^\nu \quad \Longrightarrow \quad \text{phase} \;=\; -k_\mu x^\mu \;=\; -k'_\mu x'^\mu. \tag{76}$$

The *relativistic Doppler effect* follows form the Lorentz transformation formula for the 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

$$\omega' \;=\; ck'^o \;=\; c\gamma(k^0 \;-\; \vec{\beta} \cdot \mathbf{k}) \;=\; \gamma(\omega \;-\; \mathbf{v} \cdot \mathbf{k}). \tag{77}$$

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

$$\omega' \;=\; \gamma(1 \;-\; \vec{\beta} \cdot \mathbf{n}) \times \omega. \tag{78}$$

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,

$$\omega_{\text{medium}} = \left(1 + \frac{\mathbf{v}_{\text{source}} \cdot \mathbf{n}}{u_{\text{wave}}}\right) \times \omega_{\text{source}}\,, \quad \omega_{\text{detector}} = \left(1 + \frac{\mathbf{v}_{\text{detector}} \cdot \mathbf{n}}{u_{\text{wave}}}\right)^{-1} \times \omega_{\text{medium}}\,. \tag{79}$$

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. (78) to go directly from the source frame to the detector frame. In particularly, for the source directly approaching or directly receding from the detector,

$$\frac{\omega_{\text{detector}}}{\omega_{\text{source}}} = \gamma(1 \mp \beta) = \frac{1 \mp \beta}{\sqrt{1-\beta^2}} = \sqrt{\frac{1 \mp \beta}{1 \pm \beta}}\,, \tag{80}$$

while for the relative motion $\perp$ to the wave direction

$$\frac{\omega_{\text{detector}}}{\omega_{\text{source}}} = \gamma\,. \tag{81}$$

A particularly important example of 4–vector is the derivative vector

$$\partial_\mu = \left(\frac{\partial}{\partial x^\mu}\right)_{\text{other}\,x^\nu} = \left(\frac{1}{c}\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right). \tag{82}$$

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,

$$\left(\frac{\partial}{\partial x'^\mu}\right)_{\text{other}\,x'^\alpha} = \left(\left(L^\top\right)^{-1}\right)_\mu{}^\nu \left(\frac{\partial}{\partial x^\nu}\right)_{\text{other}\,x^\beta}. \tag{83}$$

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

$$L^\top g L = g \implies gl = \left(L^\top\right)^{-1} g \implies \left(L^\top\right)^{-1} = gLg^{-1}, \tag{84}$$

or in explicit index notations,

$$\left(\left(L^\top\right)^{-1}\right)_\mu{}^\nu = g_{\mu\alpha} L^\alpha{}_\beta g^{\beta\nu} = L_\mu{}^\nu, \tag{85}$$

where the second equality is simply raising and lowering of indices. Consequently, under any

Lorentz transforms of the spacetime coordinates, the derivatives transform as

$$\partial'_\mu \;=\; L_\mu{}^\nu \partial_\nu \quad\Longrightarrow\quad \partial'^\mu \;=\; L^\mu{}_\nu \partial^\nu . \tag{86}$$

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

Corollary: The D'Alembert operator

$$\Box \stackrel{\text{def}}{=} \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2} \;=\; \partial_\mu \partial^\mu \tag{87}$$

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.

What about Green's functions of the wave equation? Since the equation itself is Lorentz-invariant, it follows that Lorentz-transforming a Green's function turns it into a Green's function, although it could be the same Green's function we have started from or a different Green's function. Fortunately, the most important Green's function for the Electrodynamics — *the retarded Green's function*

$$G_R(x-y) \;=\; \frac{1}{4\pi c\,|\mathbf{x}-\mathbf{y}|} \times \delta\left(t_x \,-\, t_y \,-\, \frac{|\mathbf{x}-\mathbf{y}|}{c}\right) \tag{88}$$

is invariant under orthochronous Lorentz symmetries.

Proof: Let me first rewrite the retarded Green's function in a manifestly invariant way

$$G_R(x-y) \;=\; \frac{1}{2\pi}\,\delta\big((x-y)^\mu (x-y)_\mu\big) \times \Theta(x^0 - y^0), \tag{89}$$

and then I'll show that this formula is equivalent to (88). The δ–function in eq. (89) is manifestly invariant under all Lorentz symmetries since it's argument $(x-y)^\mu(x-y)_\mu$ is a dot product of a Lorentz vector with itself and hence a Lorentz scalar. Moreover, the δ–function vanishes unless the vector $x^\mu - y^\mu$ is light-like. For such vectors, the orthochronous Lorentz symmetries preserve the sign of the time component $x^0 - y^0$, and this makes the step-function factor $\Theta(x^0 - y^0)$ also invariant. Thus, the entire Green's function (89) is invariant under orthochronous Lorentz symmetries,

$$\forall L \in O^+(3,1),, \; \forall x^\mu, y^\mu : \quad G_R\big(L^\mu_\nu (x-y)^\nu\big) \; = \; G_R\big(x^\mu - y^\mu\big). \tag{90}$$

