D SPECIAL RELATIVITY

D.1 *Lorentz-transforms*

The relativity principle stipulates that the laws of Nature and the constants of Nature should be the same in all frames, or in other words: There is no preferred frame in which the laws of Nature should be formulated. Space, or spacetime is homogeneous as there neither a particular location nor a particular instant in time for the formulation of laws of Nature, and the transition between one coordinate choice and the next one should be a linear, affine function: Any nonlinearity would single out a particular location or instant, breaking homogeneity. In short, the transition between frames S and S′, with their associated coordinates x^μ and x'^μ,

$$\mathrm{S} : x^\mu = \begin{pmatrix} t \\ x^i \end{pmatrix} \quad \rightarrow \quad \mathrm{S}' : x^\mu = \begin{pmatrix} t' \\ x'^i \end{pmatrix} \tag{D.252}$$

is necessarily an affine transformation.

There is a very good physical argument why this needs to be the case: Imagine now that an observer with a clock moves through spacetime on a trajectory with coordinates $x^\mu\tau$ as seen by S, and coordinates $x'^\mu(\tau)$ as seen by S′, where the parameter τ by which the trajectory is parameterised, is the proper time of the observer - the time displayed on her or his wrist watch. For an inertial trajectory, where all accelerations are zero, the velocity $\upsilon^i = \mathrm{d}x^i/\mathrm{d}\tau$ is constant, as well as the size of the time intervals $\mathrm{d}t/\mathrm{d}\tau$. In summary,

$$\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} = \text{const} \quad \text{and} \quad \frac{\mathrm{d}^2x^\mu}{\mathrm{d}\tau^2} = 0 \tag{D.253}$$

But that statement needs to be true within the frame S′ just as well:

$$\frac{\mathrm{d}x'^\mu}{\mathrm{d}\tau} = \text{const} \quad \text{and} \quad \frac{\mathrm{d}^2x'^\mu}{\mathrm{d}\tau^2} = 0 \tag{D.254}$$

Coordinate transforms can be written as an invertible, and differentiable functional relationship between the coordinate sets, i.e. in the form $x'(x)$. In this case, the velocity in the new coordinate choice becomes

$$\frac{\mathrm{d}x'^\mu}{\mathrm{d}\tau} = \frac{\partial x'^\mu}{\partial x^\nu}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} \tag{D.255}$$

with a Jacobian $\partial x'^\mu/\partial x^\nu$ mediating the coordinate change. The acceleration though acquires two terms, as both the Jacobian as well as the velocity could change with τ, albeit indirectly through the trajectory $x^\mu(\tau)$:

$$\frac{\mathrm{d}^2x'^\mu}{\mathrm{d}\tau^2} = \frac{\partial x'^\mu}{\partial x^\nu}\frac{\mathrm{d}^2x^\nu}{\mathrm{d}\tau^2} + \underbrace{\frac{\partial^2 x'^\mu}{\partial x^\nu \partial x^\rho}}_{\mathrm{A}^\mu{}_{\nu\rho}}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau}\frac{\mathrm{d}x^\rho}{\mathrm{d}\tau} \tag{D.256}$$

Only if the term $\mathrm{A}^\mu{}_{\nu\rho}$ is equal to zero, one can conclude from $\mathrm{d}^2x^\mu/\mathrm{d}\tau^2 = 0$ that $\mathrm{d}^2x'^\mu/\mathrm{d}\tau^2 = 0$. But then, the transformation between the two coordinate frames is linear:

$$\frac{\partial^2 x'^\mu}{\partial x^\nu \partial x^\rho} = 0 \quad \rightarrow \quad \frac{\partial x'^\mu}{\partial x^\nu} = \Lambda^\mu{}_\nu \quad \rightarrow \quad x'^\mu = \Lambda^\mu{}_\nu x^\mu + b^\mu \tag{D.257}$$

with integration constants $\Lambda^\mu{}_\nu$ and b^μ.

To be more specific one needs some empirical, physical input: Let's assume that the two frames S and S′ move at a constant relative speed υ. Without loss of generality, the two frames should be oriented in the same direction and the relative displacement should be along the x-axis of the coordinate frame, and the two frames should coincide in their origins at $t = t' = 0$. Then, the origin of S has the coordinate $x' = -\upsilon t'$ seen from S′, whereas the origin of S′ is at $x = +\upsilon t$ from the point of view of S.

Linearity of the transforms commands that $x' = ax + bt$ with two constant coefficients a and b, that can be functions of υ. Because $x = \upsilon t$ implies $x' = 0$, one can write: $x' = 0 = a\upsilon t + bt = (a\upsilon + b)t$, from which follows that $b = -a\upsilon$ and therefore $x' = a(x - \upsilon t)$. Reversing the roles of S and S′ then requires from $x = ax' + bt'$ that $x' = -\upsilon t$ if $x = 0$ should hold, implying $x = 0 = -a\upsilon x' + bt' = (-a\upsilon + b)t'$, and consequently $b = +a\upsilon$ and $x = a(x' + \upsilon t')$. The symmetry of the transform has effectively reduced the number of free parameters from two to a single one.

At this point Nature can make a choice. Most straightforwardly, she might choose the time to be universal, $t = t'$, and humans thought this was the case until 1905. $x' = a(x - \upsilon t)$ and $x = a(x' + \upsilon t)$ can only be compatible if $a = 1$, leading us straight to the Galilei-transforms. Or, the speed of light could be the same in all frames, $c = c'$, with $x = ct$ in S and $x' = ct'$ in S′, as the distance a light signal covers in the two respective frames. Then,

$$\begin{cases} ct &= a(ct' + \upsilon t') = a(c + \upsilon)t' \\ ct' &= a(ct - \upsilon t) = a(c - \upsilon)t \end{cases} \tag{D.258}$$

Multiplying both equations leads to $c^2 tt' = a^2(c + \upsilon)(c - \upsilon)tt'$, such that

$$a \equiv \gamma = \frac{1}{\sqrt{1 - \beta^2}} \quad \text{with} \quad \beta = \frac{\upsilon}{c}. \tag{D.259}$$

The quantity γ is known as the Lorentz-factor, and by convenience one works with a dimensionless velocity $\beta = \upsilon/c$.

A Taylor-expansion for small velocities β, $|\beta| \ll 1$, or $|\upsilon| \ll c$ of the Lorentz-factor yields

$$\gamma = 1 + \frac{d\gamma}{d\beta}\Big|_{\beta=0} \beta + \frac{d^2\gamma}{d\beta^2}\Big|_{\beta=0} \frac{\beta^2}{2} + \cdots = 1 + \frac{\beta^2}{2} + \cdots \tag{D.260}$$

showing that the Lorentz-factor depends to lowest order quadratically on the velocity before diverging as β approaches unity.

