An introduction to special relativity

Edoardo Provenzi

Contents

1 History and philosophical motivations 4
1.1 Disclaimer: Special Relativity is *not* easy 4
1.2 The teleological view 4
1.3 A problem with a bow and arrow... 5
1.4 The modern era: Newtonian mechanics 5
1.5 Motion in space... but what *is* space? 6
1.6 Motion through time... but what *is* time? 6
1.7 Laplace's demon 7
1.8 The past hypothesis 7
1.9 Standing on the shoulders of giants 7
1.10 Galilean relativity 8
1.11 Galilean relativity upset a number of people... 9
1.12 Inertial reference frames and the principle of inertia 9
1.13 Inertial reference frames and forces 10
1.14 The mathematical postulates of Galilean relativity 10
1.15 Clock synchronization 11
1.16 Galilean transformations between inertial frames in the standard configuration 12
1.17 Homogeneous Galilean transformations between inertial frames in a generic configuration 13
1.18 Galilean velocity addition law 14
1.19 The most general Galilean transformations and the Galilei group 15
1.20 The principle of Galilean relativity as a meta-law of physics 16
1.21 Is there a limit on the velocity in Newtonian mechanics? 17
1.22 Action at a distance 17
1.23 Maxwell's equations and the speed of light 18
1.24 Luminiferous aether 19
1.25 Maxwell's equations are not Galilean invariant! 20
1.26 The wave equation is not Galilean invariant either! 21
1.27 Reactions to the lack of Galilean invariance of Maxwell's equations 22
1.28 Lorentz's local time 22
1.29 Galilean vs. Lorentz transformations between inertial frames 23
1.30 Toward a systematic theory of space and time... 24
1.31 Einstein relativity (in his 'miracle year': 1905) 24
1.32 Enters Minkowski: spacetime vs. space (1907) 25
1.33 A bit of relativity lingo: events and spacetime interval 25
1.34 Spacetime geometry can't be Euclidean 26

2 Lorentz geometry **27**
2.1 Preliminaries on real quadratic forms 27
2.2 It's all about the (pseudo) inner product. 28
2.3 Lorentz and Minkowski indefinite quadratic forms 28
2.4 Relation with the Euclidean inner product 29
2.5 Likeness and orientation of a vector in the Lorentz space 29
2.6 Likeness in the 2-Lorentz space $\mathbb{R}^{1,1}$ 30
2.7 Connection with hyperbolic geometry 30
2.8 Two quadratic forms with signature $(1, n-1)$ having same null cone are proportional 32
2.9 Lorentz-orthogonal vectors 34
2.10 Lorentz-orthogonality and time-like vectors 34
2.11 Lorentz-orthogonality and space-like vectors 35
2.12 Lorentz-orthonormality 36
2.13 Lorentz transformations 36
2.14 The Lorentzian matrix and its properties 37
2.15 Noticeable subgroups of the Lorentz group 38
2.16 The restricted Lorentz group 38

3 Relativistic kinematic **39**
3.1 Back to the 4D Minkowski space. 39
3.2 Revisiting likeness regions in the Minkowski spacetime 40
3.3 Invariance of spacetime interval in inertial reference frames 41
3.4 Lorentz transformations in special relativity 43
3.5 Minkowski diagrams (1907) 44
3.6 Lorentz boost (in the standard configuration) 47
3.7 Effect of a Lorentz boosts on the axes of the Minkowski diagrams 49
3.8 The hidden role of the Lorentz factor γ in Minkowski diagrams 51
3.9 Relating the scales on the axes of a Minkowski diagram 52
3.10 Boost along a general direction and the commutativity condition 54
3.11 Comparison between the general Lorentz boost and the general Galilean transformation 55
3.12 The most general Lorentz transformations 55
3.13 Lorentz-invariance of the wave equation: compatibility between Maxwell electromagnetism and special relativity 56
3.14 Alternative derivation of Lorentz boosts and connection with hyperbolic geometry 59
3.15 Boosts as 'hyperbolic rotations' 63
3.16 Carroll transformations 65
3.17 Relativity of simultaneity 67
3.18 Relativity of simultaneity with Minkowski diagrams 68
3.19 Preservation of causality (chronological order of events) 69
3.20 Measure of time in different inertial frames: proper and coordinate time 70
3.21 Dilation of time intervals (from the viewpoint of an observer who sees a clock moving) 70
3.22 Proper time along general time-like worldlines 72
3.23 Existence of a frame in which two time-like separated events occur at the same spatial location 73

3.24 The reverse triangle inequality for future-directed time-like vectors 74
3.25 Inertial motion maximizes proper time 75
3.26 The twin 'paradox' 77
3.27 Time dilation of the mean lifetime of cosmic muons 78
3.28 Length contraction (from the viewpoint of an observer who sees an object moving) 80
3.29 Existence of a frame in which two space-like separated events occur simultaneously 82
3.30 Relativity and philosophy: the reality of the future 83
3.31 Relativistic addition of velocities 86
3.32 Relativistic aberration of velocities 88
3.33 Relativistic jets and the headlight effect 91
3.34 The group of boosts along the same axis 93
3.35 Connected components of the Lorentz group 95
3.36 Decomposition of proper orthochronous Lorentz transformations 96
3.37 Composition of non-collinear boosts and Wigner rotation 103

4 Relativistic dynamics 107
4.1 4-vectors, manifestly covariant equations and Lorentz invariants 107
4.2 The fundamental 4-vectors 109
4.2.1 4-displacement 109
4.2.2 4-velocity 109
4.2.3 4-acceleration 110
4.2.4 4-momentum 110
4.3 The 0-th component of the 4-momentum and the mass-energy equivalence relation 111
4.4 The energy–momentum relation 112
4.4.1 Consequence 1: massless particles must travel at light speed 112
4.4.2 Consequence 2: the relativistic mass is not constant 113
4.4.3 Consequence 3: particles with mass > 0 can't travel at light speed . . . 113

Chapter 1

History and philosophical motivations

1.1 Disclaimer: Special Relativity is *not* easy

- There is a widespread belief that special relativity is “easy”, whereas general relativity is “difficult”
- A more accurate statement is that *the mathematics* of special relativity is... relatively easy, compared with that of general relativity
- Special relativity is the extraordinary outcome of a philosophical journey that began with Galileo and culminated with Einstein
- What makes it difficult is that **it defies our everyday intuition**: relativistic effects become apparent only at a non-negligible fraction of the speed of light
- Thus, while the mathematical tools we will use are straightforward, their **interpretation is not**. Taking them lightly would be a serious mistake, leading to misleading or incorrect applications of the theory.

1.2 The teleological view

- Ancient thinkers, the most widely quoted being **Aristotle**, had a **teleological view** of the world that surrounds us
- Teleological: the ‘things’ of the worlds evolve toward their... final *purpose*
- What’s this purpose? Staying still! Rest was considered the natural state of things
- At that time this seemed pretty obvious, after all, a rock remains there if nobody pushes it!
- In order to make things change their ‘natural’ state, what today we call a ‘force’ must be applied, and if we want things to keep moving, that force must be continuously applied.

1.3 A problem with a bow and arrow...

- A big mystery: an arrow keeps flying through the air long after the initial force is applied by the bowstring, why?

- **Philoponus** (Egypt ∼600): the first idea of 'momentum': the bowstring imparted an 'impetus' to the arrow that eventually dissipated away

- Important deviation from the teleological view: things don't move for a predefined purpose, they have 'properties' that exist in the moment

- Further development by **Ibn Sina** (Persia ∼1000): he argued that impetus is not temporary, but always there (0 for stationary objects), and can be modified

- Ibn Sina: the first to guess that, without the air resistance, a moving body would keep moving indefinitely!

- The modern idea that objects don't evolve toward an ultimate goal, but obey laws that predict what will happen in the future based on what's happening now was slowly building up.

1.4 The modern era: Newtonian mechanics

- In *Philosophiæ Naturalis Principia Mathematica* – 1687, Isaac **Newton** (1642-1727), introduced **THE LAW** that governs motion of objects:

$$\mathbf{F}(t, \mathbf{x}, \dot{\mathbf{x}}) = m\, \ddot{\mathbf{x}}(t)$$

  - $\mathbf{x}$: **position** vector
  - $\mathbf{v} = \dot{\mathbf{x}}$: **velocity** vector, $v = \|\dot{\mathbf{x}}\|$: **speed**
  - $\mathbf{a} = \ddot{\mathbf{x}}$: **acceleration** vector
  - $\mathbf{F}$: force acting upon a body of mass m and position $\mathbf{x}$ at time t

- Einstein's quote (1931):

> '*In order to give his system mathematical form,* ***Newton had first to invent the concept of the differential quotient****, and to draw up the* ***laws of motion*** *in the form of* ***differential equations****, perhaps...* ***the greatest intellectual step that it has ever been given to one man to take***'.

1.5 Motion in space. . . but what *is* space?

- In the early 1700s the **ontological** question of what space *really is* was absolutely central
- Two philosophical views:
  - **Clarke** (England, inspired by Newton): **substantivalism**, space is a substance itself, with a separate existence from the things within it
  - **Leibniz** (Germany, inspired by Descartes): **relationalism**, the space isn't a thing, it's just defined by the distance between things within it
- **Newton** thought that **space is absolute**, while Leibniz that space couldn't be defined without a relation to the objects in it
- Most physicists nowadays content themselves with an **epistemic** view of space: what we know about it thanks to observations:
  - Einstein's general relativity pushes toward substantivalism: space (as part of spacetime) has a life on its own, it bends and curves in response to the presence of mass and energy
  - Quantum mechanics seems to push toward relationalism: space is emergent from entanglement.

1.6 Motion through time. . . but what *is* time?

- The **arrow of time**[1] is, fundamentally, an asymmetry: time moves from the past to the future, not the other way round
- Before Newtonian mechanics this seemed so obvious that no explanation was required
- However, at the microscopic level[2], the fundamental forces between particles depend only on their relative positions $\mathbf{F} = \mathbf{F}(\mathbf{x})$, this implies that Newton's equation is invariant under time reversal:

$$t' = -t \implies dt' = -dt \implies \mathbf{v}' := \frac{d}{dt'}\mathbf{x} = -\mathbf{v} \implies \mathbf{a}' := \frac{d}{dt'}\mathbf{v}' = \mathbf{a}$$

- So, the fundamental laws of Newtonian mechanics don't know the difference between moving toward the past and the future.

[1] so named by the British astronomer Arthur Eddington in 1927.

[2] dissipative forces (such as friction or viscosity) also depend on velocity, but they are effective *macroscopic* descriptions of phenomena, not fundamental ones.

1.7 Laplace's demon

- **Laplace's demon** (1814) *Essai philosophique sur les probabilités*:

> We may regard the present state of the universe as **the effect of its past and the cause of its future**. An intellect which at a certain moment would know all forces that set nature in motion, and all positions of all items of which nature is composed, if this **intellect** were also **vast enough** to submit these data to analysis, it would embrace in a single formula all movements. For such an intellect nothing would be uncertain and **the future just like the past would be present** before its eyes.

1.8 The past hypothesis

- None of us is nowhere close to be the Laplace demon! We're finite creatures with dramatically limited capacities of measurement and calculation!

- Only in the 1870s, the great Austrian physicist Ludwig **Boltzmann** (1844-1906) came up with the idea of using **entropy to explain the arrow of time**: time evolves toward states of higher entropy (2nd law of thermodynamics)

- This explanation hides an implicit assumption: entropy started low in the first place: *the past hypothesis*

- Assuming the big bang theory is correct, this means that at the moment of the big bang the entropy was the lowest possible. So, we experience an arrow of time because we live (relatively speaking) 'close' to the big bang!

- Some cosmologists claim that, very (very!) far in the future, entropy will reach a maximal state and time will effectively lose its physical meaning...

1.9 Standing on the shoulders of giants

- In a letter to Robert Hooke (1675), Newton wrote[3]:

  '*If I have seen further it is by standing on the shoulders of giants*'.

- The giants Newton was referring to are:
  - Johannes Kepler (1571-1630)
  - René Descartes (1596-1650)
  - Galileo Galilei (1564-1642), who passed the same year that Newton was born: the stars aligned!

[3] Some historians believe that this sentence was also meant to insult Hooke for being minute of stature. Newton ended up in many bitter disputes, the most well-known being that with Leibniz about the paternity of calculus.

1.10 Galilean relativity

- Many consider that Galileo and Einstein had *the greatest physical intuition* ever
- While sailing on a slowly moving boat and seeing an object falling straight on the floor, Galileo had the intuition that there would be no mechanical experiment that you could do to measure how fast you were (uniformly) moving or even... if you're moving at all!
- That's because there's no such thing as '*how fast a body is moving with uniform velocity*' without specifying *relative to* what!
- Hence the name... *relativity*.

Figure 1.1: The canopy where Galileo used to teach at the university of Padova.

1.11 Galilean relativity upset a number of people...

- Galileo was also an extraordinary talented writer, he wrote books expanding his ideas that made him the most popular scientist of his time in Europe

- His claim that there is no absolute measure of rest in the universe, implied that *there is no preferred location in the universe*, which contradicted the fact that standing still on planet Earth is a preferred reference frame...

- Even today some people have problems grasping Galilean relativity, imagine in the early 1600, when people preferred to think of themselves as living in the center of the universe...

1.12 Inertial reference frames and the principle of inertia

- Physically speaking, a **reference frame** is the **choice of a triple of Cartesian axes** endowed with **identical and synchronized clocks in every point** of space

- A reference frame in which all free body, i.e. not subjected to any interaction, is at rest or moves with a straight uniform motion is said to be **inertial**

- **Principle of inertia** (fundamental for both Newtonian mechanics and special relativity): it is always possible, at least theoretically, to find an inertial reference frame, indicated with $\mathscr{R}$ from now on

- The principle of inertia is only approximately satisfied

- For a reference frame attached to the Earth at the sea level we have:

  - Angular velocity of the Earth's rotation (a complete rotation (2π) divided by the period): $\Omega_\oplus = \frac{2\pi}{\text{day}} = \frac{2\pi}{24 \cdot 3600}\ \text{s}^{-1} = 7.3 \times 10^{-5}\ \text{s}^{-1}$

  - Earth's radius: $R_\oplus = 6.4 \times 10^6$ m

  - Centrifugal acceleration due to Earth's rotation: $a_\oplus = \Omega_\oplus^2 R_\oplus = 3.4 \times 10^{-2}\ \text{m}\,\text{s}^{-2}$, around 300 times weaker than $9.81\ \text{m}\,\text{s}^{-2}$, the acceleration of Earth's gravity at the sea level

  - The acceleration due to the revolution of the Earth around the Sun is 5 times weaker than that due to its rotation.

1.13 Inertial reference frames and forces

- Galileo ruled out the existence of a preferred position or a preferred velocity, but didn't deny the existence of a preferred acceleration. That's because, contrary to position and velocity, *there is* a preferred acceleration: zero!

- Position: $\mathbf{x}(t)$, velocity: $\mathbf{v}(t) = \dot{\mathbf{x}}(t)$, acceleration: $\mathbf{a}(t) = \dot{\mathbf{v}}(t) = \ddot{\mathbf{x}}(t)$, so:

  $$\mathbf{a}(t) = \mathbf{0} \implies \mathbf{v}(t) \equiv \mathbf{v}_0 \implies \mathbf{x}(t) = \mathbf{x}_0 + \mathbf{v}_0 t$$

  zero acceleration means a *uniform motion on a straight line*

- So, using Newton's law, we can say that **a reference system on which no resultant force acts upon is inertial**.

1.14 The mathematical postulates of Galilean relativity

- Physically, the emphasis of Galilean relativity is on the fact that **all inertial frames are equivalent**, when used to describe *mechanical experiments*[4]

- Mathematically, Galilean relativity can be translated into properties of time and space

- *Time is homogeneous* (no preferred instants) and flows in $\mathbb{R}$

- *Space* is identified with $\mathbb{R}^3$, endowed with the *Euclidean geometry* and considered

  - *homogeneous*: no point in space is preferred to another

  - *isotropic*: no privileged direction in space.

[4] we'll see why it is important to underline that Galileo and then Newton only talked about mechanical experiments...

1.15 Clock synchronization

- We have said that a reference frame must be equipped with clocks at every point in space. At first glance, this may seem an excessive requirement, as one might think that a single clock is sufficient, but in reality this is not the case

- The study of motion involves time measurements carried out *at different points in space*, e.g., to measure the average velocity of a body, we need to perform two observations at two points and at two different instants:

  - passage of the body at point 1 at time t_1
  - passage of the body at point 2 at time t_2

  therefore, in this case, two synchronized clocks are required (if you are at point 1, you need to be synchronized with point 2, and vice-versa)

- There are several ways to achieve that, at least theoretically, we discuss one proposed by Einstein, which makes use of light signals. Since the velocity at which they travel is finite, it requires an *assumption*: we establish that *the time taken by light to go from 1 to 2 is* ***the same*** *as that taken to go from 2 to 1*

- There is nothing that guarantees *a priori* that this is the case. The validity of this assumption lies in the logical consistency of the resulting theory and in the experimental verification of its empirical content (that is, in checking its predictions)

- Following the assumption, we can say that two clocks at 1 and 2 are synchronized if:

  - by sending at time $t = 0$ a light signal from 1 toward a mirror placed at 2...
  - when we read the instant t on the clock at 1, we read a time $t/2$ on the clock at 2

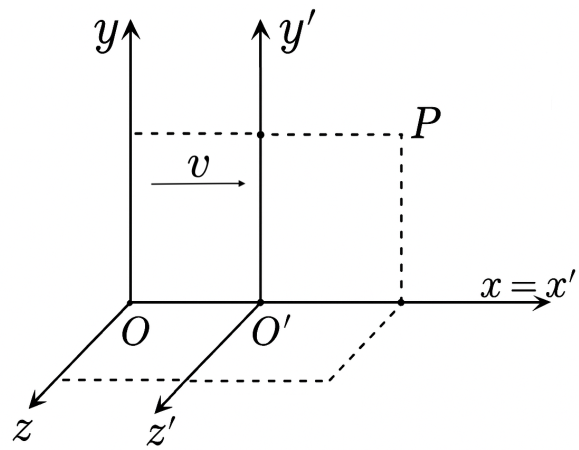
- The synchronization procedure is therefore the following. Let two observers, Alice at 1 and Bob at 2, separated by a distance L:

  - Alice sends a light signal when her clock reads time $t = 0$
  - Bob, at the instant he receive the signal, sets his clock to the time $t = L/c$
  - The assumption on the equality of the speed of light in the two directions ensures that if Bob sends, at time t, a light signal toward 1, Alice will receive the signal at time $t + L/c = t + t = 2t$, thus achieving synchronization automatically

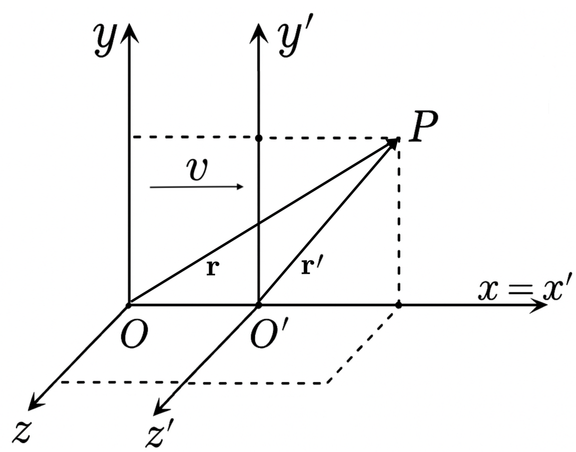
- The analysis of the concept of simultaneity is one of the fundamental elements of Einsteinian theory. *Measuring time*, Einstein emphasizes, *means establishing the simultaneity of two events*: "*We must consider that all our judgments in which time plays a role are always judgments about simultaneous events. If, for example, I say: 'That train arrives here at 7 o'clock,' I mean that the appearance of 7 on my clock and the arrival of the train are simultaneous events*". The most immediate and striking consequence of the special relativity will be precisely... the relativity of simultaneity.

1.16 Galilean transformations between inertial frames in the standard configuration

- *Standard configuration* (between two inertial reference frames $\mathscr{R}$ and $\mathscr{R}'$): parallel axes and the relative motion has constant velocity v along the horizontal axis:



- How can we relate the description of motion of $\mathscr{R}$ and $\mathscr{R}'$?



