↓Lecture 9 [12.12.23]

4. Formulation on Minkowski Space

In this section we briefly reformulate what we already learned about SPECIAL RELATIVITY in terms of tensor calculus. We use this notation in subsequent chapters to make classical and quantum mechanics relativistic, and reformulate electrodynamics in a form where its Lorentz covariance is manifest. It also allows a smooth transition into GENERAL RELATIVITY.

The formulation of SPECIAL RELATIVITY on a unified, four-dimensional spacetime manifold goes back to Hermann Minkowski, Albert Einstein's former professors of mathematics at ETH. Minkowski writes in the notes of his lecture "Raum und Zeit" delivered 1908 in Cologne [53]:

> *Die Anschauungen über Raum und Zeit, die ich Ihnen entwickeln möchte, sind auf experimentell-physikalischem Boden erwachsen. Darin liegt ihre Stärke. Ihre Tendenz ist eine radikale. Von Stund' an sollen Raum für sich und Zeit für sich völlig zu Schatten herabsinken und nur noch eine Art Union der beiden soll Selbständigkeit bewahren.*

Einstein, a physicist all through, didn't appreciate this mathematical reformulation of his theory at first. According to Sommerfeld, he (Einstein) commented:

> *Seit die Mathematiker über die Relativitätstheorie hergefallen sind, verstehe ich sie selbst nicht mehr.*

Einstein later changed his views and acknowledged that without Minkowski's introduction of spacetime as a four-dimensional manifold, the development of GENERAL RELATIVITY would have been impossible.

For a historical account on the role of Minkowski, and his relationship (or absence thereof) to Einstein, see Ref. [54].

4.1. Minkowski space

1 | Manifold:

$$\boxed{M = \langle \text{Spacetime of events / coincidence classes } \mathcal{E} \rangle \simeq \mathbb{R}^4} \tag{4.1}$$

It is a well founded, but nonetheless empirical assumption that the spacetime manifold of events has the *topology* of $\mathbb{R}^4$. Note that at this point we do not impose restrictions on the *geometry* of spacetime, e.g., whether it is flat or curved; this follows below when we settle on a metric.

2 | Charts:

In SPECIAL RELATIVITY, we restrict the coordinate systems to the ones that correspond to inertial observers / inertial coordinate systems:

$$\boxed{(\mathcal{E}, K) \quad \leftrightarrow \quad \text{Inertial (coordinate) systems } K \in \mathcal{I}} \tag{4.2}$$

The coordinates are the ones obtained by an ↑ *inertial observer*:

$$K:\ \mathcal{E} \ni E \mapsto K(E) := [E]_K = x \tag{4.3}$$

$$\text{with} \quad x^\mu = (x^0, x^1, x^2, x^3)^T = (ct, x, y, z)^T = (ct, \vec{x})^T \tag{4.4}$$

- ¡! Henceforth, *Greek* indices $\mu, \nu, \ldots$ run over $0, 1, 2, 3$ where $\mu = 0$ denotes the time component and $\mu = 1, 2, 3$ denote the spatial components. *Roman* indices $i, j, \ldots$ run only over the spatial components $1, 2, 3$.
- ¡! We multiply the time t with the speed of light to measure times and distances in the same units.
- Since we assumed that our inertial systems cover all of spacetime, the domains on which the coordinate functions are defined are the complete manifold.
- The notation above is very suggestive: You can think of our inertial systems, namely the calibrated latticework of clocks and rods, as physical manifestations of the coordinate map of the corresponding chart. That is, an inertial system is a measurement device, or function, which assigns to every event $E \in \mathcal{E}$ the coordinate tuple $x = K(E) = (ct, \vec{x})_K \in E$.

3 | Transition maps:

i | We worked hard in Section 1.4 to derive and select the correct coordinate transformations between different inertial systems. The most general ones have the form of …

$$\left.\begin{array}{r}\textit{Inhomogenous}\text{ Lorentz transformations}\\ \text{Poincaré transformations}\end{array}\right\}:\quad \bar{x} = \varphi(x) = \Lambda x + a \tag{4.5}$$

with $a \in \mathbb{R}^4$ arbitrary and $\Lambda \in \mathbb{R}^{4\times 4}$ a ↑ *Lorentz transformation*.

For the special case $a = 0 \in \mathbb{R}^4$ we found:

$$\textit{Homogeneous}\text{ Lorentz transformations:}\quad \bar{x} = \varphi(x) = \Lambda x \tag{4.6}$$

ii | Since these transformations are affine, we find immediately:

$$\frac{\partial \bar{x}^\mu}{\partial x^\nu} = \Lambda^\mu{}_\nu \quad \text{and} \quad \frac{\partial x^\mu}{\partial \bar{x}^\nu} = (\Lambda^{-1})^\mu{}_\nu \equiv \Lambda_\nu{}^\mu \tag{4.7}$$

Recall that the derivative of a linear (affine) map is simply the matrix which defines the map.

¡! We use the tensor-inspired notation $\Lambda^\mu{}_\nu$ for the matrix elements of Λ to allow for well-defined contractions with the metric (→ *later*). In $\Lambda^\mu{}_\nu$, the upper index μ denotes the *rows*, the lower index ν the *columns* of the matrix. The notation $\Lambda_\nu{}^\mu$ for the components of the inverse transformation matrix Λ^{-1} is purely conventional at this point; it will turn out to be consistent with pulling indices up and down with the Minkowski metric (→ *below*).

This allows us to rewrite the coordinate transformation Eq. (4.5) in tensor notation:

$$\bar{x}^\mu = \Lambda^\mu{}_\nu x^\nu + a^\mu \tag{4.8}$$

¡! The matrix-vector product Λx is now given by the Einstein summation (index contraction) highlighted blue. We will stick to this notation whenever possible. Since we are now in the world of tensor calculus, it is strongly discouraged to think of and write rank-2 tensors as "matrices" and contractions as matrix-vector products Λx (even though Λ does not represent the components of a tensor). It is less error-prone (and simpler) to perform computations using the index notation introduced in Chapter 3.

iii | Writing down the most general homogeneous Lorentz transformation is very complicated (and unnecessary). Here we provide the two special Lorentz transformations (boosts) discussed earlier in the new matrix notation, and an example for a spatial rotation about the z-axis:

- Lorentz boost in x-direction $K \xrightarrow{v_x} \bar{K}$ ($\beta_x = v_x/c$):

$$\boxed{\text{Eq. (1.77)} \rightarrow \quad \Lambda^{\mu}{}_{\nu} = [\Lambda_{v_x}]^{\mu}{}_{\nu} = \begin{pmatrix} \gamma & -\beta_x\gamma & 0 & 0 \\ -\beta_x\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}_{\mu\nu}} \tag{4.9}$$

- Lorentz boost in $\hat{v}$-direction $K \xrightarrow{\vec{v}} \bar{K}$ ($v = |\vec{v}|$ and $\tilde{\gamma} := \gamma - 1$):

$$\boxed{\begin{array}{l} \text{Eq. (1.75)} \rightarrow \quad \Lambda^{\mu}{}_{\nu} = [\Lambda_{\vec{v}}]^{\mu}{}_{\nu} = \\ \begin{pmatrix} \gamma & -\beta_x\gamma & -\beta_y\gamma & -\beta_z\gamma \\ -\beta_x\gamma & 1+\tilde{\gamma}{v_x}^2/v^2 & \tilde{\gamma}v_x v_y/v^2 & \tilde{\gamma}v_x v_z/v^2 \\ -\beta_y\gamma & \tilde{\gamma}v_x v_y/v^2 & 1+\tilde{\gamma}{v_y}^2/v^2 & \tilde{\gamma}v_y v_z/v^2 \\ -\beta_z\gamma & \tilde{\gamma}v_x v_z/v^2 & \tilde{\gamma}v_y v_z/v^2 & 1+\tilde{\gamma}{v_z}^2/v^2 \end{pmatrix}_{\mu\nu} \end{array}} \tag{4.10}$$

- Spatial rotation $K \xrightarrow{R_z(\theta),\vec{0}} \bar{K}$ by θ in xy-plane:

$$\boxed{\Lambda^{\mu}{}_{\nu} = [R_z(\theta)]^{\mu}{}_{\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}_{\mu\nu}} \tag{4.11}$$

4 | <u>Metric tensor:</u>

i | We elevate the spacetime manifold M to a pseudo-Riemannian (and Lorentzian) manifold by introducing the following pseudo-Riemannian metric tensor (given in inertial coordinates):

$$⁂\ \textit{Minkowski metric}\ \mathrm{d}s^2 \begin{cases} := (c\mathrm{d}t)^2 - (\mathrm{d}\vec{x})^2 \\ = (\mathrm{d}x^0)^2 - (\mathrm{d}x^1)^2 - (\mathrm{d}x^2)^2 - (\mathrm{d}x^3)^2 \\ = \eta_{\mu\nu}\,\mathrm{d}x^\mu \mathrm{d}x^\nu \end{cases} \tag{4.12a}$$

$$\text{with metric components} \quad \eta_{\mu\nu} = \eta^{\mu\nu} = \underbrace{\begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}_{\mu\nu}}_{\text{Signature } (1,3) \equiv (+,-,-,-)} . \tag{4.12b}$$

- The components $\eta_{\mu\nu}$ of this metric tensor in Eq. (4.12b) are the same for all inertial coordinate systems [→ *Eq.* (4.21) *below*].
- Recall that η^{muv} is the matrix inverse of $\eta_{\mu\nu}$.