The definition of $\beta = \upsilon/c$ allows a more consistent notation for Lorentz transforms: ct as a time coordinate is then measured in units of length, just as x, there is no ambiguity as c has by virtue of the relativity principle the same value in all frames. The term υt in the Lorentz transform becomes βct, leading to

$$\begin{cases} ct' &= \gamma(ct + \beta x) \\ x' &= \gamma(x + \beta ct) \end{cases} \tag{D.261}$$

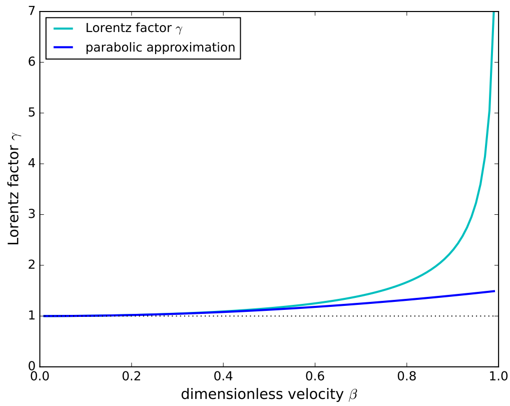

Figure 9: Lorentz γ-factor as a function of dimensionless velocity β, and the parabolic approximation for small β.

Alternatively, the transformation reads in matrix notation,

$$\underbrace{\begin{pmatrix} ct' \\ x' \end{pmatrix}}_{x'^{\mu}} = \underbrace{\begin{pmatrix} \gamma & \beta\gamma \\ \beta\gamma & \gamma \end{pmatrix}}_{\Lambda^{\mu}{}_{\nu}} \underbrace{\begin{pmatrix} ct \\ x \end{pmatrix}}_{x^{\nu}} \tag{D.262}$$

with a clearly common transformation of the ct and x coordinates, that are now combined into a single vector x^{μ}, following the transformation law $x^{\mu} \rightarrow x'^{\mu} = \Lambda^{\mu}{}_{\nu}x^{\nu}$. For small velocities, $\gamma \simeq 1$ and one obtains

$$\begin{pmatrix} ct' \\ x' \end{pmatrix} = \begin{pmatrix} 1 & \beta \\ \beta & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix} \tag{D.263}$$

With either positive or negative off-diagonal elements it is clear that a coordinate frame undergoes a shearing under Lorentz transforms, in contrast to antisymmetric transformation matrices in the case of rotations. Quantitatively for small velocities $\upsilon \ll c$ the relation reduces to $t' = t$ (neglecting $\beta x = \upsilon x/c$ for $\upsilon \ll c$) and $x' = \upsilon t + x$ in recovery of the Galilei transform.

D.2 *Lorentz-invariants*

While the coordinates depend on a chosen frame and undergo a joint change under Lorentz tranforms, one might wonder whether there are quantities that remain constant and offer the possibility to say something true for a system that would not depend on the choice of frame. Clearly, rotations leave the length of a vector, defined as its norm $r^2 = \delta_{ij}x^i x^j$ unchanged, and in this vein one can construct the quantity

$$(ct')^2 - (x')^2 = \gamma^2\left((ct)^2 - 2ct\beta x + \beta^2 x^2 - x^2 + 2ct\beta x - \beta^2(ct)^2\right) = \underbrace{\gamma^2(1-\beta^2)}_{=1}\left((ct)^2 - x^2\right) \tag{D.264}$$

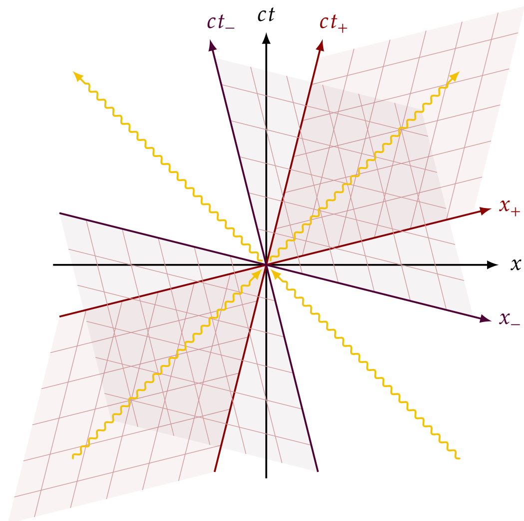

Figure 10: Spacetime diagrams under Lorentz transform for positive velocities (ct_+, x_+) and negative velocities (ct_-, x_-) relative to the frame (ct, x). Reproduction with kind permission of I. Neutelings.

which remains in fact constant under Lorentz transforms. In order to write the invariant quantity $s^2 = (ct)^2 - \delta_{ij}x^i x^j$, extended to three spatial dimensions, one introduces the Minkowski-metric,

$$s^2 = \eta_{\mu\nu}x^\nu x^\nu \quad \text{with} \quad \eta_{\mu\nu} = \begin{pmatrix} +1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} \tag{D.265}$$

which combines the Euclidean scalar product $r^2 = \gamma_{ij}x^i x^j$ mediated by by the Euclidean metric γ_{ij} to the new invariant $s^2 = \eta_{\mu\nu}x^\mu x^\nu$, as soon as Lorentz boosts are involved.

D.3 *Rapidity*

Rotations of the coordinate frame can be written in terms of a rotation matrix,

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \tag{D.266}$$

which begs the question whether (i) a similar parameterisation of the group of Lorentz transforms is possible, and if yes, (ii) which parameter ψ would replace the rotation angle α. A Lorentz-boost, written in matrix notation, would be

$$\begin{pmatrix} ct' \\ x' \end{pmatrix} = \begin{pmatrix} \gamma & \beta\gamma \\ \beta\gamma & \gamma \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix} \tag{D.267}$$

where a notable structural difference is of course the different sign in the lower left corner. But clearly, we are not looking for a rotation, as the Lorentz invariant s^2 differs from the invariant r^2! The values of the entries of the matrix are $1 \leq \gamma < +\infty$ as well as $-\infty < \beta\gamma < +\infty$, which an additional symmetry of γ for positive and negative velocities, and a sign change of $\beta\gamma$. With a bit of intuition, one might be tempted to use the hyperbolic functions to set $\gamma = \cosh\psi$ and $\beta\gamma = \sinh\psi$ (compare Fig. 11) with the so-called rapidity ψ,

$$\tanh\psi = \frac{\sinh\psi}{\cosh\psi} = \frac{\beta\gamma}{\gamma} = \beta \quad \rightarrow \quad \psi = \operatorname{artanh}\beta = \frac{1}{2}\ln\frac{1+\beta}{1-\beta}. \tag{D.268}$$

where the inverse hyperbolic tangent has a surprising representation in terms of elementary functions. More accurately, one might use the relation $\gamma^2(1+\beta^2) = 1$ to verify that

$$\gamma^2(1-\beta^2) = \gamma^2 - \gamma^2\beta^2 = \cosh^2\psi - \sinh^2\psi = 1 \tag{D.269}$$

as the defining characteristic of the hyperbolic functions.

The rapidity ψ diverges as $\beta \rightarrow 1$ and keeps, due to the antisymmetry of the hyperbolic sine, information about the direction of the boost velocity. With the rapidity as a parameter, the Lorentz-boost can be written as a hyperbolic ”rotation”,

$$\begin{pmatrix} ct' \\ x' \end{pmatrix} = \begin{pmatrix} \cosh\psi & \sinh\psi \\ \sinh\psi & \cosh\psi \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix}. \tag{D.270}$$

Then, the invariant $s^2 = (ct)^2 - x^2$ is unchanged because $\cosh^2\psi - \sinh^2\psi = \gamma^2 - \beta^2\gamma^2 = \gamma^2(1-\beta^2) = 1$, just as the invariant $r^2 = x^2 + y^2$ is unchanged because of $\cos^2\alpha + \sin^2\alpha = 1$. For a more geometric intuition, one can imagine that any point (ct, x) follows a hyperbola, purely in the timelike region for a positive norm or in the spacelike region of the spacetime diagram in the case of a negative norm. Along these hyperbolae, the norm is strictly conserved. Taking things to extremes would be a point with a lightlike norm $s^2 = 0$, which moves along the diagonals of the spacetime diagram.