- Time is considered absolute, so $t = t'$

- Motion occors in the horizontal direction, so $y = y'$ and $z = z'$

- Spatial separation in the horizontal axis at t: $x = x' + vt$ and $x' = x - vt$ (affine)

- Galilean transformations (called like this only starting by 1909 in a paper of Philipp Frank!) in the standard configuration

$$\begin{cases} t' = t \\ x' = x - vt \\ y' = y \\ z' = z \end{cases} \iff \begin{cases} t = t' \\ x = x' + vt \\ y = y' \\ z = z' \end{cases}$$

- 1 spatial dimension:

$$\begin{pmatrix} t' \\ x' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -v & 1 \end{pmatrix} \begin{pmatrix} t \\ x \end{pmatrix} \equiv \mathcal{G}(v) \begin{pmatrix} t \\ x \end{pmatrix} \iff \begin{pmatrix} t \\ x \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ v & 1 \end{pmatrix} \begin{pmatrix} t' \\ x' \end{pmatrix} \equiv \mathcal{G}(v)^{-1} \begin{pmatrix} t' \\ x' \end{pmatrix},$$

with $\mathcal{G}(v)^{-1} = \mathcal{G}(-v)$

- 3 spatial dimensions:

$$\mathcal{G}(v) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -v & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

1.17 Homogeneous Galilean transformations between inertial frames in a generic configuration

- More generally, for a generic configuration we have the affine vector sum:

$$\mathbf{r}' = \mathbf{r} - \mathbf{v}t \quad \iff \quad \mathbf{r} = \mathbf{r}' + \mathbf{v}t$$

- If the relative velocity between inertial frame is $\mathbf{v} = (v_x, v_y, v_z)$, then the Galilean transformation is:

$$\begin{cases} t' = t \\ x' = x - v_x t \\ y' = y - v_y t \\ z' = z - v_z t \end{cases}$$

and the Galilean matrix is:

$$\mathcal{G}(\mathbf{v}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -v_x & 1 & 0 & 0 \\ -v_y & 0 & 1 & 0 \\ -v_z & 0 & 0 & 1 \end{pmatrix}$$

- The appearance of affine transformations is not accidental: inertial motion is described by straight lines in space, and maps between inertial frames must preserve them

- Under the natural assumption that such transformations are bijective, it follows from the fundamental theorem of affine geometry that they must be **affine**.

1.18 Galilean velocity addition law

- Consider two inertial frames $\mathscr{R}$ and $\mathscr{R}'$ related by a Galilean transformation in a generic configuration:

$$\mathbf{r}' = \mathbf{r} - \mathbf{v}t, \qquad t' = t,$$

where $\mathbf{v} = (v_x, v_y, v_z)$ is the velocity of the $\mathscr{R}'$ frame with respect to $\mathscr{R}$

- Let a particle have trajectory $\mathbf{r}(t)$ in $\mathscr{R}$ and $\mathbf{r}'(t')$ in $\mathscr{R}'$. Since $t' = t$, we can differentiate with respect to the same time variable:

$$\frac{d\mathbf{r}'}{dt'} = \frac{d}{dt}\left(\mathbf{r} - \mathbf{v}t\right).$$

- Therefore, if

$$\mathbf{u} = \frac{d\mathbf{r}}{dt}, \qquad \mathbf{u}' = \frac{d\mathbf{r}'}{dt'},$$

then

$$\mathbf{u}' = \mathbf{u} - \mathbf{v} \qquad \Longleftrightarrow \qquad \mathbf{u} = \mathbf{u}' + \mathbf{v}$$

- Hence, in Galilean relativity, velocities add according to the ordinary vector sum:

$$\boxed{\mathbf{u} = \mathbf{u}' + \mathbf{v}}$$

- Notice that a key component to derive this law is the fact that time is absolute in Galilean kinematics: all inertial observers use the same time parameter.

1.19 The most general Galilean transformations and the Galilei group

- The previous Galilean transformations were derived under two simplifying assumptions:
  (i) the spatial origins and initial times of $\mathscr{R}$ and $\mathscr{R}'$ coincide
  (ii) the axes of $\mathscr{R}$ and $\mathscr{R}'$ are parallel

- If we remove these restrictive assumptions, the resulting Galilean transformations are no longer homogeneous and incorporate a possible rotation of the axes. Explicitly:

$$\begin{cases} \mathbf{x}' = R\mathbf{x} - \mathbf{v}t + \mathbf{x}_0, \\ t' = t + t_0, \end{cases}$$

  where $\mathbf{x}_0$ and t_0 are constants, and $R \in \mathrm{SO}(3)$ is a a special orthogonal matrix ($R^T R = I_3$, $|R| = 1$) representing a three-dimensional rotation

- These are the most general non-homogeneous Galilean transformations between two inertial frames. They depend on a total of **10 parameters**:
  - 3 free parameters of R
  - 3 components of $\mathbf{v}$
  - 3 components of $\mathbf{x}_0$
  - and t_0

- The transformations above form a group, called the **Galilei group**. An element of the group is specified by a quadruple $(R, \mathbf{v}, \mathbf{x}_0, t_0)$, where $R \in \mathrm{SO}(3)$, $\mathbf{v}, \mathbf{x}_0 \in \mathbb{R}^3$ and $t_0 \in \mathbb{R}$

- Given (R_1, v_1, x_1, t_1) and (R_2, v_2, x_2, t_2), we have
  - Group law (quite complicated!):

$$(R_2, \mathbf{v}_2, \mathbf{x}_2, t_2) \cdot (R_1, \mathbf{v}_1, \mathbf{x}_1, t_1) = \big(R_2 R_1,\ \mathbf{v}_2 + R_2\mathbf{v}_1,\ \mathbf{x}_2 + R_2\mathbf{x}_1 - \mathbf{v}_2 t_1,\ t_2 + t_1\big)$$

  - Identity: $(I_3, \mathbf{0}, \mathbf{0}, 0)$
  - Inverse:

$$(R, \mathbf{v}, \mathbf{x}_0, t_0)^{-1} = \big(R^T,\ -R^T\mathbf{v},\ -R^T(\mathbf{x}_0 + \mathbf{v}t_0),\ -t_0\big)$$

- The Galilean transformation $(R, \mathbf{v}, \mathbf{x}_0, t_0)$ can be represented as an affine action on $(t, \mathbf{x}) \in \mathbb{R} \times \mathbb{R}^3$, or equivalently as a linear action on the augmented vector $(1, t, \mathbf{x})^T \in \mathbb{R}^5$:

$$\begin{pmatrix} 1 \\ t' \\ \mathbf{x}' \end{pmatrix} = \begin{pmatrix} 1 & 0 & \mathbf{0}^T \\ t_0 & 1 & \mathbf{0}^T \\ \mathbf{x}_0 & -\mathbf{v} & R \end{pmatrix} \begin{pmatrix} 1 \\ t \\ \mathbf{x} \end{pmatrix}, \quad \mathbf{0},\, \mathbf{x}_0,\, \mathbf{v} : \text{ column vectors of } \mathbb{R}^3.$$

1.20 The principle of Galilean relativity as a meta-law of physics

- The principle of Galilean relativity allows us to limit what types of physical laws are admissible with it. In this sense it can be considered as a law on the physical laws, hence ... a meta-law

- For example, consider Newton's equation for the gravitational interaction between the the bodies a and b
$$Gm_am_b\frac{\mathbf{x}_b - \mathbf{x}_a}{|\mathbf{x}_b - \mathbf{x}_a|^3} = m_a\,\ddot{\mathbf{x}}_a,$$
under a Galilean transformation with relative velocity $\mathbf{v}$,
$$t' = t, \qquad \mathbf{x}'_a = \mathbf{x}_a - \mathbf{v}t, \qquad \mathbf{x}'_b = \mathbf{x}_b - \mathbf{v}t,$$
hence $\mathbf{x}'_b - \mathbf{x}'_a = (\mathbf{x}_b - \mathbf{v}t) - (\mathbf{x}_a - \mathbf{v}t) = \mathbf{x}_b - \mathbf{x}_a$, so that
$$\frac{\mathbf{x}'_b - \mathbf{x}'_a}{|\mathbf{x}'_b - \mathbf{x}'_a|^3} = \frac{\mathbf{x}_b - \mathbf{x}_a}{|\mathbf{x}_b - \mathbf{x}_a|^3}.$$
Moreover, since $\mathbf{v}$ is constant in t and $t' = t$,
$$\ddot{\mathbf{x}}'_a = \frac{d^2}{dt^2}(\mathbf{x}_a - \mathbf{v}t) = \ddot{\mathbf{x}}_a,$$
therefore,
$$Gm_am_b\frac{\mathbf{x}'_b - \mathbf{x}'_a}{|\mathbf{x}'_b - \mathbf{x}'_a|^3} = m_a\ddot{\mathbf{x}}'_a,$$
which has exactly the same form as the equation in $\mathscr{R}$. In general, Newton's law with a force that depends only on the distance is invariant under Galilean transformations.

- On the contrary, the law
$$k\,\frac{\mathbf{x}_a + \mathbf{x}_b}{|\mathbf{x}_a + \mathbf{x}_b|^3} = m_a\ddot{\mathbf{x}}_a(t)$$
is not Galilean invariant, indeed we still have $\ddot{\mathbf{x}}'_a = \ddot{\mathbf{x}}_a$, but
$$\mathbf{x}_a + \mathbf{x}_b = (\mathbf{x}'_a + \mathbf{v}t) + (\mathbf{x}'_b + \mathbf{v}t) = \mathbf{x}'_a + \mathbf{x}'_b + 2\mathbf{v}t.$$
therefore in $\mathscr{R}'$ the equation becomes
$$k\,\frac{\mathbf{x}'_a + \mathbf{x}'_b + 2\mathbf{v}t'}{|\mathbf{x}'_a + \mathbf{x}'_b + 2\mathbf{v}t'|^3} = m_a\ddot{\mathbf{x}}'_a,$$
which is not of the same form as the initial one, hence the law is not invariant under Galilean transformations. If we consider correct the principle of relativity, *this law of motion is not admissible*, because we could perform experiments that depend on $\mathbf{v}$.

1.21 Is there a limit on the velocity in Newtonian mechanics?

- No!
- For example, in Newton's gravitational law, instead of a local, step-by-step, propagation of interaction, the force exerted by a particle on another at a particular instant of time depends only on their distance at that precise instant, so a change in its position affects the other *immediately* (**action at a distance**)
- So, in Newtonian mechanics the velocity of interactions that depend only on spatial position is infinite!

1.22 Action at a distance

- Newton's letter to Richard Bentley (1692):

'*That gravity should be innate, inherent, and essential to matter, so that one body may act upon another* ***at a distance*** *through a vacuum, without the mediation of anything else [...]* ***is to me so great an absurdity, that I believe no man who has in philosophical matters a competent faculty of thinking, can ever fall into it****. Gravity must be caused by an agent acting constantly according to certain laws; but whether this agent be material or immaterial,* ***I have left to the consideration of my readers***'.

- 1676: after Galileo's discovery of Jupiter moons, Ole Rømer studied the eclipses of Io, noticing that they didn't always occur exactly at the predicted times:
  - when Earth was moving away from Jupiter, the eclipses of Io were observed later than expected
  - when Earth was moving closer to Jupiter, they were observed earlier
- Rømer understood that this delay was not caused by Io itself, but by the time needed by light to travel a longer distance and concluded that light has a finite speed, even though extremely large.
- Several other experiments $\Longrightarrow$ *instantaneous interaction does not exist in nature*: any change in a particle starts to have an influence on another particle only after a certain lapse of time $\Longrightarrow$ $\exists$ *limit speed of interaction propagation*:

$$\frac{\text{spatial separation between the two particles}}{\text{lapse of time after interaction occurs}}.$$

1.23 Maxwell's equations and the speed of light

- The dynamics of the electromagnetic field is governed by Maxwell's equations, using Gaussian units:

$$\begin{cases} \nabla \cdot \mathbf{E} = 4\pi\rho & \text{(Gauss's law)} \\ \nabla \cdot \mathbf{B} = 0 & \text{(absence of magnetic monopoles)} \\ \nabla \times \mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{B}}{\partial t} = 0 & \text{(Faraday–Neumann–Lenz law)} \\ \nabla \times \mathbf{B} - \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} = \frac{4\pi}{c}\mathbf{j} & \text{(Ampère–Maxwell law)}, \end{cases}$$

  $\mathbf{E}$ is the electric field, $\mathbf{B}$ the magnetic field, ρ the charge density, and $\mathbf{j}$ the current density. In the absence of sources ($\rho = 0$, $\mathbf{j} = \mathbf{0}$), the above equations reduce to

$$\begin{cases} \nabla \cdot \mathbf{E} = 0 & (*) \\ \nabla \cdot \mathbf{B} = 0 & \\ \nabla \times \mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{B}}{\partial t} = 0 & \\ \nabla \times \mathbf{B} - \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} = 0 & (**). \end{cases}$$

  Applying the curl operator ($\nabla\times$) to the third equation, we obtain

$$\nabla \times (\nabla \times \mathbf{E}) = \nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} = -\frac{1}{c}\frac{\partial}{\partial t}(\nabla \times \mathbf{B}).$$

  and using $(*)$ and $(**)$ we obtain

$$\nabla^2 \mathbf{E} - \frac{1}{c^2}\frac{\partial^2 \mathbf{E}}{\partial t^2} = 0,$$

  which is the **wave equation** and, in a similar way, one can show that $\mathbf{B}$ satisfies the same equation. According to this, an electromagnetic wave propagates in vacuum with a velocity given by the constant[5] c:

$$c = \frac{1}{\sqrt{\varepsilon_0 \mu_0}} \simeq 300000\, km\, s^{-1}$$

  ε_0, μ_0: *electrical permittivity* and *magnetic permeability* of vacuum, resp. Weber and Kohlrausch, in 1856, experimentally determined the value of c and found that it was equal to the speed of light in vacuum. This led to the conclusion that optical phenomena are electromagnetic in nature.

- Richard P. Feynman (1918-1988) celebration of Maxwell's work:

  '*From a long view of the history of mankind – seen from say, 10000 years from now – there can be little doubt that* ***the most significant event of the 19th century will be judged as Maxwell's discovery of the laws of electrodynamics****.*
  The American Civil War will pale into provincial insignificance in comparison with this important scientific event of the same decade'.

[5]Lorentz used c for *celeritas*, which means speed in latin.

1.24 Luminiferous aether

- But with respect to which reference frame does the speed of light equal c?

- In 19th century physics, all known waves required a medium to propagate. Therefore, a *hypothetical* medium filling all space was introduced: the **luminiferous aether**.

- The aether was assumed to have very contradictory properties:

  - it had to be **extremely rigid** (to sustain high-frequency waves such as light)

  - yet **extremely tenuous** (so as not to resist the motion of planets)

  - present everywhere, even in vacuum…so that is wasn't vacuum anymore…

- If the propagation of light is interpreted as a wave phenomenon occurring in this medium, its speed is expected to depend on the motion of the observer with respect to the aether. In other words, if we perform a Galilean transformation $x' = x - vt$, we find that c must change.

- This seemed to be in contradiction with an increasingly abundant amount of experimental evidences, the most famous of which was conducted with the **Michelson-Morley** interferometer in 1887 and showed no change in c.

1.25 Maxwell's equations are not Galilean invariant!

- There was also another (big) problem: Maxwell's equations are not Galilean invariant!

- To prove it, for concreteness, consider the x-component of the Faraday–Neumann–Lenz equation,

$$\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} = -\frac{1}{c}\frac{\partial B_x}{\partial t}, \tag{1.1}$$

which we assume to hold in a given inertial frame $\mathscr{R}$. To pass to a frame $\mathscr{R}'$ related to $\mathscr{R}$ by $\mathbf{x}' = \mathbf{x} - \mathbf{v}t$, $t' = t$, we use the chain rule to write the relations

$$\frac{\partial}{\partial x} = \frac{\partial x'}{\partial x}\frac{\partial}{\partial x'} + \frac{\partial t'}{\partial x}\frac{\partial}{\partial t'} = \frac{\partial(x - v_x t)}{\partial x}\frac{\partial}{\partial x'} + \frac{\partial t}{\partial x}\frac{\partial}{\partial t'} = \frac{\partial x}{\partial x}\frac{\partial}{\partial x'} + 0 \cdot \frac{\partial}{\partial t'} = \frac{\partial}{\partial x'},$$

similarly

$$\frac{\partial}{\partial y} = \frac{\partial}{\partial y'},$$

$$\frac{\partial}{\partial z} = \frac{\partial}{\partial z'},$$

but

$$\frac{\partial}{\partial t} = \frac{\partial x'}{\partial t}\frac{\partial}{\partial x'} + \frac{\partial t'}{\partial t}\frac{\partial}{\partial t'} = \frac{\partial(x - v_x t)}{\partial t}\frac{\partial}{\partial x'} + \frac{\partial}{\partial t'} = -v_x\frac{\partial}{\partial x'} + \frac{\partial}{\partial t'}.$$

- Eq. (1.1) therefore becomes

$$\frac{\partial E_z}{\partial y'} - \frac{\partial E_y}{\partial z'} = -\frac{1}{c}\left(\frac{\partial B_x}{\partial t'} - v_x\frac{\partial B_x}{\partial x'}\right)$$

and it's not possible that the field $\mathbf{B}$ transforms in such a way that this last equation can be rewritten in the same form as eq. (1.1), i.e. there is no way to eliminate the extra term proportional to v_x, since Maxwell's equations don't contain any relation linking $\frac{\partial B_x}{\partial t'}$ and $\frac{\partial B_x}{\partial x'}$.

- Therefore, the Galilean principle of relativity is violated by electromagnetism. For example, by an explicit calculation, it can be verified that Maxwell's equations in the absence of sources transform as follows under Galilean tranformations:

$$\nabla' \cdot \mathbf{E}' = -\frac{1}{c}\,\mathbf{v} \cdot (\nabla' \times \mathbf{B}')$$

$$\nabla' \cdot \mathbf{B}' = \frac{1}{c}\,\mathbf{v} \cdot (\nabla' \times \mathbf{E}')$$

$$\nabla' \times \mathbf{B}' - \frac{1}{c}\frac{\partial \mathbf{E}'}{\partial t'} = -\frac{1}{c^2}\,\mathbf{v} \times \frac{\partial \mathbf{B}'}{\partial t'}$$

$$\nabla' \times \mathbf{E}' + \frac{1}{c}\frac{\partial \mathbf{B}'}{\partial t'} = -\frac{1}{c^2}\,\mathbf{v} \times \frac{\partial \mathbf{E}'}{\partial t'}.$$

1.26 The wave equation is not Galilean invariant either!

- Consider for simplicity the wave equation in one spatial dimension:

$$\frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2} = \frac{\partial^2\phi}{\partial x^2}.$$

- Consider a Galilean transformation between two inertial frames:

$$x' = x - vt, \qquad t' = t.$$

- Using the chain rule, we compute the derivatives in the new coordinates as done before:

$$\frac{\partial}{\partial x} = \frac{\partial}{\partial x'}, \qquad \frac{\partial}{\partial t} = \frac{\partial}{\partial t'} - v\frac{\partial}{\partial x'}.$$

- We now compute the second time derivative by successive application:

$$\frac{\partial^2\phi}{\partial t^2} = \frac{\partial}{\partial t}\left(\frac{\partial\phi}{\partial t}\right) = \frac{\partial}{\partial t}\left(\frac{\partial\phi}{\partial t'} - v\frac{\partial\phi}{\partial x'}\right).$$

- Applying again the expression for $\frac{\partial}{\partial t}$, we obtain

$$\begin{aligned}\frac{\partial^2\phi}{\partial t^2} &= \left(\frac{\partial}{\partial t'} - v\frac{\partial}{\partial x'}\right)\left(\frac{\partial\phi}{\partial t'} - v\frac{\partial\phi}{\partial x'}\right)\\ &= \frac{\partial^2\phi}{\partial t'^2} - v\frac{\partial^2\phi}{\partial x'\partial t'} - v\frac{\partial^2\phi}{\partial t'\partial x'} + v^2\frac{\partial^2\phi}{\partial x'^2}.\end{aligned}$$

- Since mixed partial derivatives commute, we have

$$\frac{\partial^2\phi}{\partial t^2} = \frac{\partial^2\phi}{\partial t'^2} - 2v\frac{\partial^2\phi}{\partial t'\partial x'} + v^2\frac{\partial^2\phi}{\partial x'^2}.$$

- Moreover,

$$\frac{\partial^2\phi}{\partial x^2} = \frac{\partial^2\phi}{\partial x'^2}.$$

- Substituting into the wave equation, we obtain:

$$\frac{1}{c^2}\left(\frac{\partial^2\phi}{\partial t'^2} - 2v\frac{\partial^2\phi}{\partial t'\partial x'} + v^2\frac{\partial^2\phi}{\partial x'^2}\right) = \frac{\partial^2\phi}{\partial x'^2}.$$

- Rearranging,

$$\frac{1}{c^2}\frac{\partial^2\phi}{\partial t'^2} - \frac{2v}{c^2}\frac{\partial^2\phi}{\partial t'\partial x'} = \left(1 - \frac{v^2}{c^2}\right)\frac{\partial^2\phi}{\partial x'^2}$$

- This equation is not of the same form as the original wave equation, due to the presence of the mixed derivative term

$$\frac{\partial^2\phi}{\partial t'\partial x'}.$$

- Therefore, the wave equation is not invariant under Galilean transformations.

1.27 Reactions to the lack of Galilean invariance of Maxwell's equations

- There are two possible attitudes in response to the lack of Galilean invariance of electromagnetism:

  1. One may regard this as a serious difficulty for physical theory, and insist that the *principle of relativity should also hold for electromagnetism.* This implies that *Galilean transformations* (with their assumption of *independence of time* from the reference frame) *are not* the *correct* transformations between inertial frames, and that *mechanics must therefore be revised.*
  2. Alternatively, one may assume that the dependence of the speed of light on the reference frame is a real physical effect, to be investigated experimentally. In this case, the principle of relativity would be restricted to mechanics alone, while *electromagnetism would introduce a preferred reference frame: the aether*!

- The majority of the physics community at the turn of the nineteenth and twentieth centuries supported this latter view. However, somebody started to think that time wasn't absolute anymore...