→ We call the spacetime manifold equipped with this metric …

$$⁂\ \textit{Minkowski space}: \quad \mathbb{R}^{1,3} \equiv (\mathcal{E} \simeq \mathbb{R}^4, \mathrm{d}s^2) \tag{4.13}$$

- We will always use $\eta_{\mu\nu}$ to denote the components of the Minkowski metric (in an inertial coordinate chart) to distinguish it from a generic metric g_{ij}.
- Note that, informally speaking, $\mathrm{d}s^2$ this is the infinitesimal form of the ← *invariant spacetime interval* Eq. (1.83) we introduced earlier (→ *below*).
- Minkowski space is therefore an example of a ← *Lorentzian manifold*. By fixing a metric, we fixed the *geometry* of spacetime. As we will see in our discussion of GENERAL RELATIVITY, the distinctive feature of Minkowski space is that it is *flat* (it has no curvature). It will turn out that, in reality, this assumption is only valid to some degree: The tenet of GENERAL RELATIVITY is that the *deviations* of spacetime from flat Minkowski space are what we expercience as gravity!

ii | With the metric we can measure "lengths" of trajectories on spacetime:

$\triangleleft$ Time-like trajectory $\gamma : s \mapsto x^\mu(s)$ for $s \in [s_a, s_b]$ in $\mathbb{R}^{1,3}$ →

$$L[\gamma] \overset{3.55}{=} \int_{s_a}^{s_b} \sqrt{\eta_{\mu\nu} \frac{\mathrm{d}x^\mu(s)}{\mathrm{d}s} \frac{\mathrm{d}x^\nu(s)}{\mathrm{d}s}}\,\mathrm{d}s \tag{4.14a}$$

$$\overset{4.12b}{=} \int_{s_a}^{s_b} \sqrt{[\dot{x}^0(s)]^2 - [\dot{x}^1(s)]^2 - [\dot{x}^2(s)]^2 - [\dot{x}^3(s)]^2}\,\mathrm{d}s \tag{4.14b}$$

$$\text{Choose parametrization } s := x^0/c \equiv t \tag{4.14c}$$

$$= \int_{t_a}^{t_b} \underbrace{\sqrt{c^2 - \vec{v}^2(t)}}_{>0 \text{ (time-like)}}\,\mathrm{d}t \tag{4.14d}$$

$$\overset{2.25}{=} c\Delta\tau[\gamma] \tag{4.14e}$$

Thus the "length" $L[\gamma]$ of time-like curves in $\mathbb{R}^{1,3}$ is the ← *proper time* $\Delta\tau[\gamma]$ along the curve defined in Eq. (2.25) (multiplied by c); this explains why the Minkowski metric $\mathrm{d}s^2$ is the right choice for SPECIAL RELATIVITY.

4.2. Four vectors and tensors

5 | Tensors are defined as in Chapter 3, with the restriction to $D = 4$ and that only homogeneous Lorentz transformations Eq. (4.7) are considered as transition maps. To emphasize this, we introduce a new nomenclature:

Tensor calculus	SPECIAL RELATIVITY
Contravariant vector A^i	Contraviarant ⁂ *Lorentz vector / 4-vector* A^μ
Covariant vector B_i	Covariant ⁂ *Lorentz vector / 4-vector* B_μ
(Mixed) tensor $T^i{}_j$	(Mixed) ⁂ *Lorentz tensor / 4-tensor* $T^\mu{}_\nu$
Scalar Φ	⁂ *Lorentz scalar* Φ

Then a generic (p, q) tensor transforms under the coordinate transformation Eq. (4.7) as:

$$\boxed{\bar{T}^{\mu_1\dots\mu_p}{}_{\nu_1\dots\nu_q}(\bar{x}) = [\Lambda^{\mu_1}{}_{\rho_1}\dots\Lambda^{\mu_p}{}_{\rho_p}][\Lambda_{\nu_1}{}^{\pi_1}\dots\Lambda_{\nu_q}{}^{\pi_q}]\,T^{\rho_1\dots\rho_p}{}_{\pi_1\dots\pi_q}(x)} \tag{4.15}$$

6 | With the Minkowski metric, we can reformulate our classification for 4-vectors [recall Eq. (1.85)]:

$$\boxed{\left.\begin{array}{l} X^\mu \text{ time-like} \\ X^\mu \text{ light-like} \\ X^\mu \text{ space-like}\end{array}\right\} \;:\Leftrightarrow\; X^2 = X^\mu X_\mu = (X^0)^2 - (\vec{X})^2 \left\{\begin{array}{l} > 0 \\ = 0 \\ < 0\end{array}\right.} \tag{4.16}$$

A *light-like* 4-vector is also called ⁂ *null*.

¡! We use this classification scheme also for generic Lorentz vectors that are not coordinate differences between a pair of events (→ *below*). Since the pseudo-norm $X^\mu X_\mu = X^2$ is a Lorentz scalar, this classification is independent of the inertial system.

7 | Coordinate functions:

It is a particular feature of *linear* coordinate transformations (here: homogeneous Lorentz transformations) that the coordinate functions themselves transform as contravariant vector fields:

$\lessdot$ Coordinate field $X^\mu(x) := x^\mu \rightarrow$

$$\underbrace{\bar{X}^\mu(\bar{x})}_{\bar{x}^\mu} = \Lambda^\mu{}_\nu \underbrace{X^\nu(x)}_{x^\nu} = \underbrace{\frac{\partial \bar{x}^\mu}{\partial x^\nu}}_{\Lambda^\mu{}_\nu} X^\nu(x) \tag{4.17}$$

We make the identification $X^\mu(x) \equiv x^\mu$ and don't write $X^\mu(x)$ henceforth.

Consequently, we can construct ⁂ *covariant coordinates* (a covariant vector field) via the metric by pulling the index down:

$$\boxed{x_\mu := \eta_{\mu\nu} x^\nu = (x^0, -x^1, -x^2, -x^3) = (ct, -\vec{x})} \tag{4.18}$$

¡! To pull the index of a contravariant vector down, you multiply the spatial components by -1.

8 | ⊲ Coordinates of two events x_A^μ and $x_B^\mu \to \Delta x^\mu := x_B^\mu - x_A^\mu$ Lorentz vector

$$\Delta x^2 \equiv \Delta x^\mu \Delta x_\mu \tag{4.19a}$$
$$\stackrel{\text{def}}{=} \eta_{\mu\nu} \Delta x^\mu \Delta x^\nu \tag{4.19b}$$
$$= (\Delta x^0)^2 - (\Delta x^1)^2 - (\Delta x^2)^2 - (\Delta x^3)^2 \tag{4.19c}$$
$$\stackrel{\text{def}}{=} \Delta s^2 \tag{4.19d}$$

Remember [Eq. (1.84)]: $\Delta s^2 = \Delta \bar{s}^2$ for arbitrary Lorentz transformations

$\to$

$$\underbrace{\eta_{\mu\nu}\, \Delta x^\mu \Delta x^\nu}_{\Delta s^2} = \underbrace{\eta_{\rho\pi}\, \Delta \bar{x}^\rho \Delta \bar{x}^\pi}_{\Delta \bar{s}^2} = \left[\eta_{\rho\pi} \Lambda^\rho{}_\mu \Lambda^\pi{}_\nu\right] \Delta x^\mu \Delta x^\nu \tag{4.20}$$

Since this is true for all events $\Delta x^\mu \stackrel{\circ}{\to}$

$$\boxed{\Lambda^\rho{}_\mu \Lambda^\pi{}_\nu \eta_{\rho\pi} = \eta_{\mu\nu}} \tag{4.21}$$

Concluding Eq. (4.21) from Eq. (4.20) is non-trivial because we consider "norms" $\eta_{\mu\nu}\, \Delta x^\mu \Delta x^\nu$ and not "inner products" $\eta_{\mu\nu}\, \Delta x^\mu \Delta y^\nu$. However, for *symmetric*, real matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, it is true that if $\vec{x}^T \boldsymbol{A} \vec{x} = \vec{x}^T \boldsymbol{B} \vec{x}$ for all real vectors $\vec{x}$, then $\boldsymbol{A} = \boldsymbol{B}$. This is so because $\boldsymbol{A} - \boldsymbol{B}$ is a symmetric matrix that can be diagonalized by an orthogonal matrix and $\vec{x}^T(\boldsymbol{A} - \boldsymbol{B})\vec{x} = 0$. The last condition implies that all eigenvalues of $\boldsymbol{A} - \boldsymbol{B}$ are zero and therefore $\boldsymbol{A} - \boldsymbol{B} = \boldsymbol{0}$. Alternatively you can use the ↓ *polarization identity* to show that the invariance of the Minkowski (pseudo) norm implies the invariance of the Minkowski (pseudo) inner product.

We say:

$$\boxed{\text{Lorentz transformations are ↓ \textit{isometries} of Minkowski space.}} \tag{4.22}$$

With $\det(\eta_{\mu\nu}) \neq 0$, a corollary of Eq. (4.21) is:

$$\boxed{\det\left(\Lambda^\mu{}_\nu\right) = \pm 1} \tag{4.23}$$

If you want to write Eq. (4.21) in the old matrix notation, make the identifcations $\Lambda^\mu{}_\nu = \boldsymbol{\Lambda}_{\mu\nu}$ and $\eta_{\mu\nu} = \boldsymbol{\eta}_{\mu\nu}$. Here, subscripts of bold symbols denote the entries of *matrices* as usual (first index: row; second index: column). Equations that contain matrices (bold symbols) do *not* comply with the syntax of tensor calculus (which is why you should avoid them!).