D.4 *Spacetime symmetries*

A notion of spacetime was established fusion of the spatial and temporal coordinates into a coordinate tuple x^μ and the extension of the Euclidean scalar product $x_i y^i = \gamma_{ij} x^i y^j$ to the Minkowski scalar product $x_\mu y^\mu = \eta_{\mu\nu} x^\mu y^\nu$. Lorentz-transforms and rotations act on these coordinate tuples, $x^\mu \rightarrow \Lambda^\mu{}_\alpha x^\alpha$ and $x^i \rightarrow \mathrm{R}^i{}_a x^a$, respectively, leaving the scalar products invariant, $\eta_{\mu\nu} x^\mu x^\nu \rightarrow \eta_{\mu\nu} \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta x^\alpha x^\beta = \eta_{\alpha\beta} x^\alpha x^\beta$ and $\gamma_{ij} x^i x^j \rightarrow \gamma_{ij} \mathrm{R}^i{}_a \mathrm{R}^j{}_b x^a x^b = \gamma_{ab} x^a x^b$, expressed in coordinates $s^2 = \eta_{\mu\nu} x^\mu x^\nu = (ct)^2 - x^2 - y^2 - z^2$ and $r^2 = \gamma_{ij} x^i x^j = x^2 + y^2 + z^2$.

Clearly, the Lorentz-transforms as well as the rotations form groups: Successive transforms can be summarised into a single transform, for each transform there is an inverse (boosting with the negative velocity and rotating by a negative angle), and the neutral element is part of each group (corresponding to a boost with velocity zero or a rotation by an angle of zero). But there seems to be a peculiarity: The groups contain uncountably many elements and are parameterised by a continuous, real valued parameter (rapidity ψ or rotation angle α). As such, they are examples of Lie-groups.

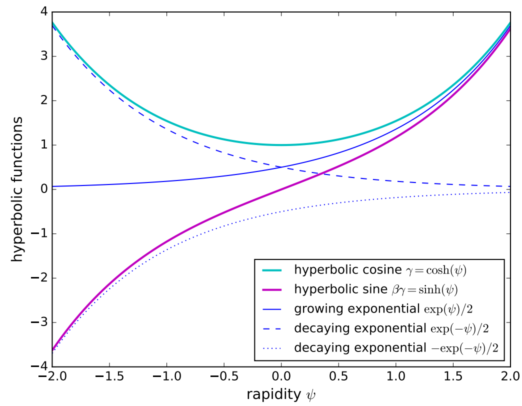

Figure 11: Hyperbolic functions $\gamma = \cosh(\psi)$ and $\beta\gamma = \sinh(\psi)$ with exponentials as their asymptotics, as a function of the rapidity ψ.

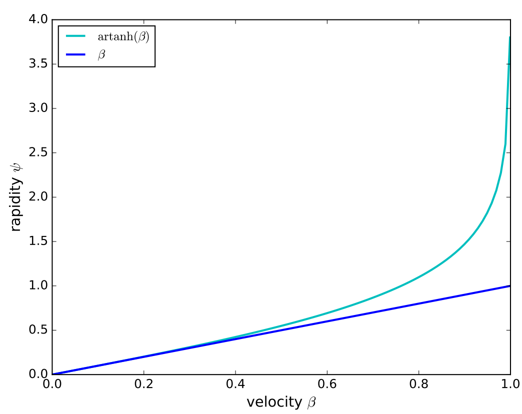

Figure 12: Rapidity ψ as a function of velocity β with the clear divergence as $\beta \to 1$.

Because of the real-valued parameter, one can actually perform a differentiation of the group element with respect to that parameter, consider an infinitesimal transform and assemble all possible group elements from this infinitesimal transform as a building block:

In the case of rotations in 2 dimensions by a small angle α one could expand the rotation matrix $R^i{}_a(\alpha)$ into a Taylor-series,

$$R^i{}_a(\alpha) = R^i{}_a\big|_{\alpha=0} + \frac{d}{d\alpha}R^i{}_a\big|_{\alpha=0}\,\alpha = \begin{pmatrix} 1 & \alpha \\ -\alpha & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \alpha\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \sigma^{(0)} + \alpha\sigma^{(2)}. \tag{D.271}$$

Such a construction with two of the Pauli-matrices

$$\sigma^{(0)} = \begin{pmatrix} +1 & 0 \\ 0 & +1 \end{pmatrix} \quad \text{and} \quad \sigma^{(2)} = \begin{pmatrix} 0 & +1 \\ -1 & 0 \end{pmatrix} \tag{D.272}$$

for an infinitesimally small angle suggest that any finite rotation by an angle α should be composable from n rotations by α/n in the limit $n \to \infty$:

$$R^i{}_a(\alpha) = \lim_{n\to\infty}\left(\sigma^{(0)} + \frac{\alpha}{n}\sigma^{(2)}\right)^n = \exp\left(\alpha\sigma^{(2)}\right) \tag{D.273}$$

where the matrix-valued exponential function is explained in terms of its series,

$$\begin{aligned} R^i{}_a = \exp\left(\alpha\sigma^{(2)}\right) &= \sum_n \frac{\alpha^n}{n!}\left(\sigma^{(2)}\right)^n = \sigma^{(0)}\sum_n \frac{\alpha^{2n}}{(2n)!}(-1)^n + \sigma^{(2)}\sum_n \frac{\alpha^{2n+1}}{(2n+1)!}(-1)^n \\ &= \sigma^{(0)}\cos\alpha + \sigma^{(2)}\sin\alpha = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \end{aligned} \tag{D.274}$$

which is the reason why $\sigma^{(2)}$ is referred to as the generator of all rotations, or equivalently, as the basis of the rotations as a Lie-group.

The same line of reasoning applies to Lorentz-transforms: They form likewise a Lie-group, parameterised by the rapidity ψ,

$$\begin{aligned} \Lambda(\psi) = \exp(\psi\sigma^{(3)}) &= \sum_n \frac{\psi^n}{n!}\left(\sigma^{(3)}\right)^n = \sigma^{(0)}\sum_n \frac{\psi^{2n}}{(2n)!} + \sigma^{(3)}\sum_n \frac{\psi^{2n+1}}{(2n+1)!} \\ &= \sigma^{(0)}\cosh\psi + \sigma^{(3)}\sinh\psi = \begin{pmatrix} \cosh\psi & \sinh\psi \\ \sinh\psi & \cosh\psi \end{pmatrix} \end{aligned} \tag{D.275}$$

where the Pauli-matrix $\sigma^{(3)}$,

$$\sigma^{(3)} = \begin{pmatrix} 0 & +1 \\ +1 & 0 \end{pmatrix} \tag{D.276}$$

can now be identified as the generator of the Lorentz-transforms. Comparing to the rotations one notices that the powers of $\sigma^{(3)}$ do not show changes in sign, but alternate between $\sigma^{(3)}$ for odd and $\sigma^{(0)}$ for even powers of n.