1.28 Lorentz's local time

- 1892: Hendrick Antoon **Lorentz** (1853-1928) worked on light propagation in reference frames that move relative to the luminiferous aether

- He managed to explain known phenomena by postulating the existence of a *local time*, a function of the 'universal time' and the position (Poincaré called it 'Lorentz's most ingenious idea')

- He also shown that the existence of a local time leads to *length contraction*, already investigated by George Francis **Fitzgerald** (1851-1901) in 1889

- In 1899 and 1904, Lorentz discussed *time dilation* and completed his transformations, which extended Galilean transformations if local time is taken into account

- In the 1904 paper he also proved that Maxwell's equations are invariant under Lorentz transformations.

1.29 Galilean vs. Lorentz transformations between inertial frames

- **Galilean transformations** (in the standard configuration)

$$\begin{cases} t' = t \\ x' = x - vt \\ y' = y \\ z' = z \end{cases} \iff \begin{cases} t = t' \\ x = x' + vt \\ y = y' \\ z = z' \end{cases}$$

- **Lorentz transformations**[6] (in the standard configuration)

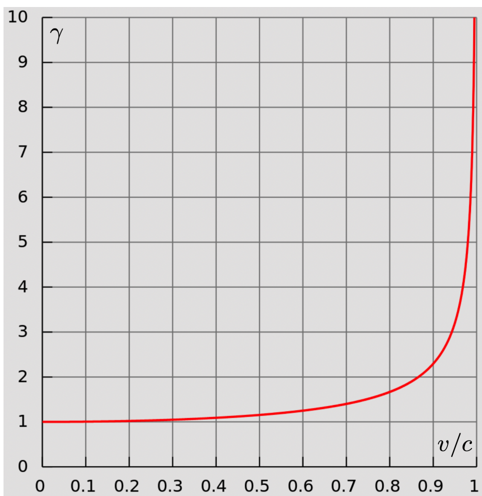
$$\begin{cases} t' = \gamma \left(t - \frac{v}{c^2}x\right) \\ x' = \gamma\,(x - vt) \\ y' = y \\ z' = z \end{cases} \iff \begin{cases} t = \gamma \left(t' + \frac{v}{c^2}x\right) \\ x = \gamma\,(x' + vt) \\ y = y' \\ z = z' \end{cases}$$

- **Lorentz factor**:

$$\gamma(v) = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \in [1, +\infty), \quad \gamma(v) \underset{v \to c_-}{\to} +\infty \text{ very slowly!}$$

- **Galilean limit**: $c \to +\infty$, so

$$\gamma(v) \to 1 \text{ and } \frac{v}{c^2}x \to 0 \implies \text{Lorentz transf.} \to \text{ Galilean ones.}$$



[6]to be proven in section 3.6.

1.30 Toward a systematic theory of space and time...

- The findings of Lorentz, Fitzgerald, Mach and others were fragmented, a comprehensive theory of space and time had to be built
- Poincaré and Einstein, inspired also by practical needs, were the leaders of this research
- **Henri Poincaré (1854–1912)**: member of the French Bureau des Longitudes
- **Albert Einstein (1879[7]–1955)**: worked at the Swiss Patent Office in Bern
- They were both concerned with the practical issues of **synchronization**: time coordination across large distances was an important issue at that time due to the spread of the telegraph system and the new European railway
- They worked on methods to define a common time standard using light signals and on the **operational meaning of simultaneity**.

1.31 Einstein relativity (in his 'miracle year': 1905)

- Albert **Einstein** '*On the Electrodynamics of Moving Bodies*' (1905) marks the beginning of special relativity
- Galilean relativity is 'updated' to take into account electromagnetic phenomena

Postulates of Einstein's theory of special relativity

1. The same postulate of Galilean relativity, but extended also to electromagnetic phenomena[8]
2. The speed of light in vacuum has the same constant value c when measured in all inertial reference frames, independently of their relative velocity

- With these postulates, Einstein could explain the underlying reason for Lorentz's local time (later called *proper*), derive in a much clearer fashion Lorentz transformations, length contraction, time dilation and predict many other facts!

[7] Stars aligned again: Einstein was born the same year Maxwell passed! (1879)
[8] First formulated by Poincaré.

1.32 Enters Minkowski: spacetime vs. space (1907)

- Hermann **Minkowski** (1864-1909) (who has been Einstein's professor) had the brilliant idea to interpret Einstein's theory in geometrical terms by blending space and time in a 4D *spacetime*. Two different visions...

- Minkowski wrote: '*space by itself, and time by itself are doomed to fade away into mere shadows, and only a kind of union of the two will preserve an independent reality*'

- Einstein: '*Minkowski's approach makes rather great demands on the reader in its mathematical aspects*'[9]

- Minkowski was Hilbert's best friend. Hilbert declared: '*I have become a mathematician thanks to the conversations during my walks with Minkowski*', he always considered himself *slow* compared to Minkowski...

1.33 A bit of relativity lingo: events and spacetime interval

- In *spacetime* both space and time are relative to an observer and the *geometry* is not Euclidean anymore, but... *Lorentzian*. To see what it means we need some typical relativity nomenclature...

- **Event**[10] $e = (x^0 = ct, x^1, x^2, x^3) = (ct, \mathbf{x})$, e.g. 'let's meet at our favourite restaurant at 8 p.m.', position and time must be given!

- Consider two events:

  - $e_1 = (ct_1, \mathbf{x}_1)$: a particle *starts* at time t_1 from position $\mathbf{x}_1$

  - $e_2 = (ct_2, \mathbf{x}_2)$: the particle *arrives* at time t_2 at position $\mathbf{x}_2$

- We have two ways to compute distances using the events coordinates:

  - Postulate 1 implicitly assumed that space has a Euclidean geometry, so the spatial distance between the *initial and final particle position* is

$$|\mathbf{x}_1 - \mathbf{x}_2|^2 = \sum_{i=1}^{3} (x_2^i - x_1^i)^2$$

  - Postulate 2 allows us to compute the distance traveled by a *light* signal between then *initial and final instant*: $c^2(t_2 - t_1)^2$.

[9] After his initial skepticism about Minkowski's work, he understood the huge importance of this geometrization by building on it his general theory of relativity!

[10] Using ct instead of t is customary in special relativity: physically, this amounts at replacing the time t with the corresponding space ct traveled by a ray of light during t, so that the coordinates (ct, x, y, z) have the **same dimensions**.

1.34 Spacetime geometry can't be Euclidean

- Minkowski had the struck of genius to notice that there is an object that encapsulates both the spatial (Postulate 1) and the temporal[11] (Postulate 2) computation of distance which happens to be the same in all inertial reference frame:

$$s_{1,2}^2 := c^2(t_2 - t_1)^2 - |\mathbf{x}_1 - \mathbf{x}_2|^2 \; : \textbf{spacetime interval} \;\; \text{between } e_1, e_2$$

- We are going to see that the spacetime interval is invariant under Lorentz transformations, so it has an *intrinsic physical meaning for spacetime*

- For light signals, the two distances agree, so: $s_{1,2}^2 = 0$, i.e. two events connected by a signal traveling at the speed of light have 0 spacetime interval

- However, if two events are connected by a signal traveling slower than the speed of light, then $s_{1,2}^2 > 0$

- Finally, if two events are so far apart that their Euclidean distance is bigger than the distance traveled by a light ray, then $s_{1,2}^2 < 0$

- Ahead of his time, Minkowski interpreted $s_{1,2}^2$ as an **indefinite quadratic form** associated to a pseudo scalar product

- A pseudo scalar product $\implies \exists$ non-null vectors with 0 pseudo-norm, called *isotropic vectors*

- Isotropic vectors for the Minkowski spacetime: vectors representing the propagation of light (null spacetime interval)

- In Euclidean geometry the norm is positive-definite $\implies \nexists$ isotropic vectors!

- With this very simple, but profound argument, Minkowski motivated why **spacetime geometry can't be Euclidean**

- We are going to see in chapter 2 that **the geometry of spacetime is hyperbolic** and the Euclidean space must be replaced with the **Lorentz spacetime**.

[11] the spatial computation of distance can be calculated with, a priori, any velocity, but the temporal distance is always computed using the speed of light, if we use another speed, then $s_{1,2}^2$ fails to be invariant in all reference frames.

Chapter 2

Lorentz geometry

Main reference: **Ratcliffe** '*Foundations of Hyperbolic Manifolds*'

2.1 Preliminaries on real quadratic forms

- A real quadratic form on $\mathbb{R}^n$ is a homogeneous polynomial of degree 2 in the coordinates of x. Thus, if $x = (x^0, \dots, x^{n-1}) \equiv (x^0, \mathbf{x})$, it can be written as

$$q(x) = \sum_{i,j=0}^{n-1} a_{ij} x^i x^j,$$

  with real coefficients $a_{ij} \in \mathbb{R}$. Equivalently, we can say that there exists a real symmetric matrix $A \in M_n(\mathbb{R})$ such that

$$q(x) = x^T A x$$

- $A \in M_n(\mathbb{R})$ can be written in block form as

$$A = \begin{pmatrix} a & \mathbf{b}^T \\ \mathbf{b} & C \end{pmatrix},$$

  where

$$a \in \mathbb{R}, \qquad \mathbf{b} \in \mathbb{R}^{n-1}, \qquad C = C^T \in M_{n-1}(\mathbb{R}).$$

- Therefore, if $q(x) = x^T A x$, then

$$q(x^0, \mathbf{x}) = \begin{pmatrix} x^0 & \mathbf{x}^T \end{pmatrix} \begin{pmatrix} a & \mathbf{b}^T \\ \mathbf{b} & C \end{pmatrix} \begin{pmatrix} x^0 \\ \mathbf{x} \end{pmatrix} = a(x^0)^2 + x^0 \mathbf{b}^T \mathbf{x} + \mathbf{x}^T \mathbf{b} x^0 + \mathbf{x}^T C \mathbf{x},$$

  hence we obtain that the general expression of a real quadratic form is

$$q(x^0, \mathbf{x}) = a(x^0)^2 + 2x^0 \mathbf{b} \cdot \mathbf{x} + \mathbf{x}^T C \mathbf{x}$$

- **Sylvester's law of inertia** states that every real quadratic form can be diagonalized by an invertible linear change of coordinates and written in the canonical form

$$q(x) = (x^0)^2 + \cdots + (x^p)^2 - (x^{p+1})^2 - \cdots - (x^{p+r})^2,$$

  and the integers

$$p = \text{number of positive signs}, \qquad r = \text{number of negative signs}$$

  do not depend on the chosen diagonalization: they are invariants of the quadratic form

- The pair (p, r) is called the **signature** of the quadratic form.

2.2 It's all about the (pseudo) inner product...

- The **Lorentz pseudo-inner product**[1] is the (indefinite) symmetric bilinear form $\circ : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ defined as follows:

$$\boxed{x \circ y = x^0 y^0 - x^1 y^1 - \cdots - x^{n-1} y^{n-1}} \qquad \forall x, y \in \mathbb{R}^n$$

- If we write $x \equiv (x^0, \mathbf{x})$, $y \equiv (y^0, \mathbf{y})$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{n-1}$, then

$$\boxed{x \circ y = x^0 y^0 - \langle \mathbf{x}, \mathbf{y} \rangle}$$

  $\langle \mathbf{x}, \mathbf{y} \rangle$: Euclidean inner product

- $(\mathbb{R}^n, \circ) = \mathbb{R}^{1,n-1}$: **Lorentz n-space**

- $(\mathbb{R}^4, \circ) = \mathbb{R}^{1,3}$: **Minkowski space $\mathcal{M}$.**

2.3 Lorentz and Minkowski indefinite quadratic forms

- The **Lorentz indefinite quadratic form** of signature $(1, n-1)$ $q : \mathbb{R}^{1,n-1} \to \mathbb{R}$ associated to $\circ$ can be obtained by computing $\circ$ on the diagonal elements on its domain, i.e.

$$\begin{aligned} q(x) := x \circ x &= (x^0)^2 - (x^1)^2 - \cdots - (x^{n-1})^2 \in -\mathbb{R}^+ \cup \{0\} \cup \mathbb{R}^+ \\ &= (x^0)^2 - |\mathbf{x}|^2 \end{aligned}$$

  $|\mathbf{x}|$: Euclidean norm of $\mathbf{x} = (x^1, \ldots, x^{n-1})$

- **Minkowski indefinite quadratic form** $q : \mathbb{R}^{1,3} \to \mathbb{R}$ associated to $\circ$:

$$q(x) = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2,$$

  in special relativity, it agrees with the **spacetime interval**

[1] The Lorentz pseudo-inner product is the key to understand hyperbolic geometry together with Möbius transformations.

- Lorentz pseudo-norm in $\mathbb{R}^{1,n-1}$:

$$\|x\| := \sqrt{q(x)} \in i\mathbb{R}^+ \cup \{0\} \cup \mathbb{R}^+$$

- Minkowski pseudo-norm in $\mathbb{R}^{1,3}$:

$$\|x\| = \sqrt{(x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2}.$$

2.4 Relation with the Euclidean inner product

- Relation with the Euclidean inner product (written either as $\langle\, , \rangle$ or with $\cdot$):

$$x \circ y = \langle x, \eta y\rangle = \langle \eta x, y\rangle$$

  where

$$\eta = \begin{pmatrix} 1 & 0 \\ 0 & -I_{n-1} \end{pmatrix}$$

  is the **Gram matrix** of $\circ$ w.r.t. the canonical basis $(e_j)_j$ of $\mathbb{R}^n$:

$$e_i \circ e_j = \eta_{ij}$$

- Crucial properties of η:

$$\eta^T = \eta^{-1} = \eta$$

$$\eta^2 = I_n,$$

  η: symmetric, orthogonal, coincides with its inverse and squares to the identity.

2.5 Likeness and orientation of a vector in the Lorentz space

- We can separate the Lorentz space in three different subsets by the **likeness** of the vectors $x \in \mathbb{R}^{1,n-1}$:

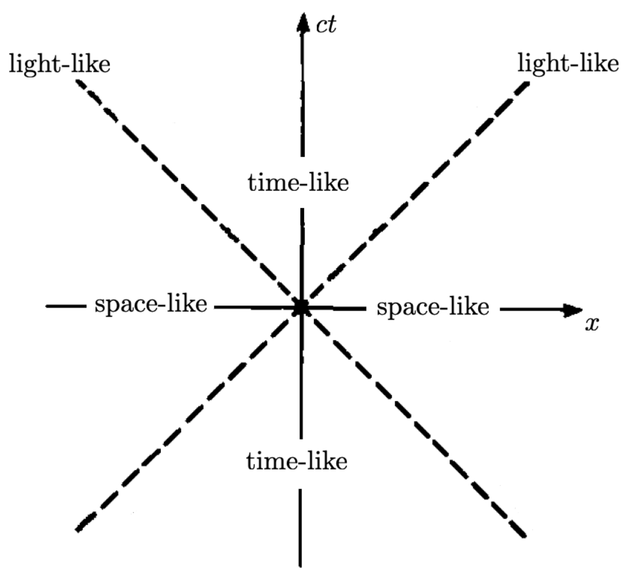
  - x **time-like** if $\|x\|^2 > 0$
  - x **light-like** if $\|x\|^2 = 0$ (the **isotropic vectors** of $\circ$)
  - x **space-like** if $\|x\|^2 < 0 \iff \|x\|$ is a pure imaginary number
  - x **causal** if it is not space-like, i.e. if it is either time-like or light-like, i.e. $\|x\|^2 \geqslant 0$

- A non-null causal vector has $x^0 \neq 0$. The **orientation of a causal vector** is

  - **positive** if $x^0 > 0$
  - **negative** if $x^0 < 0$.

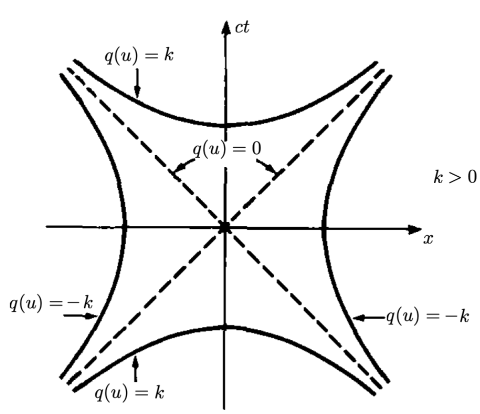
2.6 Likeness in the 2-Lorentz space $\mathbb{R}^{1,1}$

- $u = (ct, x) \in \mathbb{R}^{1,1}$, $q(u) = \|u\|^2 = (ct)^2 - x^2$
- Light-like: $x^2 = (ct)^2 \iff |x| = |ct| \iff u$ lies on the two $\pi/4$ degrees straight lines passing through the origin
- Time-like: $x^2 < (ct)^2 \iff |x| < |ct| \iff u$ belongs to the vertical upper and lower dotted triangular regions.
- Space-like: remaining areas.

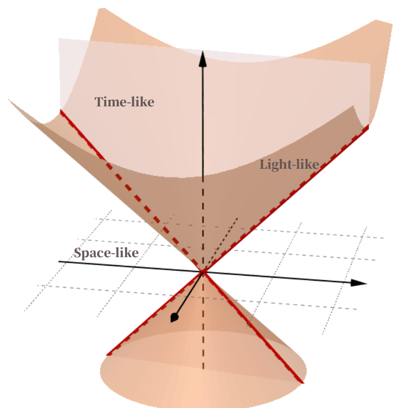


2.7 Connection with hyperbolic geometry

- Level line of q (vectors with same quadratic Lorentz pseudo-norm): **hyperbola** contained in the time-like or space-like regions light-like straight lines as asymptotes
- In fact, for all $k \neq 0$, $q(x) = k \iff (ct)^2 - x^2 = k$, which is the equation of a hyperbola



- $n \geqslant 3$, the light-like equation $\|x\|^2 = 0$ defines a hypercone $\mathcal{C}^{n-1}$ in $\mathbb{R}^n$: **lightcone** $\mathcal{L}$
- Time-like vectors $\in \mathrm{int}(\mathcal{L})$, space-like vectors $\in \mathrm{ext}(\mathcal{L})$



- Positive and negative time-like regions and lightcone: **future** and **past lightcone**
- In the **Euclidean space** E^n, vectors with same Euclidean norm lie on **spheres**
- In the Lorentz space, **vectors with same square Lorentz pseudo-norm** $\neq 0$ **lie on hyperboloids**
- This remarkable difference has very important consequences on the properties of space-time.

2.8 Two quadratic forms with signature $(1, n-1)$ having same null cone are proportional

- Let q_1 and q_2 be two quadratic forms on $\mathbb{R}^n$ with signature $(+, -, \dots, -)$. Assume that they have the same null cone, namely

$$\{x \in \mathbb{R}^n : q_1(x) = 0\} = \{x \in \mathbb{R}^n : q_2(x) = 0\},$$

  we are going to prove that q_1 and q_2 are proportional

- By Sylvester's law of inertia, after an invertible linear change of coordinates, we may assume that

$$q_1(x) = q(x) = (x^0)^2 - |\mathbf{x}|^2,$$

  where

$$x = (x^0, \mathbf{x}), \qquad x^0 \in \mathbb{R}, \qquad \mathbf{x} \in \mathbb{R}^{n-1}$$

- We notice that for every $r > 0$ and every $\mathbf{c} \in \mathbb{R}^{n-1}$ such that $|\mathbf{c}| = 1$, the vectors

$$(r, r\mathbf{c}) \qquad \text{and} \qquad (-r, r\mathbf{c})$$

  belong to the null cone of q_1 since $q_1(r, \mathbf{c}) = q_1(-r, \mathbf{c}) = r^2 - r^2|\mathbf{c}|^2 = 0$

- Now we write the second quadratic form in the general form

$$q_2(x^0, \mathbf{x}) = a(x^0)^2 + 2x^0\mathbf{b} \cdot \mathbf{x} + \mathbf{x}^T C\mathbf{x},$$

  where $a \in \mathbb{R}$, $\mathbf{b} \in \mathbb{R}^{n-1}$, and C is a real symmetric $(n-1) \times (n-1)$ matrix

- Since q_1 and q_2 have the same null cone, we have

$$q_2(r, r\mathbf{c}) = 0, \qquad q_2(-r, r\mathbf{c}) = 0$$

- Computing these two quantities gives

$$q_2(r, r\mathbf{c}) = r^2 \left(a + 2\mathbf{b} \cdot \mathbf{c} + \mathbf{c}^T C\mathbf{c}\right) = 0, \quad q_2(-r, r\mathbf{c}) = r^2 \left(a - 2\mathbf{b} \cdot \mathbf{c} + \mathbf{c}^T C\mathbf{c}\right) = 0$$

- Since $r > 0$, we get

$$a + 2\mathbf{b} \cdot \mathbf{c} + \mathbf{c}^T C\mathbf{c} = 0, \qquad a - 2\mathbf{b} \cdot \mathbf{c} + \mathbf{c}^T C\mathbf{c} = 0,$$

  subtracting these identities yields

$$4\mathbf{b} \cdot \mathbf{c} = 0$$

  for every $\mathbf{c} \in \mathbb{R}^{n-1}$ with $|\mathbf{c}| = 1$, hence $\mathbf{b} = \mathbf{0}$

- If we now *add* the two identities we find

$$2a + 2\mathbf{c}^T C\mathbf{c} = 0 \iff \mathbf{c}^T C\mathbf{c} = -a$$

  for every $\mathbf{c} \in \mathbb{R}^{n-1}$ with $|\mathbf{c}| = 1$

- Now remark that, for every nonzero $\mathbf{v} \in \mathbb{R}^{n-1}$, setting

$$\mathbf{c} = \frac{\mathbf{v}}{|\mathbf{v}|},$$

we obtain

$$\mathbf{v}^T C\mathbf{v} = |\mathbf{v}|^2\mathbf{c}^T C\mathbf{c} = -a|\mathbf{v}|^2 \iff \mathbf{v}^T C\mathbf{v} + a\mathbf{v}^T\mathbf{v} = 0 \iff \mathbf{v}^T(C + aI)\mathbf{v} = 0$$

for every $\mathbf{v} \in \mathbb{R}^{n-1}$

- Since $C + aI$ is symmetric, this implies $C + aI = 0$, and hence $C = -aI$

- Substituting $\mathbf{b} = 0$ and $C = -aI$ into the expression of q_2, we obtain

$$q_2(x^0, \mathbf{x}) = a(x^0)^2 - a|\mathbf{x}|^2 = a\big((x^0)^2 - |\mathbf{x}|^2\big),$$

therefore

$$q_2 = aq_1$$

- q_2 has signature $(+, -, \ldots, -)$, so it is not the zero form and then $a \neq 0$. This proves that q_1 and q_2 are proportional

- This result will be used to prove that the spacetime interval is the same in every inertial reference frame.