Eq. (4.21) then reads in matrix notation:

$$\boldsymbol{\Lambda}^T_{\mu\rho} \boldsymbol{\eta}_{\rho\pi} \boldsymbol{\Lambda}_{\pi\nu} = \boldsymbol{\eta}_{\mu\nu} \quad \Leftrightarrow \quad \boldsymbol{\Lambda}^T \boldsymbol{\eta} \boldsymbol{\Lambda} = \boldsymbol{\eta} \tag{4.24}$$

Here we defined the *transposed matrix* as $\boldsymbol{\Lambda}^T_{\mu\rho} := \boldsymbol{\Lambda}_{\rho\mu}$, i.e., the matrix where rows and columns are swapped. Eq. (4.24) immediately implies $\det(\boldsymbol{\Lambda}^T)\det(\boldsymbol{\eta})\det(\boldsymbol{\Lambda}) = \det(\boldsymbol{\eta})$; using $\det(\boldsymbol{\eta}) \neq 0$ and $\det(\boldsymbol{\Lambda}^T) = \det(\boldsymbol{\Lambda})$, we find $\det(\boldsymbol{\Lambda}) = \pm 1$.

9 | Eq. (4.21) →

$$\Lambda^\rho{}_\mu \left[\Lambda^\pi{}_\nu \eta_{\rho\pi} \eta^{\nu\sigma}\right] = \delta^\sigma_\mu \tag{4.25}$$

We can therefore conclude that:

$$\boxed{\Lambda_\rho{}^\sigma := \eta_{\rho\pi}\eta^{\nu\sigma}\Lambda^\pi{}_\nu = (\Lambda^{-1})^\sigma{}_\rho} \tag{4.26}$$

Note that this is consistent with our definition in Eq. (4.7).

In the literature (e.g. Schröder [1]) the concept of a "transposed" transformation is introduced. We refer to it as *"pseudo-adjoint"* transformation instead and label it by *. It is defined analogous to proper adjoints on proper inner product spaces:

$$\underbrace{\eta_{\mu\nu}\Lambda^\nu{}_\rho}_{=:\Lambda_{\mu\rho}} x^\rho y^\mu \stackrel{\text{def}}{=} \langle y, \Lambda x\rangle \stackrel{!}{=} \langle \Lambda^* y, x\rangle \stackrel{\text{def}}{=} \underbrace{\eta_{\mu\nu}(\Lambda^*)^\mu{}_\rho}_{=:(\Lambda^*)_{\nu\rho}} x^\nu y^\rho \,. \tag{4.27}$$

This yields as reasonable definition for the pseudo-adjoint:

$$(\Lambda^*)_{\mu\nu} := \Lambda_{\nu\mu} \quad \Rightarrow \quad (\Lambda^*)^\mu{}_\nu = \Lambda_\nu{}^\mu \stackrel{\text{Eq. (4.26)}}{=} (\Lambda^{-1})^\mu{}_\nu \,. \tag{4.28}$$

One can then define a corresponding matrix $\boldsymbol{\Lambda}^*$ such that $(\Lambda^*)^\mu{}_\nu = \boldsymbol{\Lambda}^*_{\mu\nu}$ and use $(\Lambda^{-1})^\mu{}_\nu = \boldsymbol{\Lambda}^{-1}_{\mu\nu}$ to rewrite the above equation as

$$\boldsymbol{\Lambda}^* = \boldsymbol{\Lambda}^{-1} \,. \tag{4.29}$$

Recall that the pseudo-adjoint is implicitly defined via the inner product. At no point did we claim that the pseudo-adjoint matrix is given by the *transposed* matrix $\boldsymbol{\Lambda}^T$ (which is defined by swapping rows and columns)! To find a relation to the latter, we can rewrite Eq. (4.26) in matrix language:

$$\boldsymbol{\Lambda}^{-1}_{\sigma\rho} = \boldsymbol{\eta}_{\rho\pi}\boldsymbol{\Lambda}_{\pi\nu}\boldsymbol{\eta}^{-1}_{\nu\sigma} = (\boldsymbol{\eta}\boldsymbol{\Lambda}\boldsymbol{\eta})_{\rho\sigma} = (\boldsymbol{\eta}\boldsymbol{\Lambda}^T\boldsymbol{\eta})_{\sigma\rho} \,. \tag{4.30}$$

Here we used that $\boldsymbol{\eta}^{-1} = \boldsymbol{\eta} = \boldsymbol{\eta}^T$ and that $\boldsymbol{M}^T_{ab} := \boldsymbol{M}_{ba}$ for any matrix $\boldsymbol{M}$. So finally:

$$\boldsymbol{\Lambda}^* = \boldsymbol{\Lambda}^{-1} = \boldsymbol{\eta}\boldsymbol{\Lambda}^T\boldsymbol{\eta} \,. \tag{4.31}$$

The take home message is that the *transpose* of a Lorentz transformation (given by swapping columns and rows) is *not* its inverse (there are additional minuses sprinkled in by the metric)! By contrast, the *pseudo-adjoint* (defined via the pseudo-inner product) *is* identical to the inverse.

Warning: In the literature you will find the notation ^{T} instead of * (e.g. Schröder [1]). Then one finds the (correct) relation $(\Lambda^T)^\mu{}_\nu = \Lambda_\nu{}^\mu = (\Lambda^{-1})^\mu{}_\nu$. The problem is that this notation *suggests* that $(\Lambda^T)^\mu{}_\nu \stackrel{\text{↯}}{=} \boldsymbol{\Lambda}^T_{\mu\nu}$ and therefore $\boldsymbol{\Lambda}^{-1} \stackrel{\text{↯}}{=} \boldsymbol{\Lambda}^T$. As shown above, *both equations are wrong!*

10 | <u>Covariant derivative:</u>

i | Since in inertial coordinate systems the Minkowski metric is given by $\eta_{\mu\nu}$, it follows immediately for the Christoffel symbols Eq. (3.74):

$$\Gamma^i{}_{kl} = \frac{1}{2}\eta^{im}(\underbrace{\eta_{mk,l}}_{0} + \underbrace{\eta_{ml,k}}_{0} - \underbrace{\eta_{kl,m}}_{0}) = 0 \tag{4.32}$$

¡! If you would transform into *curvilinear (non-inertial) coordinates*, the Christoffel symbols would *not* vanish – even on flat Minkowski space (➔ Problemset 5). That simple partial derivatives produce Lorentz tensors is therefore a special feature of Minkowski space in inertial coordinates.

$$\xrightarrow{\text{Eq. (3.79)}}$$

$$\boxed{\begin{aligned} \text{Lorentz Scalar:}\quad & \Phi_{;\mu} := \Phi_{,\mu} = \partial_\mu \Phi && (4.33a)\\ \text{Contravariant Lorentz vector:}\quad & A^\mu{}_{;\nu} := A^\mu{}_{,\nu} = \partial_\nu A^\mu && (4.33b)\\ \text{Covariant Lorentz vector:}\quad & B_{\mu;\nu} := B_{\mu,\nu} = \partial_\nu B_\mu && (4.33c) \end{aligned}}$$

ii | ⁂ *4-Gradient*:

This allows us to think of the differential operator ∂_μ itself as a *covariant* Lorentz vector and motivates the introduction of its *contravariant* components:

$$\boxed{\begin{aligned} \partial_\mu \quad &= \frac{\partial}{\partial x^\mu} = \left(\tfrac{1}{c}\partial_t, +\vec{\nabla}\right)^T && (4.34a)\\ \partial^\mu := \eta^{\mu\nu}\partial_\nu &= \frac{\partial}{\partial x_\mu} = \left(\tfrac{1}{c}\partial_t, -\vec{\nabla}\right) && (4.34b) \end{aligned}}$$

Using Eq. (3.5), the transformation laws match that of co- and contravariant Lorentz vectors, respectively:

$$\boxed{\begin{aligned} \bar{\partial}_\mu &= \frac{\partial}{\partial \bar{x}^\mu} = \Lambda_\mu{}^\nu \frac{\partial}{\partial x^\nu} = \Lambda_\mu{}^\nu \partial_\nu && (4.35a)\\ \bar{\partial}^\mu &= \frac{\partial}{\partial \bar{x}_\mu} = \Lambda^\mu{}_\nu \frac{\partial}{\partial x_\nu} = \Lambda^\mu{}_\nu \partial^\nu\,. && (4.35b) \end{aligned}}$$

¡! The *covariant* 4-gradient (index *down*) is the partial derivative wrt. the *contravariant* coordinates (index *up*) and vice versa.

iii | These transformation properties immediately suggest two Lorentz scalars that can be constructed from 4-gradients ($A^\mu = (A^0, \vec{A})$):

$$\boxed{\begin{aligned} ⁂\ \textit{4-divergence:}\quad & \partial A := \partial_\mu A^\mu = \partial^\mu A_\mu = \tfrac{1}{c}\partial_t A^0 + \vec{\nabla}\cdot\vec{A} && (4.36a)\\ ⁂\ \textit{4-Laplacian:}\quad & \Box \equiv \partial^2 := \partial_\mu \partial^\mu = \left(\tfrac{1}{c}\partial_t\right)^2 - \vec{\nabla}^2 && (4.36b) \end{aligned}}$$

The 4-Laplacian $\Box$ is also known as ↓ *d'Alembert operator*.

Examples:

- In electrodynamics (→ *later*) the gauge potential transforms as a contravariant Lorentz vector $A^\mu = (\frac{1}{c}\varphi, \vec{A})$.