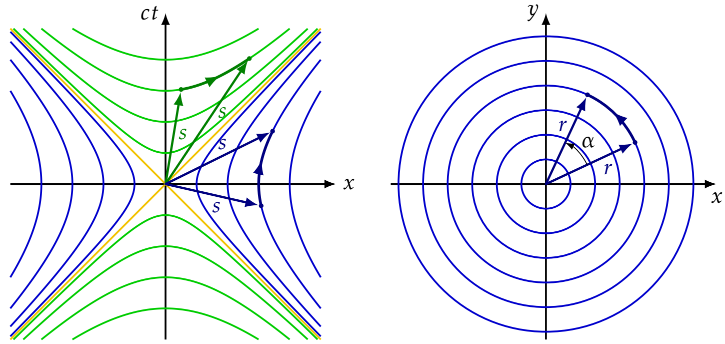

Figure 13: The rapidity ψ corresponds to the arc length that is covered by the end point of a vector with unit Minkowski norm $s^2 = (ct)^2 - x^2 = \pm 1$ under Lorentz transforms, in the same way as the rotation angle α is the arc length (or radian) covered by a point with unit Euclidean norm $r^2 = x^2 + y^2 = 1$ under a rotation. Reproduction with kind permission of I. Neutelings.

D.5 *Lorentz-group as a Lie-group*

It is intuitively clear that rotations form a group as subsequent rotations can be combined into a single rotations, and likewise, combinations of Lorentz transforms are Lorentz transforms again, Mathematically speaking, this is expressed by the group structure that is defined by the axioms: closedness, the existence of a unit element, the existence of an inverse element and lastly associativity.

For the closedness of a group one needs to show that the combination of group elements is again a group element. In a Lie-group, where the elements are generated by means of an exponential, one gets for instance for rotations

$$R(\alpha)R(\beta) = \exp\left(\alpha\sigma^{(2)}\right)\exp\left(\beta\sigma^{(2)}\right) = \left(\sum_i \frac{\alpha^i}{i!}\left(\sigma^{(2)}\right)^i\right)\left(\sum_j \frac{\beta^j}{j!}\left(\sigma^{(2)}\right)^j\right). \tag{D.277}$$

Multiplying the two exponential series can be achieved by application of the Cauchy-product

$$= \sum_i \sum_j^i \frac{\alpha^j}{j!}\frac{\beta^{i-j}}{(i-j)!}\left(\sigma^{(2)}\right)^j\left(\sigma^{(2)}\right)^{i-j} = \sum_i \frac{1}{i!}\left(\sum_j \binom{i}{j}\alpha^j\beta^{i-j}\right)\left(\sigma^{(2)}\right)^i \tag{D.278}$$

by using the definition of the binomial coefficient

$$\binom{i}{j} = \frac{i!}{j!(i-j)!}, \tag{D.279}$$

which leads to

$$= \sum_i \frac{(\alpha+\beta)^i}{i!}\left(\sigma^{(2)}\right)^i = R(\alpha+\beta) \tag{D.280}$$

by virtue of the generalised binomial formula,

$$(\alpha+\beta)^i = \sum_j \binom{i}{j}\alpha^j\beta^{i-j}, \tag{D.281}$$

which confirms the intuitive expectation that combining two rotations leads to a rotation again. In complete analogy one can show that $\Lambda(\psi)\Lambda(\varphi) = \Lambda(\psi+\varphi)$ for Lorentz transforms, with the rapidity as an additive parameter.

The unit element, which leaves a vector unchanged, is the quite obviously obtained for a rotation by the angle zero or a boost by zero rapidity:

$$R(\alpha=0) = \exp(0\times\sigma^{(2)}) = \exp(\sigma^{(2)})^0 = \mathrm{id} \tag{D.282}$$

Alternatively, one might argue that

$$R(\alpha=0) = \sigma^{(0)}\cos(0) + \sigma^{(3)}\sin(0) = \sigma^{(0)} \equiv \mathrm{id} \tag{D.283}$$

and likewise obtain the unit matrix.

Associativity is very obvious for Lie-groups as their additive parameters naturally obey associativity:

$$R\left(\alpha+(\beta+\gamma)\right) = R\left((\alpha+\beta)+\gamma\right) \tag{D.284}$$

which implies

$$R\left(\alpha+(\beta+\gamma)\right) = R(\alpha)R(\beta+\gamma) = R(\alpha)\left[R(\beta)R(\gamma)\right] = \left[R(\alpha)R(\beta)\right]R(\gamma) = R(\alpha+\beta)R(\gamma) = R\left((\alpha+\beta)+\gamma\right) \tag{D.285}$$

Conservation of the norm of vectors under transformations, or equivalently, the orthogonality of the transform is realised in the following way, keeping in mind that $(\sigma^{(2)})^t = -\sigma^{(2)}$,

$$R^t(\alpha)R(\alpha) = \exp\left(\alpha\sigma^{(2)}\right)^t\exp\left(\alpha\sigma^{(2)}\right) = \exp\left(\alpha(\sigma^{(2)})^t\right)\exp\left(\alpha\sigma^{(2)}\right) = \exp\left(-\alpha\sigma^{(2)}\right)\exp\left(\alpha\sigma^{(2)}\right) = \exp\left((-\alpha+\alpha)\sigma^{(2)}\right) = \mathrm{id} \tag{D.286}$$

which differs slightly in the case of Lorentz-transforms, as they are orthogonal with respect to the the Minkowski-metric $\eta = \sigma^{(1)}$ instead of the Euclidean metric $\sigma^{(0)} = \mathrm{id}$,

$$\Lambda(\psi)^t\eta\Lambda(\psi) = \eta \quad \text{with} \quad \eta = \sigma^{(1)} = \begin{pmatrix} +1 & 0 \\ 0 & -1 \end{pmatrix} \tag{D.287}$$

Invariants of the transform such as determinants are realised in a funky way in Lie-groups: As an auxiliary result, we need that for any transform A with eigenvalues λ_i,

$$\ln\det(A) = \ln\prod_i \lambda_i = \sum_i \ln\lambda_i = \mathrm{tr}\ln(A) \tag{D.288}$$

where the matrix-valued logarithm $\ln(A)$ is naturally defined in terms of its series.

Because the logarithm can not be expanded at zero, where it is undefined, one uses this neat trick,

$$\ln(A) = \ln(\mathrm{id} + (A - \mathrm{id})) = \sum_n \frac{(-1)^{n+1}}{n}(A - \mathrm{id})^n. \tag{D.289}$$

Then,

$$\exp\ln\det(A) = \det(A) = \exp\operatorname{tr}\ln(A), \tag{D.290}$$

and with the substitution $B = \ln(A)$ one arrives at

$$\det\exp(B) = \exp\operatorname{tr}(B), \tag{D.291}$$

which is particular suitable for our purpose, as the determinant of a Lie-generated group element is related to the trace of its generator. Applied to the rotations this implies

$$\det(R) = \det\exp\left(\alpha\sigma^{(2)}\right) = \exp\operatorname{tr}\left(\alpha\sigma^{(2)}\right) = \exp\left(\alpha\operatorname{tr}\sigma^{(2)}\right) = \exp(0) = 1 \tag{D.292}$$

because the Pauli-matrix $\sigma^{(2)}$ is traceless. The same result for the Lorentz-transforms $\Lambda(\psi)$ follows in complete analogy,

$$\det(\Lambda) = \det\exp\left(\psi\sigma^{(3)}\right) = \exp\operatorname{tr}\left(\psi\sigma^{(3)}\right) = \exp\left(\psi\operatorname{tr}\sigma^{(3)}\right) = \exp(0) = 1. \tag{D.293}$$

Essentially, the determinant of the Lie-group is fixed to unity by the tracelessness of the generator.