2.9 Lorentz-orthogonal vectors

- $x, y \in \mathbb{R}^{1,n-1}$: Lorentz-orthogonal if $x \circ y = 0$

Theorem 2.9.1 *$x, y \in \mathbb{R}^{1,n-1}$, $x, y \neq 0$, equioriented and causal, then $x \circ y \geqslant 0$, specifically*

- *$x \circ y > 0 \iff x, y$ are time-like*
- *$x \circ y = 0 \iff x, y$ are collinear light-like vectors.*

Proof: without loss of generality, assume both $x \equiv (x^0, \mathbf{x})$ and $y \equiv (y^0, \mathbf{y})$ positively oriented: $x^0 > 0$ and $y^0 > 0$. Since x is causal,

$$x \circ x \geqslant 0 \iff |\mathbf{x}|^2 \leqslant (x^0)^2 \underset{x^0>0}{\iff} |\mathbf{x}| \leqslant x^0,$$

and, similarly, $|\mathbf{y}| \leqslant y^0$. By the Cauchy-Schwarz inequality

$$\langle \mathbf{x}, \mathbf{y} \rangle \leqslant |\langle \mathbf{x}, \mathbf{y} \rangle| \leqslant |\mathbf{x}||\mathbf{y}| \leqslant x^0 y^0, \tag{2.1}$$

hence $x \circ y = x^0 y^0 - \langle \mathbf{x}, \mathbf{y} \rangle \geqslant 0$.

The equality $x \circ y = x^0 y^0 - \langle \mathbf{x}, \mathbf{y} \rangle = 0$ implies $x^0 y^0 = \langle \mathbf{x}, \mathbf{y} \rangle > 0$, thus (2.1) becomes

$$\langle \mathbf{x}, \mathbf{y} \rangle \leqslant |\mathbf{x}||\mathbf{y}| \leqslant \langle \mathbf{x}, \mathbf{y} \rangle,$$

i.e.

$$0 < \langle \mathbf{x}, \mathbf{y} \rangle = |\mathbf{x}||\mathbf{y}| \iff \mathbf{y} = t\mathbf{x}, \quad t \neq 0,$$

so

$$x \circ y = 0 \iff x^0 y^0 = t|\mathbf{x}|^2 \iff y^0 = \frac{t|\mathbf{x}|^2}{x^0} \implies \underline{t > 0}, \tag{2.2}$$

hence

$$y \circ y = (y^0)^2 - |\mathbf{y}|^2 \underset{\mathbf{y}=t\mathbf{x}}{=} \frac{t^2|\mathbf{x}|^4}{(x^0)^2} - t^2|\mathbf{x}|^2 \underset{y \text{ causal}}{\geqslant} 0 \iff |\mathbf{x}|^2 \geqslant (x^0)^2 \implies x \circ x \leqslant 0,$$

but x is causal, hence x must be light-like and so

$$\|x\|^2 = 0 \iff (x^0)^2 = |\mathbf{x}|^2 \underset{x^0>0}{\iff} x^0 = |\mathbf{x}|.$$

Thus, the central equation of formula (2.2) becomes $x^0 y^0 = t(x^0)^2 \iff y^0 = t x^0$.

In conclusion, $y = tx$, so x and y are linearly dependent light-like vectors. □

2.10 Lorentz-orthogonality and time-like vectors

- **Corollary 1**: $x, y \in \mathbb{R}^{1,n-1}$, $x, y \neq 0$. If x, y are two equioriented time-like vectors then $x \circ y > 0$, so, in particular, they can't be Lorentz-orthogonal.

- **Corollary 2**: $x, y \in \mathbb{R}^{1,n-1}$, $x, y \neq 0$, x time-like and $x \circ y = 0 \implies y$ is space-like.

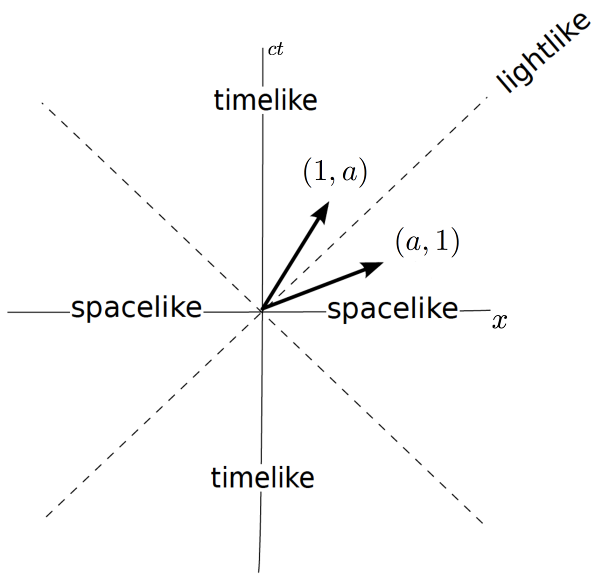
  Proof: assume x, y equioriented (it they're not, replace y by $-y$ and the result doesn't change). If x is time-like, then y can't be time-like because of Corollary 1, or light-like because of Theorem 2.9.1. Hence y must be space-like. □

2.11 Lorentz-orthogonality and space-like vectors

- The reverse statement is not true: **there are many more Lorentz-orthogonal space-like vectors than time-like ones**!

- Example in $\mathbb{R}^{1,2}$: $(0,1,0)$ and $(0,0,1)$ are Lorentz-orthogonal and both space-like

- Two Lorentz-orthogonal vectors of $\mathbb{R}^{1,n-1}$ are no longer characterized by having a relative angle of $\pi/2$. For instance:
  - By theorem 2.9.1 we know that $x, y \in \mathcal{L}$, $x \circ y = 0 \iff$ they are collinear, so two photons moving in different directions are never Lorentz-orthogonal!
  - In $\mathbb{R}^{1,1}$, any couple of vectors $u = (1, a)$ and $w = (a, 1)$ is Lorentz-orthogonal: $u \circ w = a - a = 0$. If $a > 1$:



  - The space-like vectors orthogonal to the time-like vector $u = (ct, x)$ are $w = \lambda(x, ct)$, $\lambda \neq 0$, and they are complementary[2]: their angles ϑ_u, ϑ_w w.r.t. the x-axis sum to $\pi/2$ thanks to the following identity satisfied by the arc tangent:

$$\vartheta_u + \vartheta_w = \arctan\left(\frac{ct}{x}\right) + \arctan\left(\frac{x}{ct}\right) = \frac{\pi}{2} \pmod{\pi}.$$

[2](to be used in section 3.7. . .)

2.12 Lorentz-orthonormality

- $x, y \in \mathbb{R}^{1,n-1}$ are **Lorentz-orthonormal** if they are Lorentz-orthogonal and if the *modulus* of their Lorentz pseudo-norm is 1, (so, only time-like and space-like vectors can be Lorentz-unit vectors)

- Lorentz-orthonormal family in $\mathbb{R}^{1,n-1}$: $m \leqslant n$ mutually Lorentz-orthogonal unit vectors. If $m = n$: Lorentz-orthonormal **basis**

- Results in section 2.10 $\Longrightarrow$ **in every Lorentz-orthonormal basis** of $\mathbb{R}^{1,n-1}$ there is **exactly one time-like vector** (conventionally set in the first place) and $n-1$ **space-like vectors**

- So, a set of n vectors $\mathscr{B} = (u_0, \dots, u_{n-1})$ is a Lorentz-orthonormal basis of $\mathbb{R}^{1,n-1}$ if

$$u_i \circ u_j = \begin{cases} 1 & \text{if} \quad i = j = 0 \\ -1 & \text{if} \quad 1 \leqslant i, j \leqslant n-1, \ i = j \\ 0 & \text{if} \quad i \neq j \end{cases}$$

- Canonical basis $(e_1, \dots, e_n)$ of $\mathbb{R}^n$: Lorentz-orthonormal basis of $\mathbb{R}^{1,n-1}$.

2.13 Lorentz transformations

- Lorentz transformations are for the Lorentz n-space what orthogonal transformations are for the Euclidean n-space

- A **Lorentz transformation** on $\mathbb{R}^{1,n-1}$ is a map $\phi : \mathbb{R}^{1,n-1} \to \mathbb{R}^{1,n-1}$ that preserves the Lorentz pseudo-scalar product, i.e.

$$\boxed{\phi(x) \circ \phi(y) = x \circ y} \qquad \forall x, y \in \mathbb{R}^{1,n-1}$$

- The set of Lorentz transformation on $\mathbb{R}^{1,n-1}$ form a group[3] under composition called **Lorentz group**. Notation: $\mathrm{O}(1, n-1)$ or $\mathscr{L}$

- Theorem: $\phi : \mathbb{R}^{1,n-1} \to \mathbb{R}^{1,n-1}$ is a Lorentz transformation $\iff$ is linear and, given a Lorentz-orthonormal basis $(u_0, \dots, u_{n-1})$ of $\mathbb{R}^{1,n-1}$, $(\phi(u_0), \dots, \phi(u_{n-1}))$ is a Lorentz-orthonormal basis of $\mathbb{R}^{1,n-1}$.

[3] Henri Poincaré was the one who showed, in a 1905 paper *Sur la dynamique de l'électron*, in *Comptes Rendus de l'Académie des Sciences*, that the transformations previously introduced by Lorentz form a group. Historically, this is the **first appearance of group-theoretical arguments in fundamental theoretical physics**. Poincaré's communication was presented *eight days before* the publication of Einstein's memoir in the *Annalen der Physik*!

2.14 The Lorentzian matrix and its properties

- The matrix $\Lambda \in \mathrm{M}(n,\mathbb{R})$ is Lorentzian if the map $\phi_\Lambda : \mathbb{R}^{1,n-1} \to \mathbb{R}^{1,n-1}$, $x \mapsto \phi_\Lambda(x) := \Lambda x$, is a Lorentz transformation

Theorem 2.14.1 *Λ: $n \times n$ real matrix, $\eta = \mathrm{diag}(1,-1,\dots,-1)$. These items are equivalent.*

1. *Λ is a Lorentzian matrix*
2. *Λ^T is a Lorentzian matrix*
3. *The columns of Λ form a Lorentz-orthonormal basis of $\mathbb{R}^{1,n-1}$*
4. *The rows of Λ form a Lorentz-orthonormal basis of $\mathbb{R}^{1,n-1}$*
5. *Λ verifies:*
$$\Lambda^T \eta \Lambda = \eta = \Lambda \eta \Lambda^T$$
6. *Λ is invertible,*
$$\Lambda^{-1} = \eta \Lambda^T \eta$$
and Λ^{-1} is a Lorentzian matrix
7. *Λ preserves the Lorentz quadratic form:*
$$q(\Lambda x) = q(x) \iff \|\Lambda x\|^2 = \|x\|^2 .$$

- Property 7 implies that the
  - time-like regions
  - space-like regions
  - lightcone
  - level surfaces of q, i.e. the hyperboloids $q(x) = q(\Lambda x) = k$, $k \in \mathbb{R}$,

  remain unaltered as a whole *under Lorentz transformations*, despite the fact that time-like, space-like and light-like vectors are modified by them

- If we identify Lorentz transformations ϕ with their Lorentzian matrices Λ, we can identify $\mathscr{L} = \mathrm{O}(1,n-1)$ with a subgroup of $\mathrm{GL}(n,\mathbb{R})$ as follows:
$$\begin{aligned}\mathrm{O}(1,n-1) &= \left\{\phi : \mathbb{R}^n \to \mathbb{R}^n \ : \ \phi(x) \circ \phi(y) = x \circ y, \ \ \forall x,y \in \mathbb{R}^{1,n-1}\right\} \\ &= \left\{\Lambda \in \mathrm{GL}(n,\mathbb{R}) \ : \ \Lambda^T \eta \Lambda = \eta = \Lambda \eta \Lambda^T\right\}\end{aligned}$$

- If we replace $\circ$ by the Euclidean inner product $\langle \ , \ \rangle$ and η by I_n, then $\mathrm{O}(1,n-1)$ becomes the orthogonal group $\mathrm{O}(n)$:
$$\begin{aligned}\mathrm{O}(n) &= \left\{\phi : \mathbb{R}^n \to \mathbb{R}^n \ : \ \langle \phi(x), \phi(y) \rangle = \langle x, y \rangle, \ \ \forall x,y \in \mathbb{R}^n\right\} \\ &= \left\{A \in \mathrm{GL}(n,\mathbb{R}) \ : \ A^T I_n A = I_n = A I_n A^T\right\}.\end{aligned}$$

2.15 Noticeable subgroups of the Lorentz group

- Binet's th. $\Longrightarrow$ $\det(\eta) = \det(\Lambda^T \eta \Lambda) = \det(\Lambda^T)\det(\eta)\det(\Lambda)$, i.e. $\det(\Lambda^T)\det(\Lambda) = 1$, but $\det(\Lambda^T) = \det(\Lambda)$, thus

$$\boxed{\det(\Lambda) = \pm 1} \quad \text{(as for orthogonal matrices)}$$

- We define:

$$\mathrm{SO}(1, n-1) \equiv \mathscr{L}_+ := \big\{\Lambda \in \mathrm{O}(1, n-1) \ : \ \det(\Lambda) = 1\big\},$$

called either **special** or **proper Lorentz group**.

- $\Lambda \in \mathrm{O}(1, n-1)$ is **positively oriented** if it preserves the orientation of time-like vectors: $\forall x \in \mathbb{R}^{1,n-1}$, x time-like,

$$x^0 > 0 \implies (\Lambda x)^0 > 0.$$

Subgroup of $\mathrm{O}(1, n-1)$ of positive Lorentz transformations:

$$\mathrm{PO}(1, n-1) \equiv \mathrm{O}^+(1, n-1) \equiv \mathscr{L}^\uparrow,$$

called either **positive** or **orthochronous** Lorentz group.

2.16 The restricted Lorentz group

- The subgroup

$$\mathrm{SPO}(1, n-1) \equiv \mathrm{SO}^+(1, n-1) \equiv \mathscr{L}_+^\uparrow := \mathrm{SO}(1, n-1) \cap \mathrm{PO}(1, n-1),$$

special positive or **proper orthochronous** or **restricted Lorentz group**

- $\mathscr{L}_+^\uparrow$ is the *connected component to the identity* of the Lorentz group

- **Theorem 2.16.1** *Let* $\Lambda = (\Lambda_{ij})_{0 \leqslant i,j \leqslant n-1} \in \mathrm{O}(1, n-1)$. *Then:*

$$\Lambda \in \mathrm{O}^+(1, n-1) \iff \Lambda_{00} \geqslant 1,$$

i.e. a Lorentz transformation preserves time orientation if and only if the first diagonal element of Λ *is* $\geqslant 1$. *Moreover,*

$$\Lambda \in \mathrm{O}(n) \cap \mathrm{O}^+(1, n-1) = \mathrm{O}(n-1) \iff \Lambda_{00} = 1$$

$$\mathrm{O}(n) \cap \mathrm{O}^+(1, n-1) = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix} : R \in \mathrm{O}(n-1) \right\} \cong \mathrm{O}(n-1),$$

i.e. a Lorentz transformation is an augmented rotation matrix if and only if the first diagonal element of Λ *is* 1.

Chapter 3

Relativistic kinematic

Main references:

- **Landau & Lifshitz**: 'The Classical Theory of Fields' (1971)
- **Barone**: 'Relatività' (2004)
- **Steane**: 'Relativity Made Relatively Easy' (2012)

3.1 Back to the 4D Minkowski space...

- In special relativity $\mathcal{M}$ is interpreted as an affine Minkowski spacetime modeled on the vector space $(\mathbb{R}^4, \eta) \equiv \mathbb{R}^{1,3}$: $\eta_{\mu\nu} = \mathrm{diag}(1, -1, -1, -1)$
- A point of $\mathcal{M}$ is called an **event**: $e = (x^\mu)_{\mu=0,\ldots,3}$, $x^0 = ct$, $\mathbf{x} = (x^i)_{i=1,2,3}$
  - greek indices: spacetime coordinates
  - latin indices: space coordinates
- **Wordline**: sequence of events in time of a particle or light ray moving in $\mathcal{M}$, i.e. a curve $\gamma : I \subset \mathbb{R} \to \mathcal{M}$
- **Displacement**: fixed an origin $O = (0, \mathbf{0}) \in \mathcal{M}$, an event defines the displacement vector $x = e - O \in \mathbb{R}^{1,3}$. In coordinates, $x = (x^\mu)_{\mu=0,\ldots,3}$ is interpreted as a column
- More generally, given two events e_1, e_2, their displacement vector is defined by $e_2 - e_1$
- The **infinitesimal spacetime interval** is the indefinite Minkowski quadratic form of the infinitesimal displacement $dx = (dx^\mu)_{\mu=0,\ldots,3} = (cdt, dx^1, dx^2, dx^3)$:

$$
\begin{aligned}
q(dx) = dx \circ dx &\underset{\text{(sec. 2.4)}}{=} \langle \eta dx, dx \rangle = (\eta dx)^T dx = (dx)^T \eta\, dx \\
&= (cdt, dx^1, dx^2, dx^3) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} cdt \\ dx^1 \\ dx^2 \\ dx^3 \end{pmatrix} = c^2 dt^2 - |d\mathbf{x}|^2
\end{aligned}
$$

- Physics notation: $q(dx)$ is written ds^2 and calculated in coordinates as follows:

$$ds^2 = \sum_{\mu=0}^{3}\sum_{\nu=0}^{3} \eta_{\mu\nu} dx^\mu dx^\nu$$

but $\eta_{\mu\nu} = 0$ for $\mu \neq \nu$, hence the mixed terms in the sum do not contribute and

$$\begin{aligned} ds^2 &= \eta_{00}(dx^0)^2 + \eta_{11}(dx^1)^2 + \eta_{22}(dx^2)^2 + \eta_{33}(dx^3)^2 \\ &= (dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2 = c^2dt^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2 \\ &= c^2dt^2 - |d\mathbf{x}|^2 \end{aligned}$$

- To avoid repeating the sums, **Einstein** proposed his famous **convention**: whenever an index appears once as an upper index and once as a lower index, it is implicitly summed over, so

$$\boxed{ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu = c^2dt^2 - |d\mathbf{x}|^2}\,.$$

3.2 Revisiting likeness regions in the Minkowski spacetime

- $ds^2 = 0 \iff |d\mathbf{x}|^2 = c^2dt^2 \iff$ signal traveling at the speed of light $\implies$ worldline in the lightcone $\mathcal{L}$ in $\mathcal{M}$

- $ds^2 > 0 \iff |d\mathbf{x}|^2 < c^2dt^2$: the spatial separation between events is smaller than the distance traveled by a light ray $\implies$ worldline with speed less than c in the *time-like* $\mathcal{T}$ (causal) region of $\mathcal{M}$. Events belonging to $\mathcal{T}$ may be viewed as evolution *in time* of past events
  - Particular case: if $|d\mathbf{x}| = 0$, $dt \neq 0$ (events occurring *at the same place at different times*) then $ds^2 > 0$
- $ds^2 < 0 \iff |d\mathbf{x}|^2 > c^2dt^2$: spatial separation larger than the distance traveled by a light ray $\implies$ worldline in the *space-like* $\mathcal{S}$ (non-causal) region of $\mathcal{M}$. Events belonging to $\mathcal{S}$ are so far apart *in space* that it's impossible that one can be the evolution of the other
  - Particular case: if $dt = 0$, $|d\mathbf{x}| > 0$ (*simultaneous events occurring at different places*) then $ds^2 < 0$, because if two events take place at the same time in different places, they can't be connected by any causal signal, since that would require an infinite speed
- We will turn back to the division of the Minkowski space soon...

3.3 Invariance of spacetime interval in inertial reference frames

- One of the most important consequences of the postulates of special relativity is contained in the following result

- **Theorem**: the infinitesimal (and so also the finite) spacetime interval between two events is invariant across any two inertial frames $\mathscr{R}$ and $\mathscr{R}'$:

$$\boxed{ds'^2 = ds^2}$$

- **Proof** (à la Einstein).