  The ↓ *Lorenz gauge* is defined as $\partial_\mu A^\mu = 0$; it is Lorentz invariant since the 4-divergence is a Lorentz scalar: $\bar{\partial}_\mu \bar{A}^\mu(\bar{x}) = \partial_\mu A^\mu(x)$.

  Note: The Lorenz gauge is named after ↑ *Ludvig Lorenz*; by contrast, the Lorentz transformation is named after ↑ *Hendrik Lorentz*. Thus: *The Lorenz gauge* (no "t") *is Lorentz invariant.*

- In vacuum (and in Lorenz gauge), the gauge field of electrodynamics satisfies the wave equation

$$\partial^2 A^\mu = \left[\left(\tfrac{1}{c}\partial_t\right)^2 - \vec{\nabla}^2\right] A^\mu = 0\,. \tag{4.37}$$

Since ∂^2 is a Lorentz scalar and A^μ a Lorentz vector, $\partial^2 A^\mu$ transforms as a contravariant Lorentz vector and the equation is *manifestly Lorentz covariant*:

$$\partial^2 A^\mu(x) = 0 \quad \Leftrightarrow \quad \bar{\partial}^2 \bar{A}^\mu(\bar{x}) = 0\,. \tag{4.38}$$

- If we have a scalar field Φ, we can construct a manifestly Lorentz covariant wave equation:

$$(\partial^2 + m^2)\Phi(x) = 0 \quad \Leftrightarrow \quad (\bar{\partial}^2 + m^2)\bar{\Phi}(\bar{x}) = 0\,. \tag{4.39}$$

The parameter m is arbitrary and plays the role of a mass (spectral gap) of the excitations. This equation is known as ↑ *Klein-Gordon equation* and describes, for example, the classical equation of motion of the Higgs field (without interactions).

11 | Relative tensors → ⁂ *Lorentz pseudo tensor*:

Since $\det(\Lambda) = \pm 1$, the classification of tensors simplifies:

$$\boxed{\begin{aligned} \text{Tensor:}\quad & \bar{T}^M{}_N(\bar{x}) = \Lambda^M{}_R \Lambda_N{}^P T^R{}_P(x) \qquad (4.40a) \\ \textit{Pseudo}\text{ tensor:}\quad & \bar{T}^M{}_N(\bar{x}) = \det(\Lambda)\, \Lambda^M{}_R \Lambda_N{}^P T^R{}_P(x) \qquad (4.40b) \end{aligned}}$$

Here we use again a multi-index notation: $M = \mu_1, \dots, \mu_p$ etc. Recall that $\det(\Lambda) = \pm 1$; pseudo tensors therefore pick up an additional minus sign under parity or time inversion (→ *later*).

→ Relative tensors of *odd* weight w are pseudo tensors under Lorentz transformations.

Example:

The Levi-Civita symbol is a Lorentz pseudo tensor [recall Eq. (3.42)]:

$$\bar{\varepsilon}^{\mu\nu\rho\pi} = \varepsilon^{\mu\nu\rho\pi} = \det(\Lambda)\, \Lambda^\mu{}_{\mu'} \Lambda^\nu{}_{\nu'} \Lambda^\rho{}_{\rho'} \Lambda^\pi{}_{\pi'}\, \varepsilon^{\mu'\nu'\rho'\pi'}\,. \tag{4.41}$$

This means that if you contract a Levi-Civita symbol with an actual $(0,4)$ Lorentz tensor like $F_{\mu\nu}F_{\rho\pi}$ (the tensor product of two electromagnetic field trength tensors), you obtain a *pseudo (Lorentz) scalar*:

$$\bar{\Phi}(\bar{x}) := \bar{\varepsilon}^{\mu\nu\rho\pi}\, \bar{F}_{\mu\nu}\bar{F}_{\rho\pi} \stackrel{\circ}{=} \det(\Lambda)\, \varepsilon^{\mu\nu\rho\pi} F_{\mu\nu}F_{\rho\pi} = \det(\Lambda)\,\Phi(x)\,. \tag{4.42}$$

Since this is a quadratic (pseudo) scalar quantity, you might try to add it to the Lagrangian of Maxwell theory ($\theta \in \mathbb{R}$):

$$\tilde{\mathcal{L}} = -\tfrac{1}{4} F^{\mu\nu}F_{\mu\nu} + \theta\varepsilon^{\mu\nu\rho\pi} F_{\mu\nu}F_{\rho\pi}\,. \tag{4.43}$$

(This Lagrangian is now only invariant under Lorentz transformations with $\det(\Lambda) = +1$.)

The new term is called ↑ *θ-term*. One can show that it is a total derivative and therefore does not affect the classical equations of motion (Maxwell's equations). However, for non-abelian generalizations of electrodynamics like ↑ *quantum chromodynamics* (↑ *Yang-Mills theories*), it *does* affect the theory (↑ *Strong CP-problem* [55]).

Note that we did not use the metric tensor $\eta_{\mu\nu}$ to construct the term $\varepsilon^{\mu\nu\rho\pi}F_{\mu\nu}F_{\rho\pi}$ (as compared to $F^{\mu\nu}F_{\mu\nu}$ where we need it to pull two indices up); this makes the θ-term an example of a so called ↑ *topological* term (↑ *topological field theory*): the term doesn't "see" the *geometry* of spacetime! In condensed matter physics, the term plays a role in the description of ↑ *topological insulators* [56].

12 | In the next chapter we want to construct a relativistic version of classical mechanics (using the framework of tensors calculus to make the equations Lorentz covariant). As a preparation, we can already define two 4-vectors with physical interpreation:

i | 4-velocity:

Question: What is a reasonable definition for a relativistic (= Lorentz covariant) velocity?

$\sphericalangle$ Particle trajectory $x^\mu(\lambda)$ parametrized by λ:

$$x^\mu(\lambda) = \begin{pmatrix} ct(\lambda) \\ \vec{x}(\lambda) \end{pmatrix} \quad \Rightarrow \quad \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda} = \begin{pmatrix} \frac{\mathrm{d}ct}{\mathrm{d}\lambda} \\ \frac{\mathrm{d}\vec{x}}{\mathrm{d}\lambda} \end{pmatrix} \tag{4.44}$$

First try: $\lambda = t$ (coordinate time) $\rightarrow$

$$\frac{\mathrm{d}x^\mu}{\mathrm{d}t} = \begin{pmatrix} c \\ \frac{\mathrm{d}\vec{x}}{\mathrm{d}t} \end{pmatrix} = \begin{pmatrix} c \\ \vec{v}(t) \end{pmatrix} \tag{4.45}$$

with coordinate velocity $\vec{v}(t)$.

Problem:

$\frac{\mathrm{d}x^\mu}{\mathrm{d}t}$ is *not* a contravariant Lorentz vector because $\mathrm{d}t \neq \mathrm{d}\bar{t}$ is not a Lorentz scalar. That is:

$$\frac{\mathrm{d}\bar{x}^\mu}{\mathrm{d}\bar{t}} \neq \Lambda^\mu{}_\nu \frac{\mathrm{d}x^\nu}{\mathrm{d}t} \tag{4.46}$$

$\rightarrow$ Eq. (4.45) is useless to construct Lorentz covariant equations!

Idea: The ← *Proper time* τ is a Lorentz scalar [Eq. (2.24)]: $\mathrm{d}\tau = \mathrm{d}\bar{\tau}$

$\rightarrow$ Set $\lambda = \tau$:

$$\boxed{⁂\ \textit{4-velocity}: \quad u^\mu := \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} = \begin{pmatrix} c\frac{\mathrm{d}t}{\mathrm{d}\tau} \\ \frac{\mathrm{d}\vec{x}}{\mathrm{d}\tau} \end{pmatrix} = \gamma_v \begin{pmatrix} c \\ \vec{v} \end{pmatrix}} \tag{4.47}$$

Here we used $\frac{\mathrm{d}t}{\mathrm{d}\tau} = \gamma_{v(t)}$ [recall Eq. (2.23)].

By construction, the 4-velocity is a contravariant Lorentz vector: $\bar{u}^\mu = \Lambda^\mu{}_\nu u^\nu$.

$\sphericalangle$ Pseudo-norm:

$$\boxed{u^2 = \eta_{\mu\nu} u^\mu u^\nu = (u^0)^2 - (\vec{u})^2 \stackrel{\circ}{=} c^2 > 0} \tag{4.48}$$

$\rightarrow$ *Time-like* 4-vector

In Minkowski space, u^μ is the tangent at x^μ of the world line $x^\mu(\tau)$.

ii | 4-acceleration:

Following the same line of arguments above, the 4-acceleration is then defined as the derivative of the 4-velocity wrt. the proper time:

⁂ *4-acceleration:*

$$\boxed{b^\mu := \frac{\mathrm{d}u^\mu}{\mathrm{d}\tau} = \begin{pmatrix} c\frac{\mathrm{d}\gamma_{v(t)}}{\mathrm{d}\tau} \\ \frac{\mathrm{d}[\gamma_{v(t)}\vec{v}(t)]}{\mathrm{d}\tau} \end{pmatrix} \stackrel{\circ}{=} \begin{pmatrix} \gamma_v^4 \frac{\vec{v}\cdot\vec{a}}{c} \\ \gamma_v^2 \vec{a} + \gamma_v^4 \frac{\vec{v}\cdot\vec{a}}{c^2}\vec{v} \end{pmatrix}} \tag{4.49}$$

Here $\vec{a} := \frac{d\vec{v}(t)}{dt}$ is the coordinate acceleration or 3-acceleration.