Up to this point, we have been dealing with a single generator, but in 3+1 dimensions there might be cases where one combines rotations about different axes, boosts in different directions or even considers combinations between boosts and rotations! In these cases commutativity plays an important role, as it provides a correction factor to the rule $\exp(A)\exp(B) = \exp(A + B)$ known as the Baker-Hausdorff-Campbell formula:

$$\exp(A)\exp(B) = \exp(A + B)\exp\left(-\frac{1}{2}[A, B]\right), \tag{D.294}$$

with the commutator $[A, B] = AB - BA$.

D.6 *Adding velocities*

Subsequent Lorentz-transforms can be combined into a single transformation, and we already know that the Lorentz-transforms form a Lie-group with the rapidity ψ as an additive parameter instead of the velocity $\beta = \tanh\psi$. Luckily, there is a handy addition theorem for the hyperbolic tangent function:

$$\tanh(\psi + \varphi) = \frac{\tanh\psi + \tanh\varphi}{1 + \tanh\psi\cdot\tanh\varphi} \tag{D.295}$$

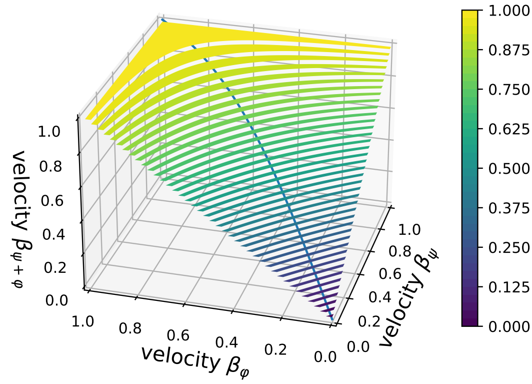

Figure 14: Relativistic addition theorem for velocities, with the particular case of $\beta_\psi = \beta_\varphi$ along the diagonal. $\beta_{\psi+\varphi}$ remains strictly below unity, excluding superluminal velocities. Clearly the relation needs to be linear if one of the velocities is zero, as seen at the edges.

Therefore, one obtains for the velocities

$$\beta_{\psi+\varphi} = \frac{\beta_\psi + \beta_\varphi}{1 + \beta_\psi \cdot \beta_\varphi} < 1 \tag{D.296}$$

leading to a combined velocity strictly smaller than the speed of light. Linearising the relationship shows a straightforward addition of velocities,

$$\beta_{\psi+\varphi} \simeq \beta_\psi + \beta_\varphi, \tag{D.297}$$

as one would expect from Galilean physics. A proof that the added velocities are strictly smaller than c might be done along these lines: Writing $\beta_\psi = 1 - x$ and $\beta_\varphi = 1 - y$ with positive x and y lead to

$$\beta_{\psi+\varphi} = \frac{(1-x) + (1-y)}{1 + (1-x)(1-y)} = \frac{2 - x - y}{2 - x - y + xy} < 1 \tag{D.298}$$

because the product xy is larger than zero.

D.7 *Relativistic effects*

There are quite a number of relativistic effects, and they all hinge on the fact that spatial and temporal coordinates change jointly under Lorentz transforms, while only invariants constructed from them are truly fixed. If one chooses to ignore that the coordinates transform jointly and only looks at a single coordinate, surprising things will happen. Invariants will have identical values in all frames and take into account all coordinates. As such, they are the means for making statements that do

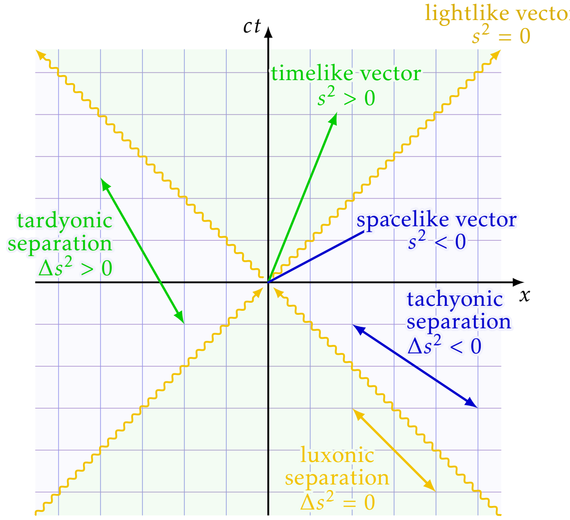

Figure 15: Classification of distances as spacelike $\Delta s^2 < 0$, timelike $\Delta s^2 > 0$ and lightlike $\Delta s^2 = 0$. Reproduction with kind permission of I. Neutelings.

not depend on a particular coordinate choice and hence transcend frames. Personally, I like skewed spacetime diagrams where the rapidities are chosen to be $\pm\psi/2$ because then the relative lengths in both frames are equal, and one can compare distances directly.

D.7.1 Constancy of the speed of light

In every frame, the speed of light comes out as constant, to the same numerical value, as illustrated by Fig. 16. This is no surprise, was it was the defining choice that differentiated Lorentz- from Galilei-transforms. In the diagram one immediately sees that a point on the diagonal, which corresponds to the light cone, acquires x- and ct-coordinates that change in proportionality to each other, indicating that their ratio is constant – the speed of light.

D.7.2 Relativity of simultaneity

Events at nonzero spatial separation, which take place at the same time (but at different positions), i.e. simultaneously on one frame, take place at different times in another frame, as shown in Fig. 17.

D.7.3 Time dilation

A time interval ct' taken at constant spatial coordinate x' gets mapped onto a time interval ct with differing spatial coordinates. The ratio between the two time intervals is proportional to the Lorentz-factor $\gamma \geq 1$. Fig. 18 shows how the time interval appears longer in projection.

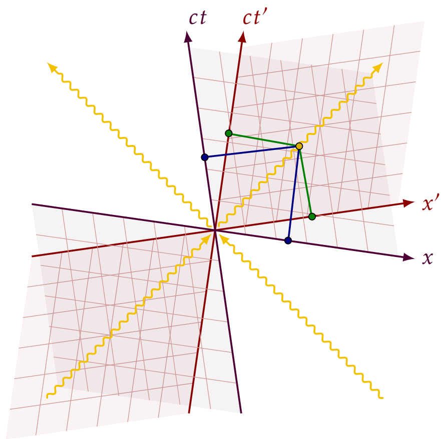

Figure 16: Constancy of the speed of light: The ratio of the two coordinates of a light-like event is always constant.

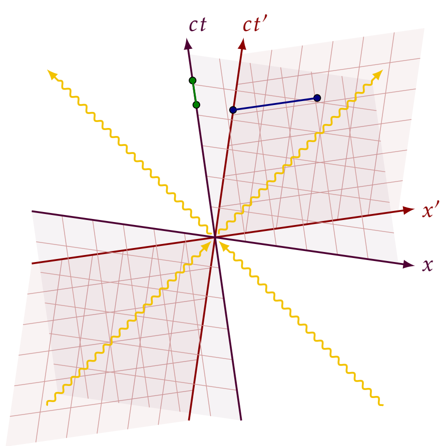

Figure 17: Relativity of simultaneity: Events that take place at the same time ct′ in one system (blue), take place at different times in another system (green).