- We start by proving the claim in the null case, i.e. $ds'^2 = 0 \iff ds^2 = 0$

- By definition of the infinitesimal spacetime interval, $ds^2 = c^2dt^2 - |d\mathbf{x}|^2$, therefore,

$$ds^2 = 0 \iff c^2dt^2 - |d\mathbf{x}|^2 = 0 \iff \frac{|d\mathbf{x}|^2}{dt^2} = c^2,$$

so, $ds^2 = 0$ means that the two infinitesimally close events are connected by a ray of light in the inertial frame $\mathscr{R}$

- By the postulate of constancy of the speed of light, the same light signal has speed c in every other inertial frame $\mathscr{R}'$. Hence

$$ds^2 = 0 \iff \frac{|d\mathbf{x}|^2}{dt^2} = c^2 \iff \frac{|d\mathbf{x}'|^2}{dt'^2} = c^2 \iff ds'^2 = c^2dt'^2 - |d\mathbf{x}'|^2 = 0,$$

so, in particular, the two quadratic forms ds^2 and ds'^2 have the same null cone

- We know that, since inertial-frame transformations preserve inertial motion, they are affine transformations of spacetime coordinates:

$$x' = Ax + b,$$

passing to infinitesimal displacements, the translation term disappears, thus giving

$$dx' = Adx$$

- Hence, the transformation acts linearly on the differentials and therefore ds'^2 is a homogeneous quadratic form in the differentials dt and $d\mathbf{x}$

- In section 2.8 we have proven that two Lorentzian quadratic forms with the same null cone are proportional, so ds'^2 and ds^2 can differ only by a multiplicative factor depending on the pair of inertial frames:

$$ds'^2 = \alpha_{\mathscr{R},\mathscr{R}'} ds^2$$

- $\alpha_{\mathscr{R},\mathscr{R}'}$ can't depend on time or space coordinates: different points in space and instants in time wouldn't be equivalent ... against postulate 1

- Only difference $\mathscr{R}$ vs. $\mathscr{R}'$: their relative motion, so $\alpha_{\mathscr{R},\mathscr{R}'}$ can only be a function of $\mathbf{v}$
- Due to the postulate of isotropy of space, α cannot depend on the *velocity* vector $\mathbf{v}$, but only on the *speed* v

$$\implies \alpha_{\mathscr{R},\mathscr{R}'} = \alpha(v) \implies ds'^2 = \alpha(v) ds^2$$

- To prove that $\alpha(v) \equiv 1$, consider the inverse transformation (from $\mathscr{R}'$ back to $\mathscr{R}$). The relative velocity is now $-\mathbf{v}$, so using again the isotropy of space and the independence on the spatial orientation we have

$$ds^2 = \alpha(v)\, ds'^2$$

- Combining the two relations, we obtain

$$ds^2 = \alpha(v)\, ds'^2 = \alpha(v)\alpha(v)\, ds^2,$$

  hence

$$\alpha(v)^2 = 1$$

- By continuity (and the fact that for $v = 0$ the transformation is the identity, so $\alpha(0) = 1$), we must take the positive branch:

$$\alpha(v) \equiv 1.$$

□

- $ds'^2 = ds^2$ has a stunning number of very important consequences ...

3.4 Lorentz transformations in special relativity

- Consider two inertial frames $\mathscr{R}$, $\mathscr{R}'$ with parallel axes and define the coordinate transformation $T : \mathbb{R}^{1,3} \to \mathbb{R}^{1,3}$, such that $x' = T(x)$

- Postulate 1 (Galilean relativity): $x' = T(x) = \Lambda x + a$, $\Lambda \in M(4, \mathbb{R})$, $a \in \mathbb{R}^{1,3}$ (shift vector)

- We use Postulate 2 to determine Λ. Constancy of the speed of light $\implies$ invariance of the spacetime interval, i.e. of the quadratic Minkowski form ($dx' = \Lambda dx$, so the burden of the preservation of the quadratic Minkowski form falls only on Λ)

- Preservation of the spacetime interval $\implies$ Λ **belongs to the Lorentz group**:

$$x' = \Lambda x + a, \quad \Lambda \in \mathrm{O}(1,3),\ a \in \mathbb{R}^{1,3}$$

- The set of these transformations is the **Poincaré group** defined by

$$\mathscr{P} = \{(\Lambda, a)\ :\ \Lambda \in \mathrm{O}(1,3),\ a \in \mathbb{R}^{1,3}\},$$

  with composition law $(\Lambda_1, a_1) \cdot (\Lambda_2, a_2) = (\Lambda_1 \Lambda_2, a_1 + \Lambda_1 a_2)$.

- The Poincaré group contains the Lorentz group and the translations as subgroups:

  - $\mathrm{O}(1,3)$ is isomorphic to the subgroup of $\mathscr{P}$ given by the elements $(\Lambda, \mathbf{0})$
  - The elements of $\mathscr{P}$ of the form (I_4, a) are translations by the constant vector a

- The affine coordinate transformations $x' = \Lambda x + a$ are called **Poincaré transformations**, and those corresponding to $a = \mathbf{0}$, i.e. the linear functions $x' = \Lambda x$, are of course the **Lorentz transformations**

- A bit of **history**: timeline of the discovery of Lorentz transformations

- **1887**: first discovered by **Voigt** working on the Doppler effect in elastic media

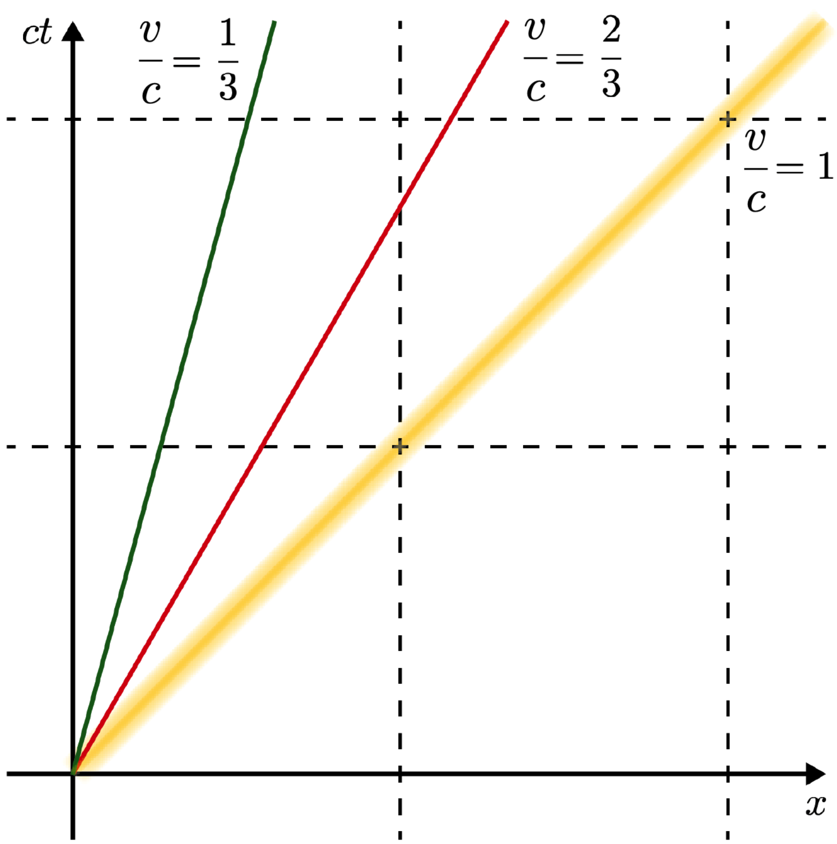
- **1892-1904**: **Lorentz** rediscovered independently (1892) the same transformations and further refined their interpretation up to length contraction and time dilation (1904)

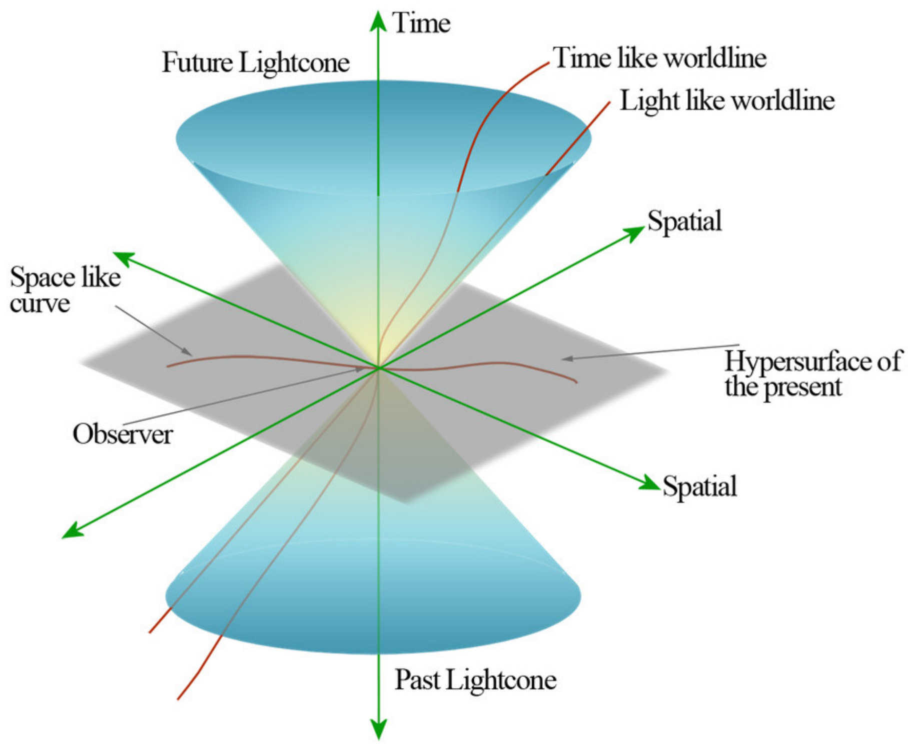
- **1905**: **Poincaré** recognized the group structure under Lorentz transformations and gave them their modern name. He also suggested that physics should be invariant under these transformations

- **1905**: **Einstein** rederived Lorentz equations from the special relativity principles.

3.5 Minkowski diagrams (1907)

- Minkowski diagram: 2-dimensional *spacetime diagram*, vertical axis parameterized by ct, horizontal by one spatial coordinate, typically x
- Worldline of a ray of light: straight line with a slope $\pi/2$
- Worldline of a particle moving uniformly with speed $v < c$: straight line with slope v/c, the fraction of the speed of light
- Worldline of a particle at rest in a position x of the inertial frame: vertical line, in particular the ct axis is the wordline of a particle at rest in $x = 0$
- Horizontal line (at constant t): line of *simultaneity*, events happening at the instant t, so the x-axis is the set of events happening at time $t = 0$
- Origin: represents an event occurring at time $t = 0$ and position $x = 0$.





- The events lying in the upper quadrant between the two bisectors ($c^2t^2 - x^2 > 0$ with $t > 0$) can be reached by a signal sent from the origin O

- Outside this quadrant, which, in the case of more than one spatial dimension, becomes a cone and is therefore called the *future cone*, there is no future event located in the half-plane $t > 0$ that is accessible from O

- The Lorentz transformations preserve (as a whole) the different regions of $\mathcal{M}$, so the future lightcone has an absolute (non-relative) nature, it is the **region of absolute future** of O in special relativity

- Similarly, the events lying in the lower quadrant between the two bisectors ($c^2t^2 - x^2 > 0$ with $t < 0$) are events in the past that are causally connected with the origin: they may have influenced the event O by sending a signal or through an interaction

- They live in the *past cone*, which represents the **region of absolute past** of O

- Therefore, the worldline of a particle that at time $t = 0$ is located at O must lie entirely within the two (past and future) lightcones accessible from O

- The events lying on the two bisectors (that is, on the *lightcones*) are future events (for $t > 0$) or past events (for $t < 0$) of interactions propagating at the speed of light

- We talk about *future lightcone* and *past lightcone* when we consider propagation of both particles that travel with subluminal and luminal speed, these are the absolute concepts in special relativity that replace the absolute time of Newtonian mechanics: time is not absolute anymore, but the futur and past lightcones are!

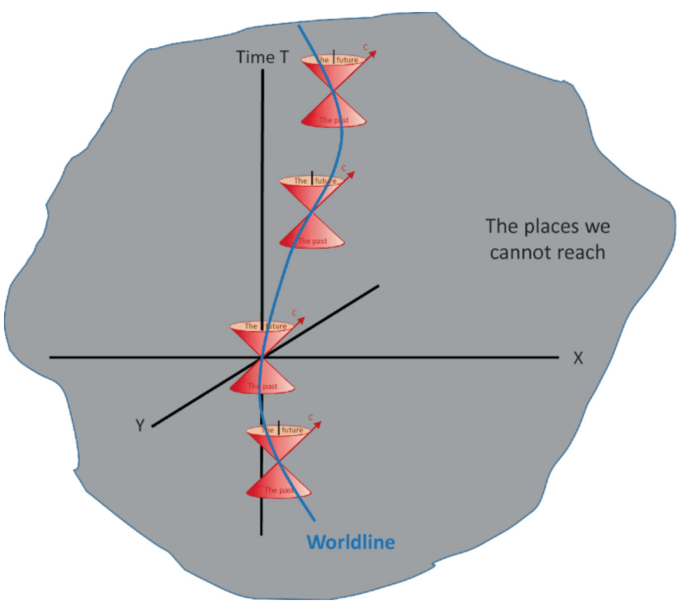
- Finally, the events lying outside both the future and past lightcones (that is, the events for which $c^2t^2 - x^2 < 0$) are separated from the origin by space-like intervals: as we will show in section 3.29, it is always possible to find inertial reference frames in which they are simultaneous with the origin event

- The region between the two lightcones therefore represents the **present** of O (and also the **absolute elsewhere** of O, since in no reference frame an event belonging to this region and the event O can occur at the same spatial point)

- To summarize, the existence of a limit velocity implies that, in relativity, the past and the future of O are **restricted**[1] compared to the non-relativistic case, in which there is no limit on the speed of signals, and therefore the past and the future are, respectively, the half-planes $t < 0$ and $t > 0$:

$$\text{Special relativity} \quad \begin{cases} \text{Future:} & |ct| > |x| \quad t > 0 \\ \text{Present:} & |ct| < |x| \\ \text{Past:} & |ct| > |x| \quad t < 0 \end{cases} \qquad \text{Newtonian mechanics} \quad \begin{cases} \text{Future:} & t > 0 \\ \text{Present:} & t = 0 \\ \text{Past:} & t < 0 \end{cases}$$

- The discussion carried out so far for the origin O actually applies to all events. Each event is associated with a future lightcone and a past lightcone as shown in the figure below (Wouter Schmitz, *Understanding Relativity*, Springer, 2022).



[1]hence the name restricted relativity!

3.6 Lorentz boost (in the standard configuration)

- Boosts: **Lorentz transformations between reference frames with parallel axes**. The coordinate transformation takes place only if the velocity vector $\mathbf{v}$ has a non-null component along the axis, the coordinates of axis $\perp$ to $\mathbf{v}$ are left unchanged

- Coordinate modification of a Lorentz boost in the *standard configuration*: $\mathscr{R}$ and $\mathscr{R}'$ have parallel axes and $\mathscr{R}'$ moves along the x-axis with speed v:

$$\begin{pmatrix} ct' \\ x' \end{pmatrix} = \mathcal{B}(v) \begin{pmatrix} ct \\ x \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix} \implies \begin{cases} ct' = Act + Bx \\ x' = Cct + Dx \quad (\star) \end{cases} \tag{3.1}$$

- The location of the $\mathscr{R}$ origin, $x = 0$, is seen by the $\mathscr{R}'$ frame to evolve in time as $x' = -vt'$. Setting $x = 0$ and $x' = vt$ in the previous system gives

$$\begin{cases} ct' = Act \iff t' = At \\ -vt' = Cct \end{cases} \iff -vAt = Cct \iff C = -\frac{A}{c}v$$

- Analogously, the location of the $\mathscr{R}'$ origin, $x' = 0$, is seen by the $\mathscr{R}$ frame to evolve in time as $x = vt$. Eq. $(\star)$ with $x' = 0$ and $x = vt$ gives:

$$0 = Cct + Dvt \iff C = -\frac{D}{c}v$$

- To summarize: $D = A$ and $C = -\frac{A}{c}v$

- A light-speed signal in $\mathscr{R}$ must also have light-speed in $\mathscr{R}'$, so $x = ct$ implies $x' = ct'$ and system (3.1) gives

$$\begin{cases} ct' = Act + Bct \\ ct' = Cct + Dct \end{cases} \iff A + B = C + D \iff B = C$$

- So, system (3.1) can be written as:

$$\begin{cases} ct' = Act + Bx = Act + Cx = Act - \frac{A}{c}vx = A\left(ct - \frac{v}{c}x\right) \\ x' = Cct + Dx = -\frac{A}{c}vct + Ax = A(-vt + x) \end{cases}$$

- That is:

$$\begin{cases} ct' = A\left(ct - \frac{v}{c}x\right) \\ x' = A(-vt + x) \end{cases} \iff \begin{pmatrix} ct' \\ x' \end{pmatrix} = A \begin{pmatrix} 1 & -\frac{v}{c} \\ -\frac{v}{c} & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix}$$

- The inverse matrix is:

$$\frac{1}{A\left(1 - \frac{v^2}{c^2}\right)} \begin{pmatrix} 1 & \frac{v}{c} \\ \frac{v}{c} & 1 \end{pmatrix} \implies \begin{pmatrix} ct \\ x \end{pmatrix} = \frac{1}{A\left(1 - \frac{v^2}{c^2}\right)} \begin{pmatrix} 1 & \frac{v}{c} \\ \frac{v}{c} & 1 \end{pmatrix} \begin{pmatrix} ct' \\ x' \end{pmatrix}$$

- We have obtained:

$$\begin{cases} ct' = A\left(ct - \frac{v}{c}x\right) \\ x' = A(-vt + x) \end{cases} \quad \text{and} \quad \begin{cases} ct = \frac{1}{A\left(1-\frac{v^2}{c^2}\right)}\left(ct' + \frac{v}{c}x'\right) \\ x = \frac{1}{A\left(1-\frac{v^2}{c^2}\right)}(vt' + x') \end{cases}$$

- By symmetry, the inverse Lorentz transformations are the same as the direct ones, except for the *sign swap of* v and the *exchange primed-unprimed coordinates*, this implies[2]:

$$A = \frac{1}{A\left(1 - \frac{v^2}{c^2}\right)} \quad \Longleftrightarrow \quad A = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} = \gamma$$

- So, the Lorentz boost in the standard configuration[3] is:

$$\begin{cases} t' = \gamma\left(t - \frac{v}{c^2}x\right) \\ x' = \gamma(x - vt) \\ y' = y \\ z' = z \end{cases} \quad \text{and} \quad \begin{cases} t = \gamma\left(t' + \frac{v}{c^2}x'\right) \\ x = \gamma(x' + vt') \\ y = y' \\ z = z' \end{cases}$$

- Galilean limit: $c \to +\infty$, so $\gamma \to 1$ and

$$\begin{cases} t' = \gamma\left(t - \frac{v}{c^2}x\right) \\ x' = \gamma(x - vt) \\ y' = y \\ z' = z \end{cases} \longrightarrow \begin{cases} t' = t \\ x' = x - vt \\ y' = y \\ z' = z \end{cases}$$

- The boost matrix tends to the Galilean matrix:

$$\begin{pmatrix} \gamma & \gamma\frac{v}{c^2} & 0 & 0 \\ -\gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \longrightarrow \mathcal{G}(v) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -v & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

[2] we can exclude $-\gamma$ using the Newtonian approximation.

[3] as we'll see in section 3.17, $\frac{v}{c^2}x$ is responsible for the relativity of simultaneity.