Now, to see that eq. (89) is equivalent to eq. (88) for the retired Green's function, let's rewrite it in 3D terms. Let $x^\mu = (ct_x, \mathbf{x})$ and $y^\mu = (ct_y, \mathbf{y})$, and let

$$(x-y)^\mu (x-y)_\mu \; = \; c^2 (t_x - t_y)^2 \; - \; |\mathbf{x} - \mathbf{y}|^2 \; = \; f(t_y) \tag{91}$$

where the second equality emphasizes the t_y dependence of this expression, since $G_R(x-y)$ usually appears in the context of an integral over t_y (and then an integral over $\mathbf{y}$, but right now we only care about the t_y). As usual

$$\delta(f(t_y)) \; = \; \sum_i \frac{\delta(t_y - t_i)}{|f'(t_i)|} \tag{92}$$

where the sum is over the points t_i where $f(t_y = t_i) = 0$ and f' is the derivative of f WRT t_y. For the $f(t_y)$ as in eq. (91), there two points where $f(t_y)$ vanishes,

$$t_{1,2} \; = \; t_x \; \mp \; \frac{|\mathbf{x} - \mathbf{y}|}{c}\,, \tag{93}$$

and at these points

$$f(t_i) \; = \; 2c^2 (t_i - t_x) \; = \; \mp 2c |\mathbf{x} - \mathbf{y}|, \tag{94}$$

hence

$$\delta(f(t_y)) \; = \; \frac{1}{2c|\mathbf{x} - \mathbf{y}|} \times \Big(\delta(t_y - t_x + |\mathbf{x} - \mathbf{y}|/c) \; + \; \delta(t_y - t_x - |\mathbf{x} - \mathbf{y}|/c)\Big). \tag{95}$$

In the context of eq. (89), this makes for

$$G_R(x-y) \; = \; \frac{1}{4\pi c\, |\mathbf{x} - \mathbf{y}|} \times \Big(\delta(t_y - t_x + |\mathbf{x} - \mathbf{y}|/c) + \delta(t_y - t_x - |\mathbf{x} - \mathbf{y}|/c)\Big) \times \Theta(t_x - t_y). \tag{96}$$

The step function factor here is 1 for $t_y = t_x - r/c$ but zero for $t_y = t_x + r/c$, so we are left with

$$G_R(x-y) \; = \; \frac{1}{4\pi c\, |\mathbf{x} - \mathbf{y}|} \times \delta(t_y - t_x + |\mathbf{x} - \mathbf{y}|/c), \tag{97}$$

exactly as in eq. (88). *Quod erat demonstrandum.*

BTW, the retarded Green's function in our formulae has an extra factor of $1/c$ compared to what we had earlier in class, *cf.* my notes on Maxwell equations (eq. (70) on page 13). Consequently,

$$\Box_x\, G_R(x-y) \;=\; \frac{1}{c}\times\delta(t_x-t_y)\times\delta^{(3)}(\mathbf{x}-\mathbf{y}) \;=\; \delta\bigl(c(t_x-t_y)=x^0-y^0\bigr)\times\delta^{(3)}(\mathbf{x}-\mathbf{y}) \;=\; \delta^{(4)}(x-y), \tag{98}$$

which is a more appropriate normalization of a relativistic Green's function.

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

$$p^\mu \;=\; \left(\frac{E}{c}\,p_x, p_y, p_z\right). \tag{99}$$

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

$$p^\mu \;=\; m_0 u^\mu \quad\Longrightarrow\quad E \;=\; \gamma m_0 c^2 \quad\text{and}\quad \mathbf{p} \;=\; \gamma m_0 \mathbf{v}. \tag{100}$$

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

$$J^\mu \;=\; (c\rho, J_x, J_y, J_z). \tag{101}$$

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

$$A^\mu \;=\; (\Phi, A_x, A_y, A_z). \tag{102}$$

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

$$F^{\mu\nu} \;=\; \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ +E_x & 0 & -B_z & +B_y \\ +E_y & +B_z & 0 & -B_x \\ +E_z & -B_y & +B_x & 0 \end{pmatrix}. \tag{103}$$

- 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 Gauss units

$$T^{\mu\nu} = \begin{pmatrix} U & S_x & S_y & S_z \\ S_x & T_{xx} & T_{xy} & T_{xz} \\ S_y & T_{yx} & T_{yy} & T_{yz} \\ S_z & T_{zx} & T_{zy} & T_{zz} \end{pmatrix}. \qquad (104)$$