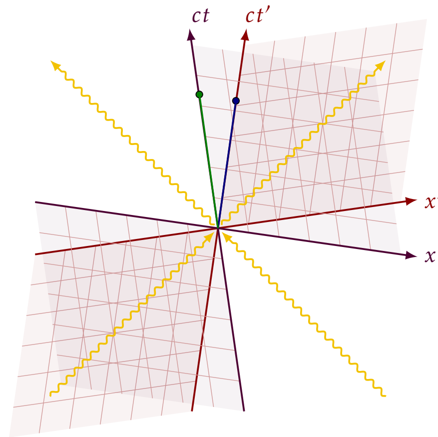

Figure 18: A duration ct' of a process in one system (blue), seems to take more time when viewed from another system (green)

D.7.4 Length contraction

An object with a given length on one frame will appear to have a shorter length as viewed from another frame. This is ultimately traced back to the relativity of simultaneity: The length of an object is defined as the distance between its ends at the same time, but in a different frame, one effectively combines coordinates at different times, as demonstrated in Fig. 19. The contraction effect is proportional to the inverse Lorentz-factor $1/\gamma \leq 1$.

D.7.5 Causal ordering inside the light cone

The temporal order of time-like events is conserved under Lorentz-transforms, lightlike-events take place simultaneously, while the order of space-like separated events depends on the frame. To formulate this in a more extreme way, there is causal ordering only inside the light cone, and no causal ordering outside the light cone, as shown in Fig. 20.

D.8 *Proper time*

If a particle moves through spacetime along a trajectory $x^\mu(\tau)$ in the sense that it passes by the coordinates x^μ as its proper time τ evolves, one can define the 4-velocity u^μ of the particle as a tangent to the trajectory

$$u^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} \tag{D.299}$$

which is consistent with the definition of infinitesimal arc length $\mathrm{d}s$ along the trajectory, as

$$\mathrm{d}s^2 = \eta_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu = \eta_{\mu\nu}u^\mu u^\nu\mathrm{d}\tau^2 = c^2\mathrm{d}\tau^2 \tag{D.300}$$

i.e. if the 4-velocity is defined with proper time τ as an affine parameter, it is normalised to $\eta_{\mu\nu}u^\mu u^\nu = c^2$, and the arc length is measurable by means of a clock.

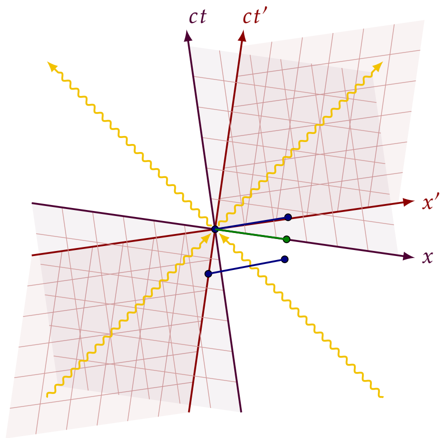

Figure 19: A yardstick at rest in the primed system (blue) seems to be contracted as viewed from another system (green).

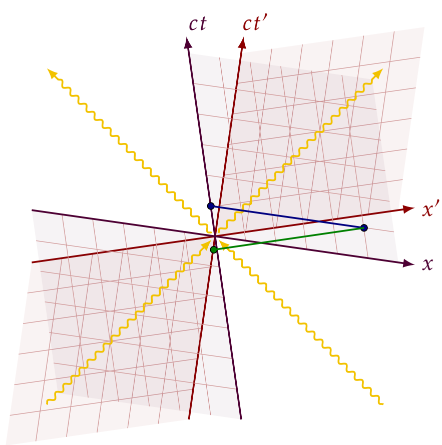

Figure 20: Spacelike separated events do not have an absolute causal ordering. The event seems to have a positive time coordinate ct (blue) and takes place after *the event at the origin, but a negative coordinate ct' (green) in the other frame and precedes the event at the origin.*

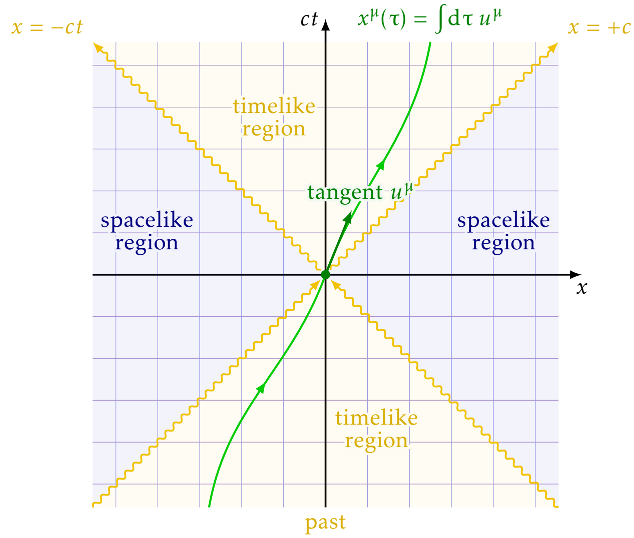

Figure 21: Spacetime diagram with spacelike (both left and right) and timelike (both past and future) regions, along with the worldline $x^\mu(\tau)$ of a massive particle, with 4-velocity $u^\mu = \mathrm{d}x^\mu/\mathrm{d}\tau$. As $\eta_{\mu\nu}u^\mu u^\nu = c^2 > 0$, the massive particle necessarily moves inside the light cone. Reproduction with kind permission of I. Neutelings.

Proper time is the time elapsing on a clock that is carried along with the particle: The infinitesimal arc length can be expressed in terms of the coordinate differentials

$$\mathrm{d}s^2 = \eta_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu = c^2\mathrm{d}t^2 - \gamma_{ij}\mathrm{d}x^i\mathrm{d}x^j = c^2\mathrm{d}\tau^2 \tag{D.301}$$

as the change of spatial coordinates $\mathrm{d}x^i$ is zero for the comoving clock. This implies three things: Proper time measures the arc length of the trajectory of a particle through spacetime,

$$s = \int_A^B \mathrm{d}s = c\int_A^B \mathrm{d}\tau = \int_A^B \sqrt{\eta_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu} \tag{D.302}$$

and is, as a Lorentz scalar, invariant under Lorentz transforms. And in addition, the normalisation of the 4-velocity is c^2 if τ is used as the affine parameter for $x^\mu(\tau)$.

Returning to the expression of s in terms of the infinitesimal coordinate changes leads to

$$s = \int_A^B \mathrm{d}s = c\int_A^B \mathrm{d}t\,\sqrt{1 - \frac{\gamma_{ij}}{c^2}\frac{\mathrm{d}x^i}{\mathrm{d}t}\frac{\mathrm{d}x^j}{\mathrm{d}t}} = c\int_A^B \mathrm{d}t\,\sqrt{1-\gamma_{ij}\beta^i\beta^j} = c\int_A^B \mathrm{d}t\,\frac{1}{\gamma} = c\int_A^B \mathrm{d}t\,\mathcal{L} \tag{D.303}$$

which can be used to compute arc lengths through spacetime.