3.7 Effect of a Lorentz boosts on the axes of the Minkowski diagrams

- In the Minkowski diagram, the ct-axis is defined by imposing $x = 0$, and the x-axis is defined by imposing $t = 0$. They are Lorentz-orthogonal to each other and the angle they form with the worldline of a light ray is $\pi/4$

- 2-dimensional x-Lorentz boost

$$\begin{cases} ct' = \gamma(ct - \beta x) \\ x' = \gamma(x - \beta ct) \end{cases} \qquad \boxed{\beta := \frac{v}{c}}$$

- To find the ct'-axis we impose $x' = 0 \iff \beta ct = x$, which is a straight line in the Minkowski diagram with slope

$$\tan(\theta_{ct'}) = \frac{ct}{x} = \frac{1}{\beta} \iff \theta_{ct'} = \arctan\frac{1}{\beta}$$

- Similarly, to find the x'-axis we impose $t' = 0 \iff ct = \beta x$, which is a straight line in the Minkowski diagram with slope

$$\tan(\theta_{x'}) = \frac{ct}{x} = \beta \iff \theta_{x'} = \arctan\beta$$

- Using the identity

$$\arctan\frac{1}{\beta} = \frac{\pi}{2} - \arctan(\beta), \qquad (\beta > 0),$$

  we obtain

$$\theta_{ct'} = \frac{\pi}{2} - \theta_{x'} \tag{3.2}$$

  so ct' and x' are tilted by the same amount w.r.t. the original axes ct and x

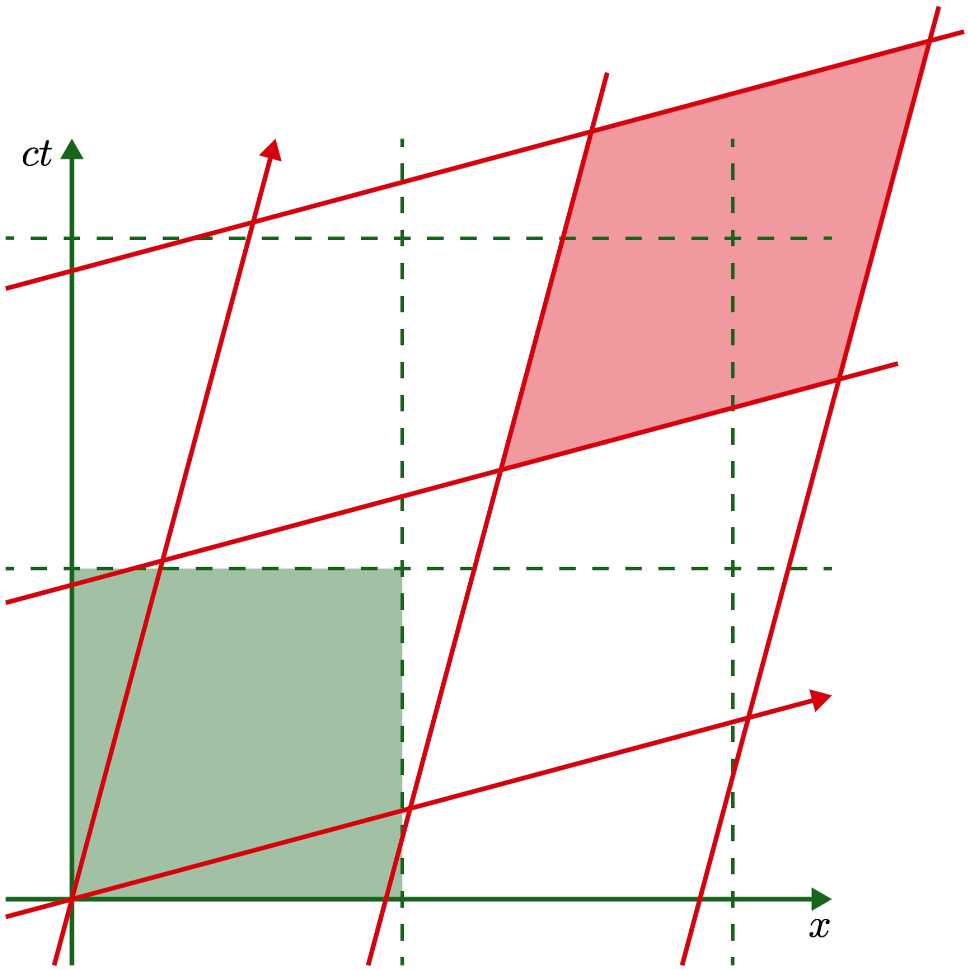
- Moreover, $\theta_{ct'} + \theta_{x'} = \frac{\pi}{2}$, so the axes are Lorentz-orthogonal (see section 2.11) as it should be since Lorentz boost preserves (by definition) Lorentz-orthogonality!

- Finally, from (3.2) we get

$$\theta_{ct'} - \frac{\pi}{4} = \frac{\pi}{4} - \theta_{x'},$$

  and $\pi/4$ is the angle that the worldline of a light ray forms with the x-axis. Therefore, the axes ct' and x' are tilted by the same angle (in magnitude) with respect to the light cone, i.e. they are symmetric with respect to the line $ct = x$.

- **Example 1**: $\beta = 1/3$, so $\gamma \simeq 1.06$ and $\theta_{x'} \approx 18^\circ$

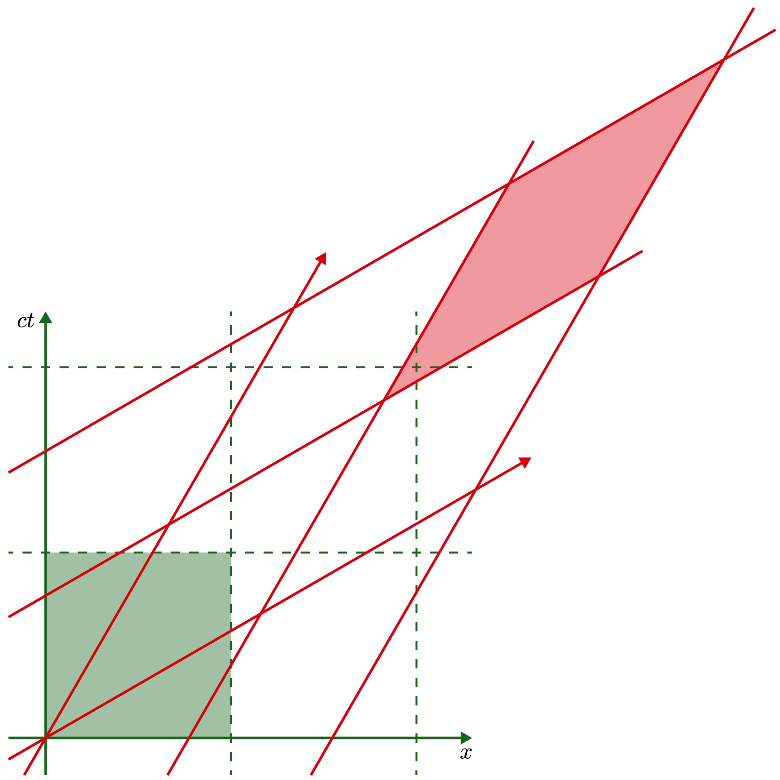


- β measures the relativistic effects[4], as $\beta = v/c$ approaches 1, these become more and more important and the axes become more and more tilted towards the worldine of the light ray

- $v = 0.999\ldots c$: **ultra-relativistic** regime.

[4]A *first order* relativistic effect is proportional to β, a *second order* relativistic effect is proportional to $\beta^2 \ldots$

- **Example 2**: $\beta = 2/3$, so $\gamma \simeq 1.34$ and $\theta_{x'} \approx 34°$



3.8 The hidden role of the Lorentz factor γ in Minkowski diagrams

- Area(green square)=Area(red parallelogram), why?

- Lorentz factor γ: **scaling coefficient** appearing in the transformation of both ct and x

- Recall: 2-dimensional Lorentzian boost matrix

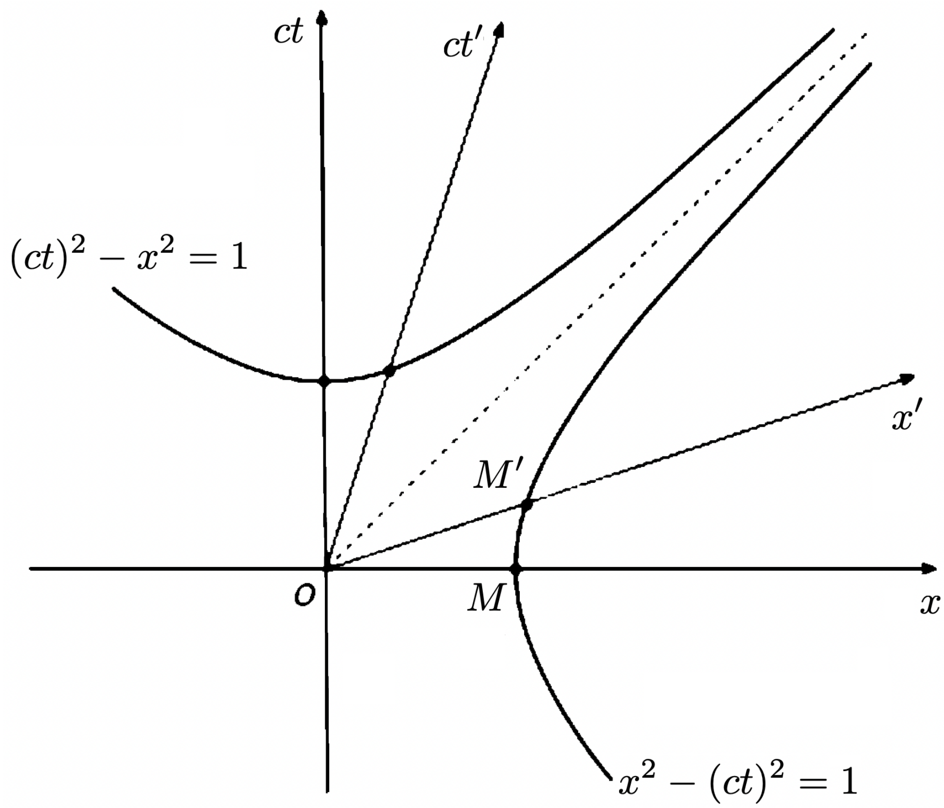
$$\mathcal{B}(\beta) = \begin{pmatrix} \gamma & -\beta\gamma \\ -\beta\gamma & \gamma \end{pmatrix}$$

- $|\det(\mathcal{B}(\beta))|$: area change of a coordinate grid box after the Lorentz boost

$$\det(\mathcal{B}(\beta)) = \gamma^2(1-\beta^2) = \frac{1-(v/c)^2}{1-(v/c)^2} = 1$$

- We obtain 1 thanks to the scaling operated by the Lorentz factor, so: γ is responsible for the preservation of the area of any grid box in a Minkowski diagram.

3.9 Relating the scales on the axes of a Minkowski diagram



- Let us relate the scales on the axes corresponding to the reference frames $\mathscr{R}$ and $\mathscr{R}'$

- Let M be the event with coordinates $(1, 0)$ in the coordinate system (x, ct). By invariance of the spacetime interval,

$$x'^2 - (ct')^2 = x^2 - (ct)^2 = 1$$

- The x'-axis is defined by $ct' = 0$, so the previous equation implies $x'^2 = 1$, and the point M' has coordinates $(1, 0)$ also in the coordinate system (x', ct')

- However, in the Euclidean geometry of the diagram, it is clear that $|\overline{OM'}| \neq |\overline{OM}| = 1$. To relate these scales, we use the fact that M' belongs to the intersection of the hyperbola $x^2 - (ct)^2 = 1$ with the x'-axis

- The x'-axis is characterized by $ct' = 0$, but the Lorentz transformation gives

$$ct' = \gamma(ct - \beta x),$$

so the equation of x'-axis in terms of x and ct is $ct = \beta x$

- The intersection of the hyperbola $x^2 - (ct)^2 = 1$ with the x'-axis is then described by the system:

$$\begin{cases} x^2 - (ct)^2 = 1 \\ ct = \beta x \end{cases} \iff \begin{cases} x^2 - \beta^2 x^2 = 1 \\ ct = \beta x \end{cases} \iff \begin{cases} x^2(1-\beta^2) = 1 \\ ct = \beta x \end{cases},$$

hence, since we are considering $x > 0$,

$$x = \frac{1}{\sqrt{1-\beta^2}}, \qquad ct = \frac{\beta}{\sqrt{1-\beta^2}}$$

- Therefore, M' can be expressed in (x, ct) coordinates as

$$M' = \left(\frac{1}{\sqrt{1-\beta^2}}, \frac{\beta}{\sqrt{1-\beta^2}} \right),$$

so

$$|\overline{OM'}|^2 = x^2 + (ct)^2 = \frac{1}{1-\beta^2} + \frac{\beta^2}{1-\beta^2} = \frac{1+\beta^2}{1-\beta^2},$$

and thus

$$|\overline{OM'}| = \sqrt{\frac{1+\beta^2}{1-\beta^2}}$$

- For example, if $\beta = 2/3$,

$$|\overline{OM'}| = \sqrt{\frac{13}{5}} \approx 1.6,$$

for the ultrarelavistic speed reached by the Large Hadron Collider we have $\beta = 0.99999999$, in this case

$$|\overline{OM'}| \approx 7071.1$$

- Of course, an analogous reasoning holds for the ct and ct' axes.

3.10 Boost along a general direction and the commutativity condition

- If the relative velocity vector $\mathbf{v} = (v_x, v_y, v_z)^T \in \mathbb{R}^3$ has a generic orientation, then we can obtain the boost along $\mathbf{v}$ by decomposing the spatial displacement $\mathbf{x} = (x, y, z)^T$ of an event e in the reference frame $\mathscr{R}$ as $\mathbf{x} = \mathbf{x}_\parallel + \mathbf{x}_\perp$, where $\mathbf{x}_\parallel$ and $\mathbf{x}_\perp$ are the parallel and orthogonal projections w.r.t. $\mathbf{v}$

- A Lorentz transformation acts as a boost on $\mathbf{x}_\parallel$ and doesn't affect $\mathbf{x}_\perp$, so, by linearity,

$$\mathbf{x}' = \gamma(\mathbf{x}_\parallel - \mathbf{v}t) + \mathbf{x}_\perp = \gamma(\mathbf{x}_\parallel - \mathbf{v}t) + \mathbf{x} - \mathbf{x}_\parallel = -\gamma\mathbf{v}t + \mathbf{x} + (\gamma - 1)\mathbf{x}_\parallel,$$

but

$$\mathbf{x}_\parallel = \frac{\mathbf{x} \cdot \mathbf{v}}{|\mathbf{v}|^2}\mathbf{v},$$

so the transformation of $\mathbf{x}$ is

$$\mathbf{x}' = \mathbf{x} + \left[(\gamma - 1)\frac{\mathbf{x} \cdot \mathbf{v}}{|\mathbf{v}|^2} - \gamma t\right]\mathbf{v} \tag{3.3}$$

- For the time coordinate we observe that, in the standard configuration, ct transforms as $ct' = \gamma(ct - (v/c)x)$, so here we must replace x by the component of $\mathbf{x}$ in the $\mathbf{v}$ direction:

$$ct' = \gamma\left(ct - \frac{\mathbf{x} \cdot \mathbf{v}}{c}\right) \tag{3.4}$$

- With tedious calculations, we find that the matrix representation of the boost in the general direction $\mathbf{v}$ is

$$\mathcal{B}(\mathbf{v}) = \begin{pmatrix} \gamma & -\frac{\gamma}{c}\mathbf{v}^T \\ -\frac{\gamma}{c}\mathbf{v} & I_3 + \frac{\gamma^2}{c^2(1+\gamma)}\mathbf{v}\mathbf{v}^T \end{pmatrix}, \quad \mathbf{v}\mathbf{v}^T = \begin{pmatrix} v_x^2 & v_xv_y & v_xv_z \\ v_xv_y & v_y^2 & v_yv_z \\ v_xv_z & v_yv_z & v_z^2 \end{pmatrix} \tag{3.5}$$

$$B(\mathbf{v})^{-1} = B(-\mathbf{v}) = \begin{pmatrix} \gamma & \frac{\gamma}{c}\mathbf{v}^T \\ \frac{\gamma}{c}\mathbf{v} & I_3 + \frac{\gamma^2}{c^2(1+\gamma)}\mathbf{v}\mathbf{v}^T \end{pmatrix} \tag{3.6}$$

- Since the first diagonal matrix element of $\mathcal{B}(\mathbf{v})$ is $\gamma \geqslant 1$ and its determinant is 1, **boosts are proper orthochronous Lorentz transformations**

- We have

$$[\mathcal{B}(\mathbf{v}), \mathcal{B}(\mathbf{w})] := \mathcal{B}(\mathbf{v})\mathcal{B}(\mathbf{w}) - \mathcal{B}(\mathbf{w})\mathcal{B}(\mathbf{v}) = \begin{pmatrix} 0 & \star \\ * & \bullet \end{pmatrix},$$

and all nonzero entry is a linear combination of terms containing $\mathbf{v}\mathbf{w}^T - \mathbf{w}\mathbf{v}^T$.

- But $\mathbf{v}\mathbf{w}^T - \mathbf{w}\mathbf{v}^T = \mathbf{0}$ if and only if $\mathbf{w} = \lambda\mathbf{v}$ for some $\lambda \in \mathbb{R}$. In fact, the matrix $\mathbf{v}\mathbf{w}^T$ has entries v_iw_j, so equality with $\mathbf{w}\mathbf{v}^T$ implies

$$v_iw_j = w_iv_j, \quad \forall i, j,$$

which is equivalent to collinearity.

- Therefore, two boosts $\mathcal{B}(\mathbf{v})$, $\mathcal{B}(\mathbf{w})$ commute if and only if their velocities are collinear.

3.11 Comparison between the general Lorentz boost and the general Galilean transformation

- For the comparison, consider the transformation $t \mapsto t'$ and not $ct \mapsto ct'$:

$$\mathcal{B}(\mathbf{v}) = \begin{pmatrix} \gamma & -\frac{v_x}{c^2}\gamma & -\frac{v_y}{c^2}\gamma & -\frac{v_z}{c^2}\gamma \\ -\gamma v_x & 1+\frac{\gamma^2}{c^2(1+\gamma)}v_x^2 & \frac{\gamma^2}{c^2(1+\gamma)}v_x v_y & \frac{\gamma^2}{c^2(1+\gamma)}v_x v_z \\ -\gamma v_y & \frac{\gamma^2}{c^2(1+\gamma)}v_y v_x & 1+\frac{\gamma^2}{c^2(1+\gamma)}v_y^2 & \frac{\gamma^2}{c^2(1+\gamma)}v_y v_z \\ -\gamma v_z & \frac{\gamma^2}{c^2(1+\gamma)}v_z v_x & \frac{\gamma^2}{c^2(1+\gamma)}v_z v_y & 1+\frac{\gamma^2}{c^2(1+\gamma)}v_z^2 \end{pmatrix}$$

- It tends of course to the general Galilean matrix in the Galilean limit:

$$\mathcal{G}(\mathbf{v}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -v_x & 1 & 0 & 0 \\ -v_y & 0 & 1 & 0 \\ -v_z & 0 & 0 & 1 \end{pmatrix}$$

  (see section 1.17), but it's much more complicated because of the interdependence between time and space coordinates.

3.12 The most general Lorentz transformations

- Recall that boosts are Lorentz transformations for reference frames with parallel axes

- Dropping the assumption of parallel axes, one obtains the most general homogeneous Lorentz transformations, which combine the boost just found with a rotation represented by an orthogonal matrix R. Their form is

$$\mathbf{x}' = R\mathbf{x} + \left[(\gamma-1)\frac{R\mathbf{x}\cdot\mathbf{v}}{v^2} - \gamma t\right]\mathbf{v},$$
$$t' = \gamma\left(t - \frac{R\mathbf{x}\cdot\mathbf{v}}{c^2}\right)$$

- Here we have first applied the rotation R, and then the boost $\mathcal{B}(\mathbf{v})$. The same Lorentz transformation could also be obtained by reversing the order of the operations, that is, by first performing a boost $\mathcal{B}(\mathbf{v}')$ and then a rotation R'. Note that, in general, $\mathbf{v}' \neq \mathbf{v}$ and $R' \neq R$

- Finally, if we do not require that the two reference frames $\mathscr{R}$ and $\mathscr{R}'$ have coincident origins at the instant $t = t' = 0$, we obtain the most general inhomogeneous Lorentz transformations (Poincaré transformations), which incorporate a spatial translation by $\mathbf{x}_0$ and a time translation by t_0, and are given by

$$\mathbf{x}' = R\mathbf{x} + \left[(\gamma-1)\frac{R\mathbf{x}\cdot\mathbf{v}}{v^2} - \gamma t\right]\mathbf{v} + \mathbf{x}_0,$$
$$t' = \gamma\left(t - \frac{R\mathbf{x}\cdot\mathbf{v}}{c^2}\right) + t_0.$$

3.13 Lorentz-invariance of the wave equation: compatibility between Maxwell electromagnetism and special relativity

- In one spatial dimension, the d'Alembert operator is

$$\Box := \frac{\partial^2}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}$$

We start showing that $\Box$ is invariant under Lorentz transformations.