It is now easy to show that $b^2 = b_\mu b^\mu < 0$ is a *space-like* Lorentz vector and that

$$\boxed{\frac{d(u^\mu u_\mu)}{d\tau} = \frac{d(c^2)}{d\tau} = 0 \quad \Rightarrow \quad u^\mu b_\mu = 0\,,} \tag{4.50}$$

i.e., the 4-acceleration is always "orthogonal" (in terms of the Minkowski metric) to the 4-velocity.

4.3. The complete Lorentz group

Details: ➔ Problemset 5

1 | The Lorentz group is a matrix group defined as the homogenous isometry group of the Minkowski metric η:

$$\boxed{⁂\ \textit{Lorentz group}: \quad \mathrm{O}(1,3) := \left\{ \boldsymbol{\Lambda} \in \mathbb{R}^{4\times 4} \,\middle|\, \boldsymbol{\Lambda}^T \boldsymbol{\eta} \boldsymbol{\Lambda} = \boldsymbol{\eta} \right\}} \tag{4.51}$$

with identification $\Lambda^\mu{}_\nu = \boldsymbol{\Lambda}_{\mu\nu}$ and $\eta_{\mu\nu} = \boldsymbol{\eta}_{\mu\nu}$.

- As shown previously [Eq. (4.21) and Eq. (4.24)], the matrix constraint in Eq. (4.51) is equivalent to the property

$$\eta_{\mu\nu}\, x^\mu y^\nu \overset{\text{def}}{=} \eta(x,y) \overset{!}{=} \eta(\Lambda x, \Lambda y) \overset{\text{def}}{=} \left[\eta_{\rho\pi} \Lambda^\rho{}_\mu \Lambda^\pi{}_\nu\right] x^\mu y^\nu \tag{4.52}$$

  for all 4-vectors x, y. Namely, the transformations Λ do not change the inner product (and thereby length) of arbitrary vectors; maps with this feature are called ↑ *isometries*.
- If you replace the Minkowski metric $\eta_{\mu\nu} = \mathrm{diag}\,(+1,-1,-1,-1)$ by the Euclidean metric $\delta_{\mu\nu} = \mathrm{diag}\,(+1,+1,+1,+1)$, the homogeneous isometry constraint becomes $\boldsymbol{\Lambda}^T\boldsymbol{\Lambda} = \mathbb{1}$ since $\boldsymbol{\delta} = \mathbb{1}$ is the identity matrix; this constraint characterizes *orthogonal matrices*. The homogenous isometry group of a $D = 4$ Euclidean space is therefore O(4): the group of four-dimensional *rotations* and *reflections*.

2 | Continuous Lorentz transformations:

i | Mathematical fact: $\mathrm{O}(1,3)$ is a ↑ *Lie group* (= a group that is also a differentiable manifold)

To be precise: $\mathrm{O}(1,3)$ is a 6-dimensional (→ *below*) ↑ *non-compact* ↓ *non-abelian* disconnected (→ *below*) real matrix Lie group with components that are not ↑ *simply connected*.

→ In a neighborhood of $\mathbb{1}$, elements of Lie groups can be written as exponentials:

$$\boldsymbol{\Lambda} = \exp(\boldsymbol{X}) \quad \text{with} \quad \boldsymbol{X} \in \mathfrak{o}(1,3) \tag{4.53}$$

where $\mathfrak{o}(1,3)$ denotes the ↑ *Lie algebra* (= vector space with a Lie bracket):

$$\mathfrak{o}(1,3) = \left\{ \boldsymbol{X} \in \mathbb{R}^{4\times 4} \,\middle|\, \exp(t\boldsymbol{X}) \in \mathrm{O}(1,3) \ \text{ for all } \ t \in \mathbb{R} \right\}. \tag{4.54}$$

ii | The isometry constraint on the group elements can be translated into the Lie algebra:

$$\Lambda^T \eta \Lambda = \eta \quad \overset{\text{Eq. (4.53)}}{\Longleftrightarrow} \quad X^T = -\eta X \eta \tag{4.55}$$

→ Most general form of X:

$$X = \begin{pmatrix} 0 & a & b & c \\ a & 0 & -d & -e \\ b & d & 0 & -f \\ c & e & f & 0 \end{pmatrix} \quad \text{with} \quad a, \dots, f \in \mathbb{R} \tag{4.56}$$

Proof: ➔ Problemset 5

→

- $\dim(\mathfrak{o}(1,3)) = 6$

  This is why $O(1,3)$ is a 6-dimensional Lie group.

- $\mathrm{Tr}[X] = 0 \;\Rightarrow\; \det \Lambda = \det[\exp(X)] = \exp(\mathrm{Tr}[X]) = 1$

  → All Lorentz transformations connected to the identity have positive determinant. Recall that we found previously $\det \Lambda = \pm 1$, so we should not expect to find *all* elements of $O(1,3)$ in this way.

iii | Generators = Basis of $\mathfrak{o}(1,3)$ [57]:

We use the shorthand $+$ $(-)$ for $+1$ (-1).

$$\boxed{L_x = \begin{pmatrix} 0&0&0&0\\0&0&0&0\\0&0&0&-\\0&0&+&0 \end{pmatrix}, \quad L_y = \begin{pmatrix} 0&0&0&0\\0&0&0&+\\0&0&0&0\\0&-&0&0 \end{pmatrix}, \quad L_z = \begin{pmatrix} 0&0&0&0\\0&0&-&0\\0&+&0&0\\0&0&0&0 \end{pmatrix}} \tag{4.57a}$$

$$\boxed{K_x = \begin{pmatrix} 0&+&0&0\\+&0&0&0\\0&0&0&0\\0&0&0&0 \end{pmatrix}, \quad K_y = \begin{pmatrix} 0&0&+&0\\0&0&0&0\\+&0&0&0\\0&0&0&0 \end{pmatrix}, \quad K_z = \begin{pmatrix} 0&0&0&+\\0&0&0&0\\0&0&0&0\\+&0&0&0 \end{pmatrix}} \tag{4.57b}$$

Interpretation:

$$\exp(\varphi L_x) \triangleq \begin{pmatrix} 1&0&0&0\\0&1&0&0\\0&0&\cos\varphi&-\sin\varphi\\0&0&\sin\varphi&\cos\varphi \end{pmatrix} = \Lambda_{R_x(\varphi)} \quad \rightarrow \textit{Rotation} \text{ around } x\text{-axis} \tag{4.58a}$$

$$\exp(-\theta K_x) \triangleq \begin{pmatrix} \cosh\theta & -\sinh\theta & 0 & 0\\ -\sinh\theta & \cosh\theta & 0 & 0\\ 0&0&1&0\\0&0&0&1 \end{pmatrix} = \Lambda_{v_x} \quad \rightarrow \textit{Boost} \text{ in } x\text{-direction} \tag{4.58b}$$

with ← *rapidity* $\tanh\theta = \frac{v_x}{c} \in (-1,1)$ (➔ Problemset 3) and rotation angle $\varphi \in [0, 2\pi)$.

→

$$\boxed{\begin{aligned} L_x, L_y, L_z &: \quad \text{Generators of } \textit{rotations} \\ K_x, K_y, K_z &: \quad \text{Generators of } \textit{boosts} \end{aligned}} \quad \begin{aligned}(4.59a)\\(4.59b)\end{aligned}$$

An arbitrary element of O(1,3) that is connected to the identity can then be written as

$$\boldsymbol{\Lambda} = \exp\left(\sum_i \varphi_i \boldsymbol{L}_i - \sum_i \theta_i \boldsymbol{K}_i\right) \quad \text{with} \quad i \in \{x,y,z\}\,. \tag{4.60}$$

In particular [57]:

$$\text{Pure boost:} \quad \boldsymbol{\Lambda}_{\vec{v}} \equiv \boldsymbol{\Lambda}_{\vec{\theta}} = \exp\left(-\vec{\theta}\cdot\vec{\boldsymbol{K}}\right) \tag{4.61a}$$

$$\text{Pure rotation:} \quad \boldsymbol{\Lambda}_{R_{\vec{\varphi}}} = \exp\left(\vec{\varphi}\cdot\vec{\boldsymbol{L}}\right) \tag{4.61b}$$

with rotation angle $\varphi = |\vec{\varphi}|$, rotation axis $\hat{\varphi} = \vec{\varphi}/\varphi$, and rapidity vector

$$\vec{\theta} \equiv \vec{\theta}(\vec{v}) := \hat{v}\tanh^{-1}\left(\frac{v}{c}\right)\,. \tag{4.62}$$

¡! The rapidity vector $\vec{\theta}$ is *not* given by the rapidities $\tanh^{-1}\frac{v_i}{c}$ of the components v_i of $\vec{v}$.

iv | Lie algebra:

The Lie bracket (= commutator) on the Lie *algebra* determines the multiplicative structure of the Lie *group* via the ↓ *Baker-Campbell-Hausdorff formula*:

$$\exp(X)\cdot\exp(Y) = \exp\left(X + Y + \frac{1}{2}[X,Y] + \dots\right)\,. \tag{4.63}$$

→ The Lie algebra $\mathfrak{o}(1,3)$ determines the (local) group structure of O(1,3):

Eq. (4.57) $\overset{\circ}{\rightarrow}$

$$\boxed{\begin{aligned} [L_i, L_j] &= \varepsilon^{ijk} L_k & (4.64a)\\ [L_i, K_j] &= \varepsilon^{ijk} K_k & (4.64b)\\ [K_i, K_j] &= -\varepsilon^{ijk} L_k & (4.64c)\end{aligned}}$$

Some comments and implications:

- ¡! Because of Eq. (4.64) [and Eq. (4.63)], you *cannot* simply combine exponentials:

$$\exp\left(-\vec{\theta}\cdot\vec{\boldsymbol{K}}\right)\cdot\exp\left(\vec{\varphi}\cdot\vec{\boldsymbol{L}}\right) \neq \exp\left(\vec{\varphi}\cdot\vec{\boldsymbol{L}} - \vec{\theta}\cdot\vec{\boldsymbol{K}}\right)\,, \tag{4.65a}$$

$$\exp\left(-\vec{\theta}\cdot\vec{\boldsymbol{K}}\right)\cdot\exp\left(-\vec{\theta}'\cdot\vec{\boldsymbol{K}}\right) \neq \exp\left(-(\vec{\theta}+\vec{\theta}')\cdot\vec{\boldsymbol{K}}\right)\,, \tag{4.65b}$$

$$\exp\left(\vec{\varphi}\cdot\vec{\boldsymbol{L}}\right)\cdot\exp\left(\vec{\varphi}'\cdot\vec{\boldsymbol{L}}\right) \neq \exp\left((\vec{\varphi}+\vec{\varphi}')\cdot\vec{\boldsymbol{L}}\right)\,. \tag{4.65c}$$

  This is why the concatenation of Lorentz transformations is quite complicated in general.