Trajectories that have extremal values for s would result from a variational principle applied to $\mathcal{L} = 1/\gamma$. Hamilton's principle $\delta S = 0$ implies

$$\delta \int_A^B \mathrm{d}t\, \mathcal{L}(x^i, \upsilon^i) = \int_A^B \mathrm{d}t \left(\frac{\partial \mathcal{L}}{\partial x^i}\delta x^i + \frac{\partial \mathcal{L}}{\partial \upsilon^i}\delta \upsilon^i\right) = 0 \tag{D.304}$$

with the typical replacement

$$\delta \upsilon^i = \delta \frac{\mathrm{d}x^i}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\delta x^i \tag{D.305}$$

which enables integration by parts, yielding the Euler-Lagrange equation

$$\delta S = \int_A^B \mathrm{d}t \left(\frac{\partial \mathcal{L}}{\partial x^i} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \upsilon^i}\right) = 0 \quad \rightarrow \quad \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \upsilon^i} = \frac{\partial \mathcal{L}}{\partial x^i} \tag{D.306}$$

The identical calculation can be done if the velocities are 4-velocities, expressed in terms of proper time τ

$$\delta \int_A^B \mathrm{d}\tau\, \mathcal{L}(x^\mu, u^\mu) = \int_A^B \mathrm{d}\tau \left(\frac{\partial \mathcal{L}}{\partial x^\mu}\delta x^\mu + \frac{\partial \mathcal{L}}{\partial u^\mu}\delta u^\mu\right) = 0 \tag{D.307}$$

with the typical replacement

$$\delta u^\mu = \delta \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} = \frac{\mathrm{d}}{\mathrm{d}\tau}\delta x^\mu \tag{D.308}$$

which enables integration by parts, yielding the Euler-Lagrange equation

$$\delta s = \int_A^B \mathrm{d}\tau \left(\frac{\partial \mathcal{L}}{\partial x^\mu} - \frac{\mathrm{d}}{\mathrm{d}\tau}\frac{\partial \mathcal{L}}{\partial u^\mu}\right) = 0 \quad \rightarrow \quad \frac{\mathrm{d}}{\mathrm{d}\tau}\frac{\partial \mathcal{L}}{\partial u^\mu} = \frac{\partial \mathcal{L}}{\partial x^\mu} \tag{D.309}$$

for motion through 4-dimensional spacetime.

D.9 *Relativistic motion*

It would be a good idea to see if relativistic motion with the correct transformation property of all quantities involved would result naturally from a variational principle: This will be the case, and sometimes it appears to me that variational principles, always presented as the pinnacle of classical physics, are in fact relativistic: In some sense they are a piece of mathematics that has been discovered a few hundred years too early to appreciate them properly. They incorporate the idea that an invariant (under coordinate transforms) Lagrange-function gives rise to a covariant equation of motion. To see how this works, let's start at a classical Lagrange-function $\mathcal{L}(x^i, \upsilon^i)$

$$\mathcal{L}(x^i, \upsilon^i) = \frac{\gamma_{ij}}{2}\upsilon^i \upsilon^j - \Phi(x^i) \tag{D.310}$$

where both terms are invariant under e.g. rotations, Φ is scalar anyways and $\gamma_{ij}\upsilon^i\upsilon^j$ as the norm of the vector $\dot{x}$. Hamilton's principle $\delta S = 0$ with the action

$$S = \int_{t_i}^{t_f} dt\, \mathcal{L}(x^i, \upsilon^i) \tag{D.311}$$

yields the Newtonian equation of motion

$$\ddot{x}^i = -\gamma^{ij}\partial_j\Phi \tag{D.312}$$

which sets the vector $\ddot{x}$ in relation with the gradient $\partial\Phi$, which is likewise a vector. While this is perfectly nice, there are some points of criticism for the variational principle that one can not answer from a classical point of view: There is no obvious interpretation of $\mathcal{L}$ or S, they are not measurable in a direct way and they behave funnily under Galilei transforms:

$$x^i \to x^i + u^i t \quad \text{and consequently} \quad \upsilon^i \to \upsilon^i + u^i \quad \text{for a constant relative velocity } u^i \tag{D.313}$$

This implies for the Lagrange function

$$\mathcal{L}(x^i, \upsilon^i) \to \frac{\gamma_{ij}}{2}\upsilon^i\upsilon^j + \gamma_{ij}\upsilon^i u^j + \frac{\gamma_{ij}}{2}u^i u^j = \frac{\gamma_{ij}}{2}\upsilon^i\upsilon^j + \frac{d}{dt}\left(\gamma_{ij}x^i u^j + \frac{\gamma_{ij}}{2}u^i u^j t\right) \tag{D.314}$$

In fact, the Lagrange function is not invariant under Galilei-transforms, but the additional term appearing is a total time derivative and does therefore not play a role in the variational principle. It might strike you as odd (and rightfully so), that rotations and Galilei-transforms are treated so differently.

Thinking about a relativistic Lagrange-function that should be intuitive, measurable and invariant leads to proper time

$$c\tau = c\int_A^B d\tau = \int_A^B ds = \int_A^B \sqrt{\eta_{\mu\nu}dx^\mu dx^\nu} \tag{D.315}$$

It is the time that is displayed as elapsed on a clock that is moving along with the particle and is, geometrically, the arc length of the trajectory through spacetime, measured with the Minkowski-metric $\eta_{\mu\nu}$. As this metric defines an invariant, the arc-length $c\tau = s$ will be identical in any Lorentz frame, and it will be a convex functional in the velocity $\upsilon = c\beta$, making sure that the variational principle finds a uniquely defined minimum and enabling Legendre-transforms to find the associated energy. As affine transformations $\mathcal{L} \to a\mathcal{L} + b$ of the Lagrange-function or the action do not have any influence on the Euler-Lagrange-equation, we can include a prefactor $-mc$ to yield

$$S = -mc^2\int_A^B d\tau = -mc\int_A^B ds = -mc\int_A^B \frac{dt}{\gamma} \quad \to \quad \mathcal{L} = -\frac{mc}{\gamma} \tag{D.316}$$

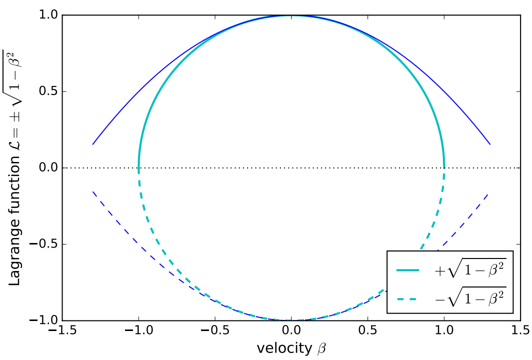

Figure 22: Relativistic Lagrange function $\mathcal{L}(\beta) = \pm\sqrt{1-\beta^2}$ in comparison to its classical limits $\mathcal{L}(\beta) = \pm 1 \mp \beta^2/2$.

where the difference between the arc length s and the action S has vanished, or in other words: We've found a geometric interpretation of the action.