- Consider the Lorentz transformation

$$\begin{cases} x' = \gamma(x - vt), \\ t' = \gamma\left(t - \frac{v}{c^2}x\right), \end{cases} \qquad \gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}.$$

By the chain rule,

$$\frac{\partial}{\partial x} = \frac{\partial x'}{\partial x}\frac{\partial}{\partial x'} + \frac{\partial t'}{\partial x}\frac{\partial}{\partial t'} = \gamma\frac{\partial}{\partial x'} - \gamma\frac{v}{c^2}\frac{\partial}{\partial t'},$$

and

$$\frac{\partial}{\partial t} = \frac{\partial x'}{\partial t}\frac{\partial}{\partial x'} + \frac{\partial t'}{\partial t}\frac{\partial}{\partial t'} = -\gamma v\frac{\partial}{\partial x'} + \gamma\frac{\partial}{\partial t'}.$$

Hence

$$\partial_x = \gamma\left(\partial_{x'} - \frac{v}{c^2}\partial_{t'}\right), \qquad \partial_t = \gamma\left(-v\partial_{x'} + \partial_{t'}\right)$$

- Let f be a smooth function. Applying ∂_x twice, we obtain

$$\begin{aligned} \partial_x^2 f &= \partial_x\left[\gamma\left(\partial_{x'} - \frac{v}{c^2}\partial_{t'}\right)f\right] \\ &= \gamma\left(\partial_{x'} - \frac{v}{c^2}\partial_{t'}\right)\left[\gamma\left(\partial_{x'} - \frac{v}{c^2}\partial_{t'}\right)f\right] \\ &= \gamma^2\left[\partial_{x'}\left(\partial_{x'} - \frac{v}{c^2}\partial_{t'}\right)f - \frac{v}{c^2}\partial_{t'}\left(\partial_{x'} - \frac{v}{c^2}\partial_{t'}\right)f\right] \\ &= \gamma^2\left[\partial_{x'}^2 f - \frac{v}{c^2}\partial_{x'}\partial_{t'}f - \frac{v}{c^2}\partial_{t'}\partial_{x'}f + \frac{v^2}{c^4}\partial_{t'}^2 f\right] \end{aligned}$$

since mixed partial derivatives commute, we get

$$\partial_x^2 f = \gamma^2\left[\partial_{x'}^2 f - 2\frac{v}{c^2}\partial_{x'}\partial_{t'}f + \frac{v^2}{c^4}\partial_{t'}^2 f\right]$$

- Similarly, applying ∂_t twice, we get

$$\begin{aligned} \partial_t^2 f &= \partial_t[\gamma\left(-v\partial_{x'} + \partial_{t'}\right)f] \\ &= \gamma\left(-v\partial_{x'} + \partial_{t'}\right)[\gamma\left(-v\partial_{x'} + \partial_{t'}\right)f] \\ &= \gamma^2\left[-v\partial_{x'}\left(-v\partial_{x'} + \partial_{t'}\right)f + \partial_{t'}\left(-v\partial_{x'} + \partial_{t'}\right)f\right] \\ &= \gamma^2\left[v^2\partial_{x'}^2 f - v\partial_{x'}\partial_{t'}f - v\partial_{t'}\partial_{x'}f + \partial_{t'}^2 f\right] \end{aligned}$$

Again, since mixed partial derivatives commute,

$$\partial_t^2 f = \gamma^2\left[v^2\partial_{x'}^2 f - 2v\partial_{x'}\partial_{t'}f + \partial_{t'}^2 f\right]$$

- Therefore

$$\begin{aligned}\Box f &= \partial_x^2 f - \frac{1}{c^2}\partial_t^2 f \\ &= \gamma^2\left[\partial_{x'}^2 f - 2\frac{v}{c^2}\partial_{x'}\partial_{t'} f + \frac{v^2}{c^4}\partial_{t'}^2 f\right] - \frac{\gamma^2}{c^2}\left[v^2\partial_{x'}^2 f - 2v\partial_{x'}\partial_{t'} f + \partial_{t'}^2 f\right]\end{aligned}$$

we see that the mixed terms cancel, so

$$\Box f = \gamma^2\left(1 - \frac{v^2}{c^2}\right)\partial_{x'}^2 f - \frac{\gamma^2}{c^2}\left(1 - \frac{v^2}{c^2}\right)\partial_{t'}^2 f.$$

- Since

$$\gamma^2\left(1 - \frac{v^2}{c^2}\right) = 1,$$

we obtain

$$\Box f = \partial_{x'}^2 f - \frac{1}{c^2}\partial_{t'}^2 f = \Box' f,$$

thus, the d'Alembert operator is Lorentz invariant.

- We now consider a **scalar field** ψ, defined as a smooth function of (t, x) invariant under changes of inertial frames:

$$\psi'(x', t') = \psi(x, t)$$

- The one-dimensional wave equation for ψ is

$$\Box\psi = 0, \qquad \text{i.e.} \qquad \frac{\partial^2\psi}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2} = 0$$

Since $\Box = \Box'$ and $\psi'(x', t') = \psi(x, t)$, we obtain

$$\Box\psi(x, t) = 0 \quad \Longleftrightarrow \quad \Box'\psi'(x', t') = 0,$$

that is,

$$\frac{\partial^2\psi'}{\partial x'^2} - \frac{1}{c^2}\frac{\partial^2\psi'}{\partial t'^2} = 0$$

- Therefore the one-dimensional wave equation of a scalar field is invariant under Lorentz transformations. It is straightforward to extend this result to the full wave equation

- The Lorentz invariance of the wave equation is not merely a mathematical property: it is a fundamental physical requirement in the context of special relativity

- Maxwell's equations predict the existence of electromagnetic waves propagating at a fixed speed c, identified experimentally with the speed of light

- Since special relativity postulates that c is the same in all inertial frames, the equations describing electromagnetic waves must have the same form in every inertial frame

- In fact, if the wave equation were not Lorentz invariant, different inertial observers would obtain different equations for the propagation of light, and in particular would measure different propagation speeds. This would contradict both Maxwell's theory and the postulate of special relativity

- Thus, in the case of electromagnetism, the Lorentz invariance of the wave equation guarantees the consistency of Maxwell's equations with the principle of relativity and the invariance of the speed of light, i.e. the compatibility with the geometric structure of spacetime.

3.14 Alternative derivation of Lorentz boosts and connection with hyperbolic geometry

- There is an alternative expression of the Lorentz boost which will reveal the connection with hyperbolic geometry

- For technical reasons, this time is convenient to start with the *inverse boost*

$$\begin{pmatrix} ct \\ x \end{pmatrix} = \begin{pmatrix} E & F \\ G & H \end{pmatrix} \begin{pmatrix} ct' \\ x' \end{pmatrix} \iff \begin{cases} ct = Ect' + Fx' \\ x = Gct' + Hx' \end{cases}$$

then impose $ds^2 = ds'^2 \iff (ct)^2 - x^2 = (ct')^2 - x'^2$

- Replacing ct and x written as in the linear system above we obtain:

$$(E^2 - G^2)(ct')^2 + (F^2 - H^2)x'^2 + 2(EF - GH)ct'x' = (ct')^2 - x'^2 \implies \begin{cases} E^2 - G^2 = 1 \\ H^2 - F^2 = 1 \\ EF - GH = 0 \end{cases}$$

- By the identity $\cosh^2(u) - \sinh^2(u) = 1$, $\exists\, \zeta, \xi \in \mathbb{R}$ such that:

$$\begin{cases} E^2 - G^2 = 1 \\ H^2 - F^2 = 1 \\ EF - GH = 0 \end{cases} \iff \begin{cases} E = \cosh(\zeta),\ G = \sinh(\zeta) \\ H = \cosh(\xi),\ F = \sinh(\xi) \\ \cosh(\zeta)\sinh(\xi) - \cosh(\xi)\sinh(\zeta) = \sinh(\zeta - \xi) = 0 \end{cases}$$

- The third equation $\implies \xi = \zeta$, so we are left with the transformation

$$\begin{cases} ct = \cosh(\zeta)ct' + \sinh(\zeta)x' \\ x = \sinh(\zeta)ct' + \cosh(\zeta)x' \end{cases} \iff \begin{pmatrix} ct \\ x \end{pmatrix} = \begin{pmatrix} \cosh(\zeta) & \sinh(\zeta) \\ \sinh(\zeta) & \cosh(\zeta) \end{pmatrix} \begin{pmatrix} ct' \\ x' \end{pmatrix} \tag{3.7}$$

- To obtain the *direct boost*, we just have to invert the matrix:

$$\begin{pmatrix} ct' \\ x' \end{pmatrix} = \begin{pmatrix} \cosh(\zeta) & -\sinh(\zeta) \\ -\sinh(\zeta) & \cosh(\zeta) \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix} \iff \begin{cases} ct' = \cosh(\zeta)ct - \sinh(\zeta)x \\ x' = -\sinh(\zeta)ct + \cosh(\zeta)x \end{cases}$$

- Express now ζ in terms of v by considering $x' = 0$ in the system (3.7):

$$\begin{cases} \not{c}t = \cosh(\zeta)\not{c}t' \\ x = \sinh(\zeta)ct' \end{cases} \implies v = \frac{x}{t} = c\frac{\sinh(\zeta)}{\cosh(\zeta)} \iff \frac{v}{c} = \tanh(\zeta)$$

- ζ: **rapidity** of the boost

$$\tanh(\zeta) := \frac{v}{c} \iff \zeta := \tanh^{-1}\left(\frac{v}{c}\right) \tag{3.8}$$

- **Fraction of speed of light** of relative motion:

$$\beta := \frac{v}{c} = \tanh(\zeta) \in [0,1), \quad \beta \underset{v\to 0}{\to} 0,\ \beta \underset{v\to c}{\to} 1$$

- $\forall \zeta \in \mathbb{R}$, the identity

$$\cosh(\zeta) = \frac{1}{\sqrt{1-\tanh^2(\zeta)}} \quad \text{implies} \quad \cosh(\zeta) = \frac{1}{\sqrt{1-\beta^2}} = \gamma$$

- Finally, by (3.8), $\beta = \sinh(\zeta)/\cosh(\zeta) = \sinh(\zeta)/\gamma$, i.e.

$$\sinh(\zeta) = \beta\gamma$$

- We arrive to final form of the $x - \mathit{boost}$:

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cosh(\zeta) & -\sinh(\zeta) & 0 & 0 \\ -\sinh(\zeta) & \cosh(\zeta) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \equiv \mathcal{B}_x(\zeta) \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}$$

or

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \equiv \mathcal{B}_x(\beta) \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}$$

- $\mathcal{B}_x(\zeta)^{-1} = \mathcal{B}_x(-\zeta)$, $\mathcal{B}_x(\beta)^{-1} = \mathcal{B}_x(-\beta)$.
- To compare the general Lorentz boost matrix with the Galilean one we must consider the transformation $t \mapsto t'$ and not $ct \mapsto ct'$
- In this case:

$$\mathcal{B}(v) = \begin{pmatrix} \gamma & -\frac{v}{c^2}\gamma & 0 & 0 \\ -\gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

- It tends of course to the general Galilean matrix in the Galilean limit:

$$\mathcal{G}(v) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -v & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

- The $y-$boost is:

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cosh(\zeta) & 0 & -\sinh(\zeta) & 0 \\ 0 & 1 & 0 & 0 \\ -\sinh(\zeta) & 0 & \cosh(\zeta) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \equiv \mathcal{B}_y(\zeta) \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}$$

or

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & 0 & -\beta\gamma & 0 \\ 0 & 1 & 0 & 0 \\ -\beta\gamma & 0 & \gamma & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \equiv \mathcal{B}_y(\beta) \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}$$

- $\mathcal{B}_y(\zeta)^{-1} = \mathcal{B}_y(-\zeta)$, $\mathcal{B}_y(\beta)^{-1} = \mathcal{B}_y(-\beta)$

- The $z-$boost is:

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cosh(\zeta) & 0 & 0 & -\sinh(\zeta) \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\sinh(\zeta) & 0 & 0 & \cosh(\zeta) \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \equiv \mathcal{B}_z(\zeta) \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}$$

or

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & 0 & 0 & -\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\beta\gamma & 0 & 0 & \gamma \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \equiv \mathcal{B}_z(\beta) \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}$$

- $\mathcal{B}_z(\zeta)^{-1} = \mathcal{B}_z(-\zeta)$, $\mathcal{B}_z(\beta)^{-1} = \mathcal{B}_z(-\beta)$

- In addition to revealing a clearer link with hyperbolic geometry, the expressions obtained above are very useful when we compose collinear boosts

- Consider e.g. the standard configuration, then

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = B_x(\zeta_1) \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}, \qquad \begin{pmatrix} ct'' \\ x'' \\ y'' \\ z'' \end{pmatrix} = B_x(\zeta_2) \begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = B_x(\zeta_2) B_x(\zeta_1) \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix},$$

with

$$\begin{aligned} B_x(\zeta_2)B_x(\zeta_1) &= \begin{pmatrix} \cosh\zeta_2 & -\sinh\zeta_2 & 0 & 0 \\ -\sinh\zeta_2 & \cosh\zeta_2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cosh\zeta_1 & -\sinh\zeta_1 & 0 & 0 \\ -\sinh\zeta_1 & \cosh\zeta_1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \cosh\zeta_2\cosh\zeta_1 + \sinh\zeta_2\sinh\zeta_1 & -(\cosh\zeta_2\sinh\zeta_1 + \sinh\zeta_2\cosh\zeta_1) & 0 & 0 \\ -(\sinh\zeta_2\cosh\zeta_1 + \cosh\zeta_2\sinh\zeta_1) & \sinh\zeta_2\sinh\zeta_1 + \cosh\zeta_2\cosh\zeta_1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \cosh(\zeta_1+\zeta_2) & -\sinh(\zeta_1+\zeta_2) & 0 & 0 \\ -\sinh(\zeta_1+\zeta_2) & \cosh(\zeta_1+\zeta_2) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\ &= B_x(\zeta_1+\zeta_2), \end{aligned}$$

and similarly for all the other boosts

- Thus, for relative motion in the same direction, the rapidities simply add. The composition of two collinear boosts with rapidities ζ_1 and ζ_2 is still a boost with rapidity:

$$\zeta = \zeta_1 + \zeta_2.$$

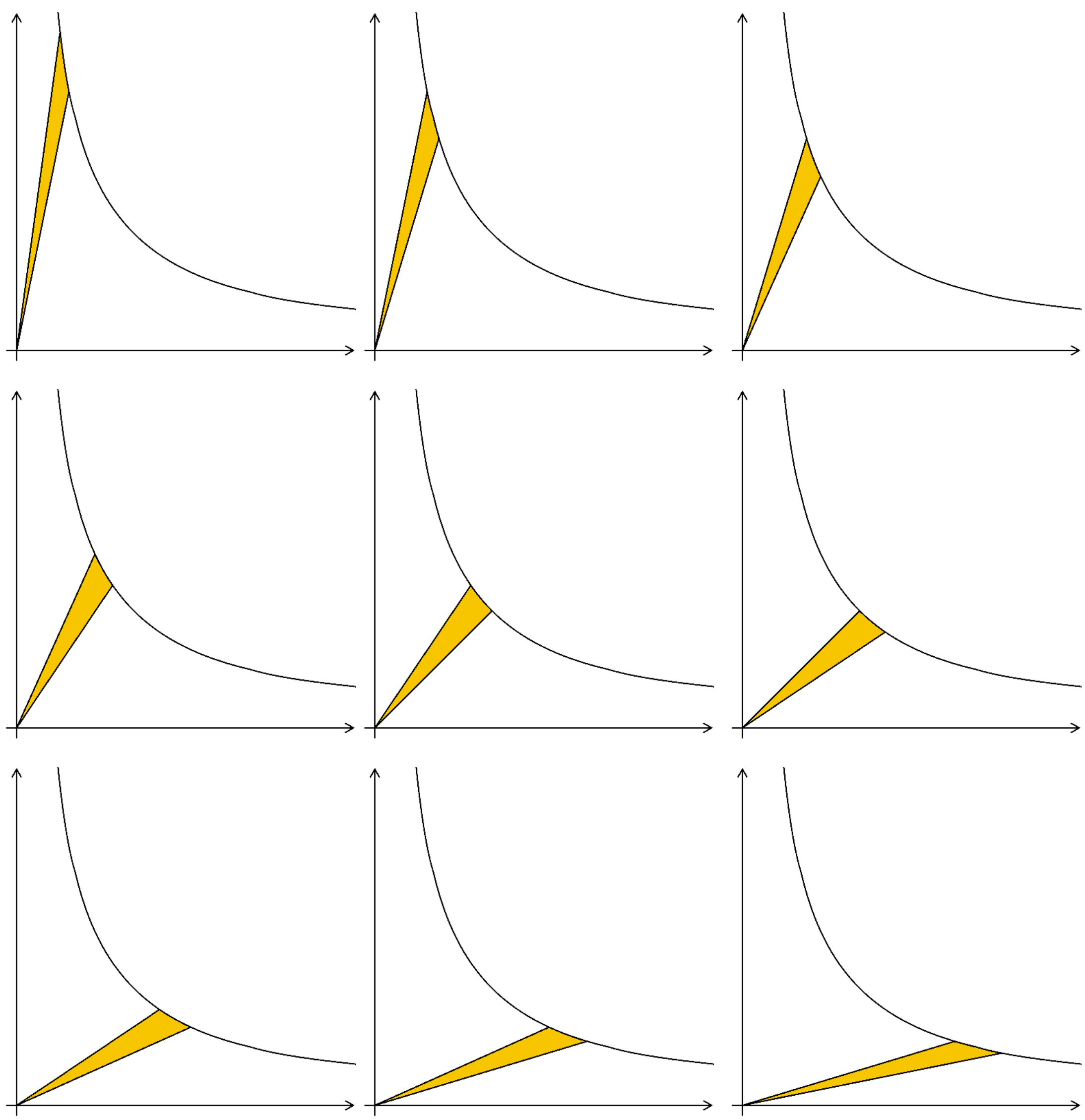
3.15 Boosts as 'hyperbolic rotations'

- The similarities between the boost and the rotation matrices explain why Lorentz boost are also called **hyperbolic rotations**

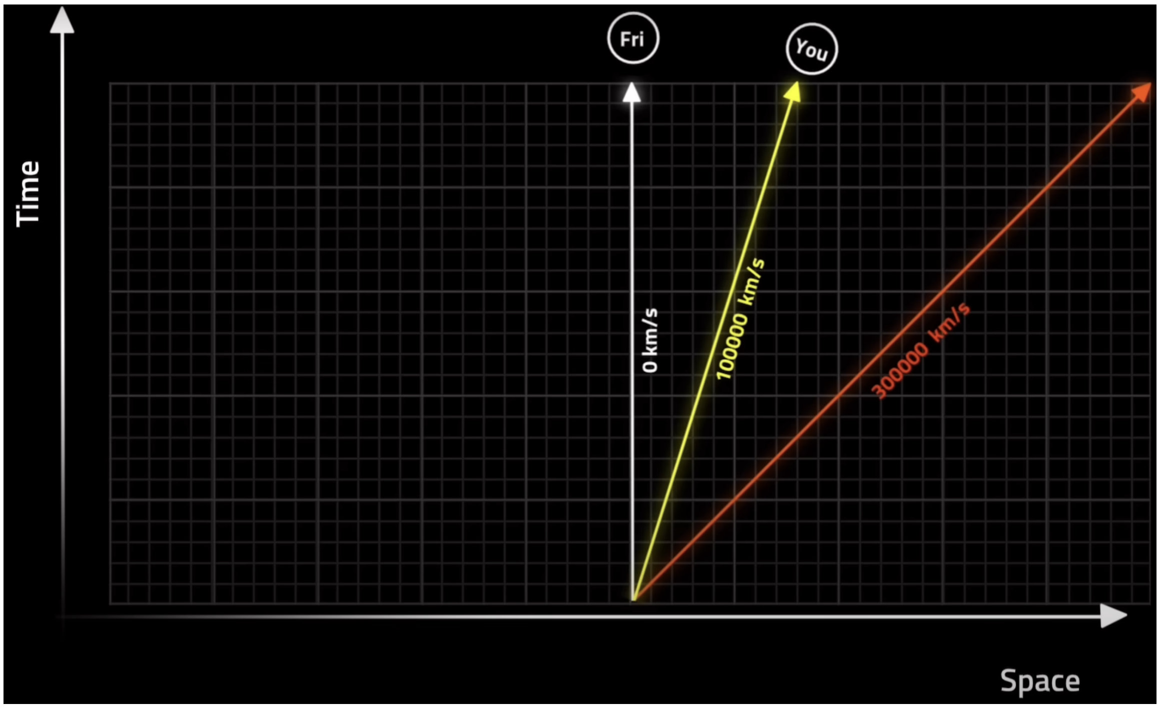
$$\begin{pmatrix} \cosh(\zeta) & -\sinh(\zeta) \\ -\sinh(\zeta) & \cosh(\zeta) \end{pmatrix} \quad \text{vs.} \quad \begin{pmatrix} \cos(\vartheta) & -\sin(\vartheta) \\ \sin(\vartheta) & \cos(\vartheta) \end{pmatrix}$$

- Lorentz boosts transform a vector in $\mathbb{R}^{1,3}$ in a vector on the same hyperboloid (due to the preservation of the Minkowski quadratic form!)

- Consequence: vectors in different frames are not rotated but **stretched** in order to remain on the same hyperboloid (figures from wikipedia)



- Why Lorentz transformations are hyperbolic rotations and not simply rotations?
- Because a rotation of axes would imply a modification of the speed of light! Klonusk (min. 13) https://www.youtube.com/watch?v=FGQseB-pvIg



3.16 Carroll transformations

- The Lorentz transformations of a spatial and temporal interval Δx, Δt are, as we know,

$$\Delta x' = \gamma(\Delta x - v\Delta t), \qquad \Delta t' = \gamma(\Delta t - \frac{v}{c^2}\Delta x)$$

- In the Galilean (or Newtonian) limit $c \to +\infty$, they reduce to

$$\Delta x' = \Delta x - v\Delta t, \qquad \Delta t' = \Delta t$$

  This limit corresponds to $v \ll c$ and $\Delta x \ll c\,\Delta t$: low velocities, *large time-like intervals*

- Now consider another limit, that of low velocities and *large space-like intervals*, i.e. $v \ll c$ and $\Delta x \gg c\,\Delta t$, so that $\Delta s^2 < 0$. The Lorentz transformations in this case become

$$\begin{cases} \Delta x' = \Delta x \\ \Delta t' = \Delta t - \frac{v}{c^2}\Delta x \end{cases} \qquad \Longleftrightarrow \qquad \begin{cases} x' = x \\ t' = t - \xi x, \end{cases} \tag{3.9}$$

  where the parameter $\xi = v/c^2 \geqslant 0$ has dimensions of the inverse of a velocity

- Eqs (3.9) are called **Carroll transformations** and represent a low-velocity, but non-Galilean, limit of the Lorentz transformations due to the fact that Δx is so large that $\Delta x/\Delta t \gg c$, which is impossible in the Galilean limit $c \to +\infty$

- Actually, the Carroll limit is *formally* equivalent to $v \ll c$ and $c \to 0$!