- Eq. (4.64a) is written in physics often as $[L_i, L_j] = i\hbar\varepsilon^{ijk}L_k$ with angular momentum operators L_k. In this notation, they generate rotations $U_{\vec{\omega}} = \exp(\frac{i}{\hbar}\vec{\omega}\vec{L})$. The additional phase i in the commutation relation matches a corresponding factor in an alternative definition of the generators $\vec{L}$. (Recall that the L_i in Eq. (4.57) are *anti-Hermitian* whereas in physics we often prefer *Hermitian* operators.)
- Eq. (4.64a) shows that $\mathfrak{o}(3) := \operatorname{span}\{L_x, L_y, L_z\}$ forms a *subalgebra* of $\mathfrak{o}(1,3)$. On the group level, this means that the group of spatial rotations SO(3) is a subgroup of the full Lorentz group O(1,3).

By contrast, Eq. (4.64c) shows that the boost generators $\{K_x, K_y, K_z\}$ do *not* form a subalgebra, but mix with rotations. This implies that there is no "subgroup of pure boosts" in $\mathrm{O}(1,3)$. In particular:

$$\Lambda_{\vec{v}}\Lambda_{\vec{u}} = \Lambda_{\vec{u}\oplus\vec{v}}\Lambda_{R(\vec{u},\vec{v})} \tag{4.66}$$

with the ← *Thomas-Wigner rotation* $R(\vec{u},\vec{v}) \in \mathrm{SO}(3)$ [recall Section 2.3].

- There is a more compact, 4-vector-inspired notation for the 6 generators in Eq. (4.57), namely [58]:

$$\left(J^{\alpha\beta}\right)^{\mu}{}_{\nu} \equiv \left(J^{\alpha\beta}\right)_{\mu\nu} := \eta^{\alpha\mu}\delta^{\beta}_{\nu} - \eta^{\beta\mu}\delta^{\alpha}_{\nu}\,. \tag{4.67}$$

Inspection shows that (➔ Problemset 5)

$$L_x = J^{23} = -J^{32}\,,\qquad K_x = J^{01} = -J^{10}\,, \tag{4.68a}$$

$$L_y = J^{31} = -J^{13}\,,\qquad K_y = J^{02} = -J^{20}\,, \tag{4.68b}$$

$$L_z = J^{12} = -J^{21}\,,\qquad K_z = J^{03} = -J^{30}\,. \tag{4.68c}$$

The three equations of the Lie algebra Eq. (4.64) can then be condensed into a single equation [58]:

$$[J^{\mu\nu}, J^{\rho\sigma}] = \eta^{\nu\rho}J^{\mu\sigma} - \eta^{\mu\rho}J^{\nu\sigma} - \eta^{\nu\sigma}J^{\mu\rho} + \eta^{\mu\sigma}J^{\nu\rho}\,. \tag{4.69}$$

This form is useful to construct other representations of the Lorentz group, especially in relativistic quantum mechanics (→ *Dirac equation*).

v | It is a useful mathematical fact that every continuous Lorentz transformation of the form Eq. (4.60) can be decomposed *uniquely* as follows:

$$\boxed{\Lambda = \Lambda_{\vec{v}}\Lambda_R = \Lambda_R\Lambda_{\vec{w}}} \tag{4.70a}$$

with parameters:

$$\frac{v_i}{c} = -\frac{\Lambda_{i0}}{\Lambda_{00}}\,,\quad \frac{w_i}{c} = -\frac{\Lambda_{0i}}{\Lambda_{00}}\quad\text{and}\quad R_{ij} = \Lambda_{ij} - \frac{\Lambda_{i0}\Lambda_{0j}}{1+\Lambda_{00}} \tag{4.70b}$$

$\Lambda_{\vec{v}}$ and Λ_R are defined in Eq. (4.61a) [or Eq. (1.75)] and Eq. (4.61b) [or Eq. (1.40)].

The proof can be found in Ref. [59]. This decomposition, sometimes referred to as ⁂ *rotation-boost decomposition*, relates to the mathematical concept of ↑ *Cartan decompositions* [60].

If we use the multiplicative law $\Lambda_R\Lambda_{\vec{v}}\Lambda_{R^{-1}} = \Lambda_{R\vec{v}}$ [recall Eq. (1.43a)] and choose R such that $R\vec{v} = (v_x, 0, 0)^T$, we can also find a decomposition of the form

$$\boxed{\Lambda = \Lambda_{R_1}\Lambda_{v_x}\Lambda_{R_2}} \tag{4.71}$$

with appropriately chosen rotations $R_1, R_2 \in \mathrm{SO}(3)$ and a boost in x-direction by v_x.

3 | Discrete generators:

It is easy to verify that the following two matrices also belong to $\mathrm{O}(1,3)$:

⁂ *Parity*:
$P\,:\,(t,\vec{x})\mapsto(t,-\vec{x})$

$$\Rightarrow\quad P^{\mu}{}_{\nu}\equiv \boldsymbol{P}_{\mu\nu}:=\begin{pmatrix}+1&0&0&0\\0&-1&0&0\\0&0&-1&0\\0&0&0&-1\end{pmatrix}_{\mu\nu} \tag{4.72}$$

⁂ *Time reversal*:
$T\,:\,(t,\vec{x})\mapsto(-t,\vec{x})$

$$\Rightarrow\quad T^{\mu}{}_{\nu}\equiv \boldsymbol{T}_{\mu\nu}:=\begin{pmatrix}-1&0&0&0\\0&+1&0&0\\0&0&+1&0\\0&0&0&+1\end{pmatrix}_{\mu\nu} \tag{4.73}$$

In contrast to the continuous group elements above: $\det(P)=\det(T)=-1$

$\rightarrow$ P and T are not generated by boosts or rotations!

4 | Structure of the Lorentz group:

Combining the discrete transformation $\boldsymbol{P}$ and $\boldsymbol{T}$ with the continuous transformations $\boldsymbol{\Lambda}=\exp(\boldsymbol{X})$ yields the complete group $\mathrm{O}(1,3)$. Let us study its structure:

i | $\det(\Lambda)=\pm 1\rightarrow$

$$O(1,3)=\underbrace{L_+}_{\det(\Lambda)=+1}\cup\underbrace{L_-}_{\det(\Lambda)=-1} \tag{4.74}$$

All Lorentz transformations that are continuously connected to $\mathbb{1}$ are in L_+. One can transition between L_+ and L_- by applying either T or P.

ii | In addition, we find:

$$1=\eta_{00}\overset{4.21}{=}\left(\Lambda^0{}_0\right)^2-\sum_{k=1}^{3}\left(\Lambda^k{}_0\right)^2\leq\left(\Lambda^0{}_0\right)^2 \tag{4.75}$$

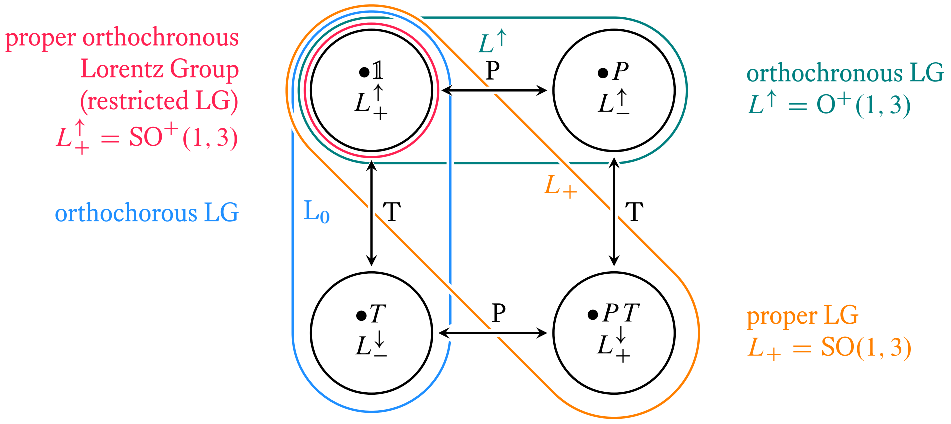
Thus $\Lambda^0{}_0\neq 0$ and $\operatorname{sign}(\Lambda^0{}_0)=\pm 1$ can be used to characterize Lorentz transformations. Note that $\operatorname{sign}(P^0{}_0)=+1$ but $\operatorname{sign}(T^0{}_0)=-1$ and $\operatorname{sign}((PT)^0{}_0)=-1$.

iii | Neither $\det(\Lambda)=\pm 1$ nor $\operatorname{sign}(\Lambda^0{}_0)=\pm 1$ can be changed by continuously deforming a Lorentz transformation.