It is very instructive to reformulate time proper time integral in terms of the 4-velocity u^μ,

$$u^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} = \frac{\mathrm{d}t}{\mathrm{d}\tau}\frac{\mathrm{d}x^\mu}{\mathrm{d}t} = \gamma\begin{pmatrix} c \\ \upsilon^i \end{pmatrix} \quad \text{for} \quad x^\mu = \begin{pmatrix} ct \\ x^i \end{pmatrix} \tag{D.317}$$

with the definition of the conventional velocity as $\upsilon^i = \mathrm{d}x^i/\mathrm{d}t$. Then,

$$\mathrm{d}s^2 = \eta_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu = \eta_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau}\mathrm{d}\tau^2 = \gamma^2\left(c^2 - \upsilon_i\upsilon^i\right)\mathrm{d}\tau^2 = c^2\underbrace{\gamma^2(1-\beta^2)}_{=1}\mathrm{d}\tau^2 = c^2\mathrm{d}\tau^2 \tag{D.318}$$

and the normalisation of the 4-velocity is timelike, $\eta_{\mu\nu}u^\mu u^\nu = c^2 > 0$, as the particle moves necessarily inside the light cone.

D.10 *Relativistic dispersion relations*

With the relativistic Lagrange function $\mathcal{L}$ being equal to the inverse Lorentz-factor,

$$\mathcal{L} = -\frac{1}{\gamma} = -\sqrt{c^2 - \upsilon^2} \tag{D.319}$$

one can derive the canonical momentum p

$$p = \frac{\partial\mathcal{L}}{\partial\upsilon} = \frac{\upsilon}{\sqrt{c^2-\upsilon^2}} \quad \text{such that} \quad \upsilon = \frac{cp}{\sqrt{1+p^2}} \tag{D.320}$$

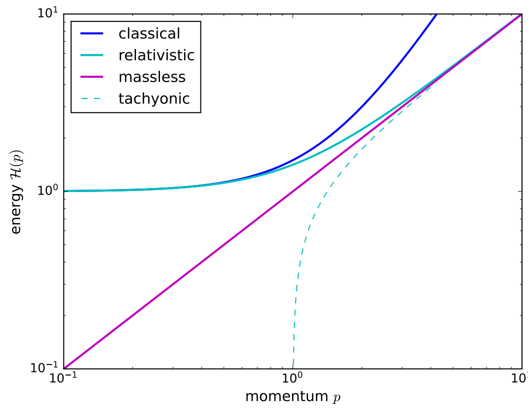

Figure 23: Relativistic dispersion relation $\mathcal{H} = \sqrt{p^2 + m^2}$ for $m = 1$, in comparison to the massless dispersion relation $\mathcal{H} = p$, the classical dispersion relation $\mathcal{H} = p^2/2 + 1$ and a tachyonic dispersion relation $\mathcal{H} = \sqrt{p^2 - 1}$, which is only defined for $p \geq 1$.

Carrying out the Legendre-transform for obtaining $\mathcal{H}$ from $\mathcal{L}$

$$\mathcal{H} = \upsilon(p)p - \mathcal{L}(\upsilon(p)) \tag{D.321}$$

then implies

$$\mathcal{H} = \upsilon\frac{\upsilon}{\sqrt{c^2 - \upsilon^2}} + \sqrt{c^2 - \upsilon^2} = \upsilon p + \frac{\upsilon}{p} = \upsilon\left(p + \frac{1}{p}\right) = c\frac{p}{\sqrt{1 + p^2}}\frac{1 + p^2}{p} = c\sqrt{1 + p^2} \tag{D.322}$$

and if one would include mc as a prefactor,

$$\mathcal{H} = \sqrt{(cp)^2 + (mc^2)^2} \tag{D.323}$$

which is exactly the relativistic dispersion relation. Surprisingly, the energy $\mathcal{H}$ is not zero even for $p = 0$, which is why we associate this energy mc^2 to the rest mass of a particle. With this dispersion relation it is straightforward to compute the group and phase velocities of a wave packet associated with a relativistic particle,

$$\upsilon_{\mathrm{gr}} = \frac{\mathrm{d}\mathcal{H}}{\mathrm{d}p} = c^2\frac{p}{\mathcal{H}} \quad \text{and} \quad \upsilon_{\mathrm{ph}} = \frac{\mathcal{H}}{p} \tag{D.324}$$

such that their geometric average is exactly c^2:

$$\upsilon_{\mathrm{gr}} \times \upsilon_{\mathrm{ph}} = c^2 \tag{D.325}$$

Because for any momentum $\mathcal{H} > cp$, it is the case that $\upsilon_{\mathrm{gr}} < c$ while $\upsilon_{\mathrm{ph}} > c$. It is reassuring to see that the group velocity, associated with the motion of massive particles, is always subluminal.

The Euler-Lagrange equation for minimising the arc-length $s = \int \mathrm{d}s$ reads

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\frac{\partial\mathcal{L}}{\partial u^\alpha} = \frac{\partial\mathcal{L}}{\partial x^\alpha} \quad \text{for} \quad \mathcal{L} = \sqrt{\eta_{\mu\nu}u^\mu u^\nu} \tag{D.326}$$

where the right side is automatically zero in this case, because $\mathcal{L}$ does not depend on x^α. Evaluating the left side gives:

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{1}{2c}\eta_{\mu\nu}\left[\frac{\partial u^\mu}{\partial u^\alpha}u^\nu + u^\mu\frac{\partial u^\nu}{\partial u^\alpha}\right]\right) = \frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{1}{2c}\eta_{\mu\nu}\left[\delta^\mu_\alpha u^\nu + u^\mu\delta^\nu_\alpha\right]\right) =$$
$$\frac{1}{2c}\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\eta_{\alpha\nu}u^\nu + \eta_{\mu\alpha}u^\mu\right) = \frac{1}{c}\frac{\mathrm{d}u_\alpha}{\mathrm{d}\tau} = 0 \tag{D.327}$$

implying that in the absence of forces, the particle moves through spacetime at a constant 4-velocity, or equivalently, that a straight line corresponds to motion free of acceleration: This is exactly the relativistic version of Newton's law of inertia. And it remains true, even in Minkowski-space, that inertial motion along a straight line minimises the arc length: The straightest trajectory is the shortest. It is quite astonishing to see the geometric picture behind Newton's axioms that is somewhat hidden in classical mechanics.

Expanding the arc length s in terms of a Taylor-expansion for small velocities

$$s = \int_A^B \mathrm{d}s = c\int_A^B \mathrm{d}t\,\sqrt{\eta_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}t}\frac{\mathrm{d}x^\nu}{\mathrm{d}t}} = c\int_A^B \mathrm{d}t\,\sqrt{c^2 - \upsilon^2} \simeq \int_A^B \mathrm{d}t\left(1 - \frac{\upsilon^2}{2}\right) \tag{D.328}$$

recovering the square of the velocity familiar from classical mechanics, in the approximation $\sqrt{1-\beta^2} \simeq 1 - \beta^2/2$ for $\beta \ll 1$. Weirdly enough, we see that it doesn't have anything to do with kinetic energies, it is just the lowest-order Taylor-expansion of the relativistic arc length and is a purely geometrical object. With the suggestive identification of the arc length as the action and the line element or proper time interval as the Lagrange function, one really falls back onto the kinetic energy as the Lagrange function of classical mechanics, because it is only ever defined up to an affine transform, negating the influence of the additive 1, and allowing to multiply the line element with the negative mass.