$\rightarrow$ Four disconnected components of $\mathrm{O}(1,3)$:

$$L_+^{\uparrow}:\ \det(\Lambda)=+1\ \text{ and }\ \operatorname{sign}(\Lambda^0{}_0)=+1\quad(\mathbb{1}\in L_+^{\uparrow}) \tag{4.76a}$$
$$L_-^{\uparrow}:\ \det(\Lambda)=-1\ \text{ and }\ \operatorname{sign}(\Lambda^0{}_0)=+1\quad(P\in L_-^{\uparrow}) \tag{4.76b}$$
$$L_+^{\downarrow}:\ \det(\Lambda)=+1\ \text{ and }\ \operatorname{sign}(\Lambda^0{}_0)=-1\quad(PT\in L_+^{\downarrow}) \tag{4.76c}$$
$$L_-^{\downarrow}:\ \det(\Lambda)=-1\ \text{ and }\ \operatorname{sign}(\Lambda^0{}_0)=-1\quad(T\in L_-^{\downarrow}) \tag{4.76d}$$

Graphically:



$$L_\pm^{\uparrow/\downarrow} \begin{cases} \uparrow & \text{no time inversion } (\operatorname{sign} \Lambda^0{}_0 = +1) \\ \downarrow & \text{time inversion } (\operatorname{sign} \Lambda^0{}_0 = -1) \\ + & \det \Lambda = +1 \text{ (proper)} \\ - & \det \Lambda = -1 \text{ (improper)} \end{cases}$$

iv | Subgroups: We can define the following four subgroups of $\mathrm{O}(1,3)$:

$$\text{⁂ } \textit{Proper LG}: \quad \mathrm{SO}(1,3) \equiv L_+ := L_+^\uparrow \cup L_+^\downarrow \tag{4.77a}$$

$$\text{⁂ } \textit{Orthochronous LG}: \quad \mathrm{O}^+(1,3) \equiv L^\uparrow := L_+^\uparrow \cup L_-^\uparrow \tag{4.77b}$$

$$\text{⁂ } \textit{Proper orthochronous LG}: \quad \mathrm{SO}^+(1,3) := L_+^\uparrow \tag{4.77c}$$

$$\text{⁂ } \textit{Orthchorous LG}: \quad L_0 := L_+^\uparrow \cup L_-^\downarrow \tag{4.77d}$$

Note that subgroups must contain the identity $\mathbb{1}$!

In Greek, "chrónos" (*χρόνος*) means "time" and "chóros" (*χώρος*) means "space".

According to modern physics, Einstein's principle of relativity **SR** reads formally:

> All fundamental theories of nature must be invariant under the *proper orthochronous Lorentz group* $\mathrm{SO}^+(1,3)$.

- This does not prevent specific theories to have *additional* symmetries. ↑ *Quantum electrodynamics* (QED), for example, is invariant under the *full* Lorentz group $O(1,3)$. This means that phenomena of electromagnetism – and its interaction with charged particles – are also symmetric under time inversion T and parity P.

  So far, observations suggest that, besides the electromagnetic force, also gravity and the strong force are symmetric under P and T. (Interestingly, there is no formal reason *why* the strong force should *not* break P and T; the fact that it does *not* violate these symmetries is called the ↑ *strong CP problem*).

- However, today we know that there *are* terms in the standard model of particle physics that *violate* both P and T. For example, the weak interaction (responsible for radioactive β-decay) violates parity P strongly (↑ *Wu experiment*). This means that you can use experiments that depend on the weak interaction to tell the difference between our world and its mirror image (or a right-handed and a left-handed coordinate system). There are also weak terms (concerning quarks) that violate time reversal T (↑ CP *violation*). As a consequence, the standard model as a whole is only invariant under the proper orthochronous Lorentz group $\mathrm{SO}^+(1,3)$.

This explains why we can only require symmetry under $\mathrm{SO}^+(1,3)$, and not the full Lorentz group $\mathrm{O}(1,3)$: We already *know* by experiments that the latter is *not* a fundamental symmetry of nature!

- The fact that there are processes that violate parity symmetry P contradicts our everyday experience: If you run an experiment using equipment found in a school physics lab and put a mirror next to it, there is no way to decide whether you are watching the experiment directly or via the mirror (i.e., parity inverted). The reason is that the physics we experience in everyday life is goverened by electrodynamics and gravity, both of which are invariant under P. To unveil that nature secretly violates P, you must perform an experiment that involves the weak interaction (that is: a particle physics experiment). This is what Chien-Shiung Wu did in her now famous ↑ *Wu experiment*. At the time, the result (that P is not a symmetry of nature) was unexpected and groundbreaking.

  So if you are surprised that P is not a symmetry of nature, you are not alone. Here is how Wolfang Pauli reacted to the result of the Wu experiment [61]:

  > *At one point, Temmer found himself in the presence of eminence grise Wolfgang Pauli, who asked for the latest news from the United States. Temmer told him that parity was no longer to be assumed conserved. "That's total nonsense" averred the great man. Temmer: "I assure you the experiment says it is not." Pauli (curtly): "Then it must be repeated!"*

4.4. ‡ Why is spacetime 3+1 dimensional?

Given the discussions in Chapter 3 and Chapter 4 it is clear that the mathematical formalism allows for straightforward generalizations to higher- (or lower-) dimensional spacetime manifolds with arbitrary signatures; these suggest spacetimes with various numbers of spatial and temporal dimensions.

It is therefore natural to ask:

Is there anything special about our $3+1$-dimensional world?

What follows is not a *proof* that spacetime must be $3+1$ dimensional. Our goal is to argue that all spacetimes, except ours with *three* space and *one* time dimension, face severe problems that, most likely, would not allow for complex life.

The following discussion is based on Tegmark [39, 62].

1 | ◁ Pseudo-Riemannian manifold of signature (t,s) with metric

$$g_{ij} = \mathrm{diag}(\underbrace{+1,\dots,+1}_{t},\underbrace{-1,\dots,-1}_{s}) \tag{4.78}$$

- This is the generalization of Minkowski space to a (flat) spacetime manifold with, naïvely, t time and s space-dimensions.
- Most of our discussions in this chapter can be transferred to this more general setting.

2 | ∢ ↑ *Klein-Gordon equation* for signature (t,s):

$$\boxed{\left(\partial^2+m^2\right)\Phi=\underbrace{\sum_{i=1}^{t}\frac{\partial^2\Phi}{\partial {x^i}^2}}_{t\,\times\,\text{Time (?)}}-\underbrace{\sum_{i=t+1}^{s+t}\frac{\partial^2\Phi}{\partial {x^i}^2}}_{s\,\times\,\text{Space (?)}}+m^2\Phi=0} \tag{4.79}$$

- Recall that $\partial^2=g^{ij}\partial_i\partial_j$ where g^{ij} is given by (the inverse of) Eq. (4.78).
- The Klein-Gordon equation (KGE) is the simplest covariant field equation. It describes the time evolution of a scalar field of mass m. It is ubiquitous in relativistic physics (especially in ↑ *quantum field theory*).
- For example, the components of the electromagnetic field in vacuum are described by the KGE for $m=0$ and $(t,s)=(1,3)$ (which is then referred to as ↓ *wave equation*):

$$\partial^2 E_i=\tfrac{1}{c^2}\partial_t^2 E_i-\nabla^2 E_i=0\,, \tag{4.80a}$$

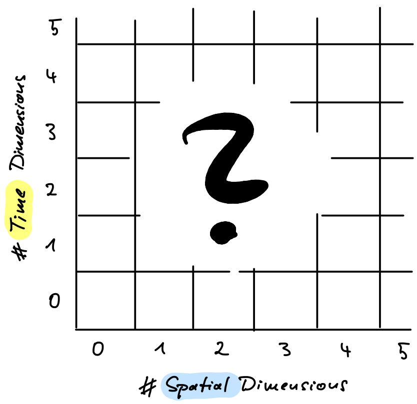
$$\partial^2 B_i=\tfrac{1}{c^2}\partial_t^2 B_i-\nabla^2 B_i=0\,. \tag{4.80b}$$

  This motivates in Eq. (4.79) the (tentative) identification of the coordinates with positive sign as "time coordinates", and the ones with a negative sign as "space coordinates":

  The difference between time and space is just a sign!

  In the following, we use the KGE as a proxy for more general relativistic field equations.

→ Possible combinations of t time and s space dimensions:



3 | Partial differential equations (PDE):

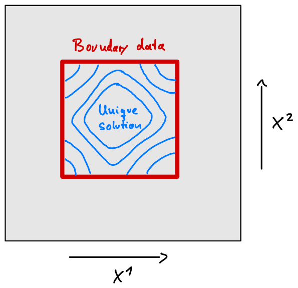
The general KGE in Eq. (4.79) is an example of a partial differential equation (PDE). The theory of PDEs has been thoroughly developed by mathematicians and a lot is known about their solvability. The problem of solving a PDE, given some boundary/initial conditions, is known as ↑ *Cauchy problem*:

- ⁂ *Well-posed (Cauchy) problem*: Given some boundary/initial data, there exists a *unique* solution to the PDE that satisfies these conditions, and this solution is *robust*. Here "robust" means that if you slightly modify the boundary/initial conditions, the solution also changes only slightly. Put differently: The solutions are not *chaotic* and you can use them to extrapolate reliably from boundary/initial states with finite errorbars. This is a crucial feature to use PDEs for *predictions* in the real world.
- ⁂ *Ill-posed (Cauchy) problem*: Given some boundary/initial data, there either exist *multiple* solutions to the PDE that satisfy these conditions, or the unique solution is not robust. In both cases, the PDE cannot be used for predictions in the real world.

i | $\sphericalangle\ (t = 0, s)$ or $(t, s = 0) \to$ Eq. (4.79) = ↑ *Elliptic PDE*

This corresponds to spacetimes that are ← *Riemannian manifolds*.

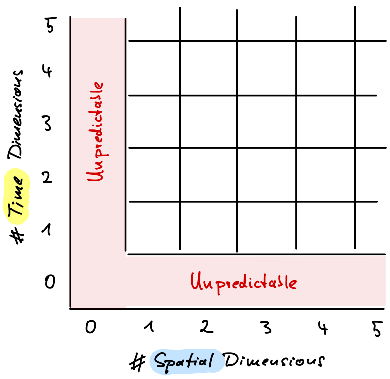
Elliptic PDEs have well-posed *boundary problems*:



→ One *cannot* use Eq. (4.79) to make predictions ☹

→ No coordinate in Eq. (4.79) qualifies as a "time coordinate".

→



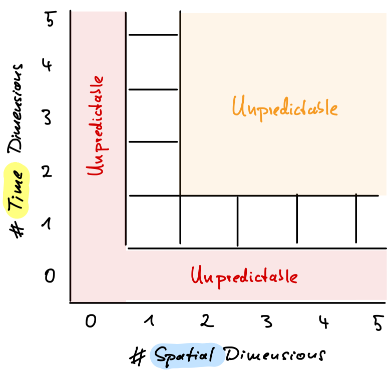
ii | $\sphericalangle\ (t \geq 2, s \geq 2)$ or $(t \geq 2, s \geq 2) \to$ Eq. (4.79) = ↑ *Ultrahyperbolic PDE*

This corresponds to spacetimes that are generic ← *pseudo-Riemannian manifolds*.

A similar but more involved chain of arguments holds also for ultrahyperbolic PDEs [39, 62].

→ One *cannot* use Eq. (4.79) to make predictions ☹

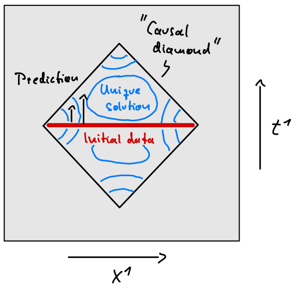
→



iii | $(t = 1, s \geq 1)$ or $(t \geq 1, s = 1) \to$ Eq. (4.79) = ↑ *Hyperbolic PDE*

This corresponds to spacetimes that are ← *Lorentzian manifolds*.

Hyperbolic PDEs have well-posed *initial value problems*:



→ We *can* use Eq. (4.79) to make predictions ☺

4 | Stability:

- ∢ Newtonian Gravity in $s \geq 4$ spatial dimensions:

  → ↓ *Two-body problem* has no stable orbits (only scattering and attraction solutions).

  → No stable planetary systems possible ☹

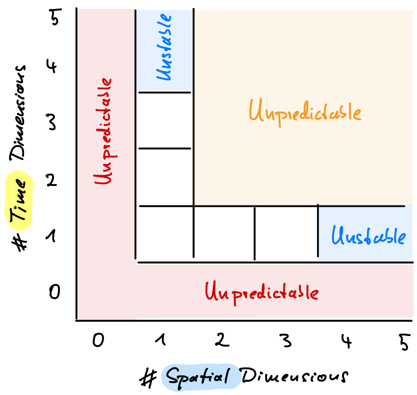
- ∢ Hydrogen atom in $s \geq 4$ spatial dimensions:

  → Schrödinger equation has no bound states.

  → No stable atoms possible ☹

The opposite cases with $t \geq 4$ and $s = 1$ are equivalent if one interprets space as time and vice versa (which is necessary to use hyperbolic PDEs to predict "the future", → *below*).

→



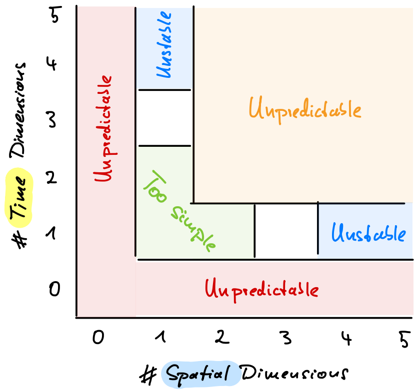
5 | Simplicity:

GENERAL RELATIVITY in $s \leq 2$ spatial dimensions → No gravity (→ *later*)!

→ No stars, no planets, no orbits ☹

The opposite cases with $t \leq 2$ and $s = 1$ are equivalent if one interprets space as time and vice versa (which is necessary to use hyperbolic PDEs to predict "the future", → *below*).

→



6 | Tachyon world:

¡! In the literature both Lorentzian signatures $(1,3)$ and $(3,1)$ are used to formulate SPECIAL RELATIVITY. Formulations in signature $(3,1)$ have nothing to do with the Tachyon sector discussed here since they compensate for the global minus in their equations. For example, the KGE in signature $(-,+,+,+,)$ reads $(-\partial^2+m^2)\Phi=0$ which is equivalent to the KGE $(\partial^2+m^2)\Phi=0$ in signature $(+,-,-,-)$. The point here is that we do *not* add this additional minus:

$$\text{Eq. (4.79)} \xrightarrow[\text{Time}\leftrightarrow\text{Space}]{(1,3)\mapsto(3,1)} (-\partial^2+m^2)\Phi=0 \;\Leftrightarrow\; (\partial^2-m^2)\Phi=0 \tag{4.81}$$

In more detail:

For $t=3$ and $s=1$ the KGE reads

$$\underbrace{\frac{\partial^2\Phi}{\partial(x^1)^2}+\frac{\partial^2\Phi}{\partial(x^2)^2}+\frac{\partial^2\Phi}{\partial(x^3)^2}}_{3\times\text{Time (?)}}-\underbrace{\frac{\partial^2\Phi}{\partial(x^4)^2}}_{1\times\text{Space (?)}}+m^2\Phi=0\,. \tag{4.82}$$

But because this an hyperbolic PDE, the Cauchy problem is only well-posed with initial conditions on a hypersurface spanned by $\{x^1,x^2,x^3\}$. Put differently: The PDE allows predictions only in x^4-direction! Thus we should interpret x^4 as *time* and $\{x^1,x^2,x^3\}$ as *space*:

$$\underbrace{\frac{\partial^2\Phi}{\partial(x^1)^2}+\frac{\partial^2\Phi}{\partial(x^2)^2}+\frac{\partial^2\Phi}{\partial(x^3)^2}}_{3\times\text{Space (!)}}-\frac{1}{c^2}\underbrace{\frac{\partial^2\Phi}{\partial t^2}}_{1\times\text{Time (!)}}+m^2\Phi=0 \tag{4.83}$$

with $ct\equiv x^4$. But this KGE is equivalent to

$$\left(\frac{1}{c^2}\partial_t^2-\nabla^2-m^2\right)\Phi=(\partial^2-m^2)\Phi=0\,. \tag{4.84}$$

Thus the "transposed" situation $(t\geq 1, s=1)$ is equivalent to the situation $(t=1, s\geq 1)$ with *negative* square-masses in the equations. Fields with negative square-mass (equivalently: imaginary mass) are called ↑ *tachyonic fields* or ↑ *tachyons* for short.

→ All massive particles are ↑ *tachyons* [63]

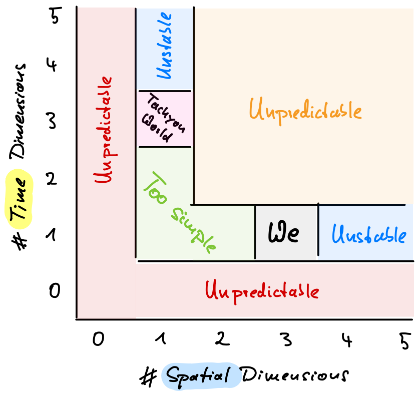
¡! Tachyonic fields are not science fiction; they do exist (→ *below*) but, contrary to the features assigned to them in science fiction, do *not* allow for faster-than-light propagation of information.

→ Tachyon fields herald vacuum instabilities [64] ☹

The spontaneous symmetry breaking of the ↑ *Higgs mechanism* is an example of this phenomenon: The Higgs field has a negative square-mass which is responsible for the "Mexican hat potential."

The consequence is spontaneous symmetry breaking, which, in this context, can be reframed as "tachyon condensation." On the new, symmetry broken vacuum, excitations are *not* tachyons with negative square-mass but Higgs bosons with positive square-mass.

→



7 | These arguments support the following *hypothesis*:

> Only a spacetime with **1 time** and **3 space** dimensions supports observers like us.

What does this line of arguments explain? Well, if you would randomly construct universes by dicing the number of space and time dimensions, only the ones with *one time* and *three space* dimensions have the chance to develop complex observers like us (who then wonder why their universe is $3+1$-dimensional). Thus the arguments above are important for "ensemble interpretations" of reality, like certain ↑ *multiverse hypotheses* or superstring theories (which can predict a plethora of different spacetime dimensions) [39, 65].