- Their *formal* similarity with the Galilean transformations is evident, but with the roles of space and time coordinates exchanged!

- In the Carroll limit, the time difference in $\mathscr{R}'$ depends on the competition between Δt and $\xi\Delta x$:

  1. if $\Delta t > \xi\Delta x$, then $\Delta t' > 0$
  2. if $\Delta t = \xi\Delta x$, then $\Delta t' = 0$
  3. if $\Delta t < \xi\Delta x$, then $\Delta t' < 0$.

  In option 3, the temporal order of events in $\mathscr{R}$ is reversed in $\mathscr{R}'$!

- This is the reason why J.-M. Lévy-Leblond, who studied these transformations in 1965, payed tribute to **Lewis Carroll**, the author of *Alice in Wonderland*: a universe having the Carroll group of transformations as its invariance group would be dominated, like Alice's world, by the relativity of time and by the absence of causality.

Geometric interpretation of the limits $c \to +\infty$ and $c \to 0$

- In the spacetime plane (ct, x), the light cone is defined by $x^2 - c^2t^2 = 0$, i.e. $x = \pm ct$ and these two straight lines separate:
  - the *time-like* region $|x| < ct$
  - the *space-like* region $|x| > ct$

  The parameter c determines the opening of the cone

- **Relativistic case (c finite)**: the lines $x = \pm ct$ have slope ± 1 in the (ct, x) plane. The cone has a finite opening angle.

- **Galilean limit ($c \to +\infty$)**: the equation $x = \pm ct$ can be satisfied by any fixed x only if $t \to 0$, hence in the (ct, x) plane:
  - the light cone collapses toward the horizontal axis $ct = 0$
  - the cone aperture becomes infinitely wide.

  Geometrically, all directions become time-like, *time becomes absolute* and we are allowed to talk about the absolute past and future

- **Carroll limit ($c \to 0$)**: the equation $x = \pm ct$ can be satisfied for any fixed t only if $x \to 0$, hence in the (ct, x) plane:
  - the light cone collapses toward the vertical axis $x = 0$
  - the cone aperture becomes infinitely narrow.

  Geometrically, all directions become space-like and *space becomes absolute*

- To summarize, the parameter c controls the geometry of the light cone:

Limit	Shape of the cone	Absolute quantity
c finite	finite opening	none
$c \to +\infty$	flattened (horizontal)	time
$c \to 0$	collapsed (vertical)	space

3.17 Relativity of simultaneity

- A direct consequence of Lorentz transformations is the relativity of simultaneity: events occurring simultaneously at different places in a frame $\mathscr{R}$ aren't simultaneous anymore in a frame $\mathscr{R}'$ with relative speed $v \neq 0$

- This was one of the most shocking and hard to accept aspects during the first years after the development of special relativity

- A and B: *simultaneous events* in $\mathscr{R}$

$$e_A = (ct, x_A, y_A, z_A), \quad e_B = (ct, x_B, y_B, z_B), \qquad \underline{t_A = t_B \equiv t}$$

- Consider for simplicity the standard configuration, then in $\mathscr{R}'$ we have:

$$t'_A = \gamma \left(t - \frac{v}{c^2} x_A \right), \quad t'_B = \gamma \left(t - \frac{v}{c^2} x_B \right)$$

- So, in $\mathscr{R}'$, e_A, e_B are separated by a time interval

$$\Delta t' = t'_B - t'_A = \gamma \frac{v}{c^2} (x_A - x_B) \neq 0 \iff x_A \neq x_B,$$

  therefore, they are *not simultaneous anymore* in $\mathscr{R}$!

- Of course, neither our visual perception nor normal clocks are capable of measuring the loss of simultaneity due to the presence of c^2 at the denominator

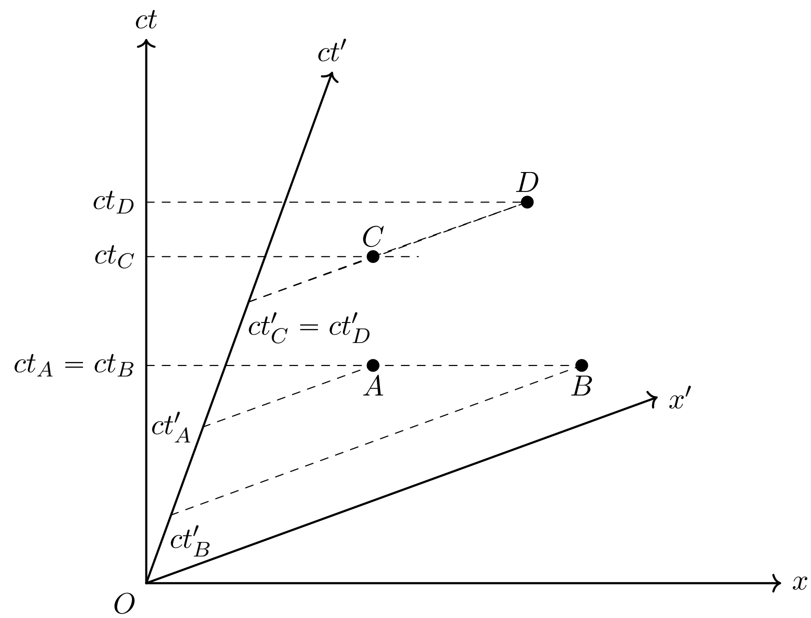
- However, if $x_A - x_B$ is big enough, the effect can be measured for 'small' speed v with very precise (atomic) clocks and it's perfectly in line with the relativistic prediction

- If two events *coincide* in $\mathscr{R}$, i.e. $t_A = t_B$ and $x_A = x_B$, then they coincide in every other reference frame. In fact, in that case $\Delta t' = 0 \iff t'_B = t'_A$, and

$$\Delta x' = \gamma(\Delta x - v \Delta t) = 0 \implies x'_A = x'_B$$

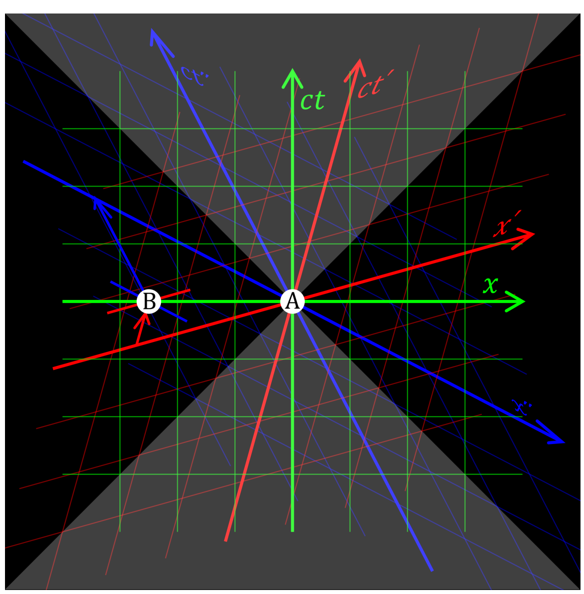
  and similarly for the other coordinates if $y_A = y_B$ and $z_A = z_B$.

3.18 Relativity of simultaneity with Minkowski diagrams



- Events A and B are simultaneous in $\mathscr{R}$, but not in $\mathscr{R}'$
- Events C and D are simultaneous in $\mathscr{R}'$, but not in $\mathscr{R}$

https://upload.wikimedia.org/wikipedia/commons/7/78/Relativity_of_Simultaneity_Animation.gif



3.19 Preservation of causality (chronological order of events)

- Suppose now that a particle or a photon *first* passes in (x_A, y, z) at time t_A and *then* arrives in (x_B, y, z) at time t_B moving along the x-axis with constant velocity u, then

$$e_A = (ct_A, x_A, y, z), \quad e_B = (ct_B, x_B, y, z), \qquad \Delta t = t_B - t_A > 0$$

and

$$x_B - x_A = u(t_B - t_A) \tag{3.10}$$

- In $\mathscr{R}'$, the Lorentz transformations imply

$$\Delta t' = t'_B - t'_A = \gamma\left[(t_B - t_A) - \frac{v}{c^2}(x_B - x_A)\right] \underset{(3.10)}{=} \gamma\left[1 - \frac{vu}{c^2}\right]\Delta t$$

- If $v < c$ and $u \leqslant c$, then $[1 - vu/c^2] > 0$, so $\Delta t'$ and Δt have the same sign: causality (t_A, t'_A : first, t_B, t'_B : then) is preserved in every inertial reference frame

- If $v < c$ and $u > c$, then $c/u < 1$, so

$$\left[1 - \frac{vu}{c^2}\right] < 0 \iff 1 < \frac{vu}{c^2} \iff \frac{c^2}{u} = \left(\frac{c}{u}\right) c < v < c$$

there would exist infinite reference frames that move at subluminal velocity v in which *the effect is seen before the cause*!

- So, admitting the existence of particles traveling at superluminal speed, means admitting that causality can be reversed!

- Special relativity *doesn't forbid* the existence of superluminal particles, i.e. that travel at speed $> c$, called **tachyons**, it just predicts that for such particles causality is not respected

- Tachyons have been studied (theoretically) a lot and it has been proven that they should be emitted with superluminal speed and bound to travel faster than light

- More than a maximal speed, c must be seen as an (upper or lower) **limit speed**! This limit differentiates between a causal and a non-causal universe.

3.20 Measure of time in different inertial frames: proper and coordinate time

- Suppose:
  - from the perspective of $\mathscr{R}$: a clock is moving uniformly with velocity $\mathbf{v}$
  - from the perspective of $\mathscr{R}'$: the clock at rest ($\mathscr{R}'$ and clock: *co-moving*)
- τ, called **proper time**[5], it's the time measured by an inertial observer in $\mathscr{R}'$
- t, called **coordinate time**, it's the time that an inertial observer in $\mathscr{R}$ assigns to events in $\mathscr{R}'$, it's defined like this because it's part of the coordinate system that the observer must choose to describe events
- We are going to prove that t and τ agree only when $\mathbf{v} = \mathbf{0}$.

3.21 Dilation of time intervals (from the viewpoint of an observer who sees a clock moving)

- Often, the dilation of time intervals is improperly reduced to the misleading, but very useful, slogan 'time dilation' (that we'll also use... because of its conciseness)
- Time doesn't slow or accelerate, time always flows at 1 second per second in all circumstances, what changes is the value that two observers in relative motion attribute to the measure of time intervals
- We derive the dilation of time intervals in two ways: using the invariance of the spacetime intervals and the Lorentz transformations

Time dilation from the invariance of the spacetime interval

- An observer in $\mathscr{R}$ measures, during a time dt, the clock in $\mathscr{R}'$ traveling the distance: $|d\mathbf{x}| = \sqrt{dx^2 + dy^2 + dz^2}$
- An observer in $\mathscr{R}'$ moves with the clock, measures the proper time $d\tau$ and a null displacement: $|d\mathbf{x}| = 0$
- By the invariance of spacetime interval:

$$\underbrace{c^2 d\tau^2 - 0}_{ds'^2} = \underbrace{c^2 dt^2 - |d\mathbf{x}|^2}_{ds^2}$$

[5] the distinction between proper and coordinate time was introduced by Minkowski in 1908.

dividing by c^2

$$d\tau^2 = \frac{1}{c^2}ds^2 = dt^2 - \frac{1}{c^2}|d\mathbf{x}|^2 < dt^2$$

which shows that *proper time is smaller than coordinate time*, more explicitly, taking the square root

$$d\tau = dt\sqrt{1 - \frac{|d\mathbf{x}|^2}{c^2(dt)^2}} = dt\sqrt{1 - \frac{v^2}{c^2}} = \frac{dt}{\gamma(v)}$$

- For finite time intervals we have

$$\Delta\tau = \frac{\Delta t}{\gamma(v)} \qquad \text{or} \qquad \Delta t = \gamma(v)\Delta\tau$$

so the measure of time intervals by clocks in $\mathscr{R}$ is dilated by the factor $\gamma(v)$ w.r.t. that of clocks in $\mathscr{R}'$ and the dilation increases with the speed v

- Of course time dilation is a symmetric effect for the two reference frames

- In the following, we will often use these formulae

$$\boxed{d\tau^2 = \frac{ds^2}{c^2}}, \qquad \boxed{d\tau = \frac{dt}{\gamma(v)}}. \tag{3.11}$$

Time dilation from the Lorentz transformations

- Consider again a clock is at rest in $\mathscr{R}'$ and suppose that it measures the proper time interval $\Delta\tau$

- The time interval measured by a clock in $\mathscr{R}$ can be obtained using the inverse Lorentz transformation:

$$\Delta t = \gamma\left(\Delta\tau + \frac{v}{c^2}\Delta x'\right) \underset{\Delta x'=0}{\Longrightarrow} \Delta t = \gamma\Delta\tau$$

confirming the previous result

- To summarize, we are perfectly welcome to introduce arbitrary coordinates, but that mathematical comfort comes at a price: accepting *time dilation from the point of view of the observer who sees the motion.*

- Limit cases:

  - Galilean limit: $\gamma \to 1 \implies \Delta t \simeq \Delta\tau$, this is why at our common speed of motion, we don't experience time dilation

  - $v \to c \implies \Delta\tau \to 0$: it seems that from a photon's perspective time doesn't pass and every future events happens simultaneously, but v never reaches c, so this interpretation is misleading!

- If v is a non-negligible fraction of the speed of light, $\gamma \gg 1$ and so the time Δt measured in $\mathscr{R}$ is considerably larger than the proper time $\Delta\tau$ in $\mathscr{R}'$

- Example 1: Speed $= 0.8c$, $\Delta t = 10$ years

$$\Delta\tau = \int_0^{10} \sqrt{1 - \frac{v^2}{c^2}}\, dt = 10 \cdot \sqrt{1 - \left(\frac{0.8c}{c}\right)^2} \approx 6 \text{ years}$$

- Example 2: Speed $= 0.99c$, $\Delta t = 7$ years

$$\Delta\tau = 10 \cdot \sqrt{1 - \left(\frac{0.99c}{c}\right)^2} \approx 1 \text{ years.}$$

- So why Newton (who was kinda smart...) thought he could group anyone's proper time in one absolute time? For speed as 100km/h $\approx$ 28m/s,

$$\gamma \approx 1 + 4.28 \times 10^{-15},$$

and this correction to the unit value is negligible for practically all intents and purposes.

3.22 Proper time along general time-like worldlines

- The previous formula for $d\tau$ was derived assuming a straight, inertial motion with constant velocity $\mathbf{v}$

- However, the concept of proper time remains valid also for an observer moving along a general differentiable time-like curve $t \mapsto \mathbf{x}(t)$ in spacetime

- In fact, *locally*, such curved worldlines can be linearized

- So also for curved (accelerated) non-inertial paths, we can compute the proper time increment *at each point* $\mathbf{x}(t)$ using:

$$d\tau(t) = dt\sqrt{1 - \frac{v(t)^2}{c^2}}$$

- An important counter-intuitive fact of special relativity is that straight (inertial) worldlines maximize the proper time. To prove it, we need two preliminary results.

3.23 Existence of a frame in which two time-like separated events occur at the same spatial location

- Let $e_1, e_2 \in \mathcal{M}$ be two events and let $x = e_2 - e_1 = (ct, \mathbf{x}) \in \mathbb{R}^{1,3}$ be a future-directed time-like vector, i.e. $q(x) = (ct)^2 - |\mathbf{x}|^2 > 0$, $t > 0$.

- We want to determine an inertial frame in which the two events associated with the displacement vector x occur at the same spatial point, i.e. such that the spatial component of the transformed vector vanishes

- We consider a Lorentz boost along the direction of $\mathbf{x}$, with velocity $\mathbf{v} = \lambda\mathbf{x}$, with $\lambda > 0$ (with physical dimension T^{-1}) to be determined in such a way that $\mathbf{x}' = \mathbf{0}$

- From the general boost formulas (section 3.10) we have

$$ct' = \gamma\left(ct - \frac{\mathbf{v}\cdot\mathbf{x}}{c}\right)$$

$$\begin{aligned}\mathbf{x}' &= \mathbf{x} + \frac{\gamma-1}{|\mathbf{v}|^2}(\mathbf{v}\cdot\mathbf{x})\mathbf{v} - \gamma t\,\mathbf{v} = \mathbf{x} + \frac{\gamma-1}{\lambda^2|\mathbf{x}|^2}\lambda^2|\mathbf{x}|^2\,\mathbf{x} - \gamma t\,\lambda\mathbf{x} \\ &= \mathbf{x} + (\gamma-1)\mathbf{x} - \gamma t\,\lambda\mathbf{x} = \gamma(\mathbf{x} - t\,\lambda\mathbf{x})\end{aligned}$$

so

$$\mathbf{x}' = \mathbf{0} \iff \lambda = 1/t \implies \mathbf{v} = \frac{\mathbf{x}}{t}$$

- The condition $q(x) > 0$ implies $|\mathbf{x}| < ct$, hence $|\mathbf{v}| < c$ and the boost is admissible

- With this choice of $\mathbf{v}$, the time component becomes

$$ct' = \gamma\left(ct - \frac{|\mathbf{x}|^2}{ct}\right) = \gamma\,\frac{(ct)^2 - |\mathbf{x}|^2}{ct}$$

and γ becomes

$$\gamma = \frac{1}{\sqrt{1 - \frac{|\mathbf{x}|^2}{c^2t^2}}} = \frac{ct}{\sqrt{(ct)^2 - |\mathbf{x}|^2}}$$

so

$$ct' = \sqrt{(ct)^2 - |\mathbf{x}|^2} = \sqrt{q(x)} = \|x\| \qquad (\|x\| \text{ has the physical dimension } L)$$

- In conclusion, the boost $\mathcal{B}(\mathbf{v})$, with $\mathbf{v} = \frac{\mathbf{x}}{t}$, realizes a change of reference frame in which the displacement vector has spacetime coordinates

$$x' = (\|x\|, \mathbf{0}),$$

i.e. in this frame the events e_1, e_2 occur at the same spatial location, but at different instants of time separated by a proper time

$$\tau = \frac{\|x\|}{c}.$$

3.24 The reverse triangle inequality for future-directed time-like vectors

- **Theorem**: if $x, y \in \mathbb{R}^{1,3}$ are future-directed time-like vectors, then

$$\|x + y\| \geqslant \|x\| + \|y\| \qquad (\textbf{reverse triangle inequality})$$

- **Proof**: since x is future-directed time-like, by the previous result there exists an inertial frame in which x has spacetime coordinates $(\|x\|, \mathbf{0})$

- Let y be another future-directed time-like vector with coordinates $(y^0, \mathbf{y})$ in this frame

- Since y is future-directed time-like, we have $y^0 > 0$ and $q(y) = (y^0)^2 - |\mathbf{y}|^2 > 0$, so

$$\|y\| = \sqrt{(y^0)^2 - |\mathbf{y}|^2} \leqslant y^0$$

  which implies

$$x \circ y = \|x\| \, y^0 - \mathbf{0} \cdot \mathbf{y} = \|x\| \, y^0 \geqslant \|x\| \, \|y\|,$$

- By definition, both $\circ$ and $\| \ \|$ are Lorentz invariant, so this inequality holds *in every inertial frame*:

$$x \circ y \geqslant \|x\| \, \|y\| \qquad (\textbf{reverse Cauchy-Schwarz inequality})$$

- Hence:

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2x \circ y \geqslant \|x\|^2 + \|y\|^2 + 2\|x\| \, \|y\| = (\|x\| + \|y\|)^2$$

- Both x and y belong to the future lightcone, which is a convex cone, so it's stable w.r.t. convex combinations and multiplications by positive scalars, so $2(x/2 + y/2) = x + y$ still belongs to the future lightcone, i.e. $x + y$ is a future-directed time-like vector

- Thanks to this remark, $\|x + y\| > 0$, so taking the square roots of the previous inequality gives

$$\|x + y\| \geqslant \|x\| + \|y\|.$$

□

3.25 Inertial motion maximizes proper time

- **Theorem**: among all piecewise inertial future-directed time-like motions joining the origin O to an event e, the inertial motion maximizes the proper time.

- **Proof**: consider a piecewise inertial worldline, i.e. a finite union of inertial future-directed time-like segments joining O to e:

$$O = P_0, P_1, \ldots, P_N = e$$

and set

$$x := e - O, \quad x_i := P_i - P_{i-1}, \qquad i = 1, \ldots, N,$$

then, we can write x as the telescopic (affine) sum

$$x = P_N - P_0 = (P_N - P_{N-1}) + (P_{N-1} - P_{N-2}) + \cdots + (P_1 - P_0) = \sum_{i=1}^{N} (P_i - P_{i-1}) = \sum_{i=1}^{N} x_i$$

- The proper time along the i-th inertial segment is

$$\tau_i = \frac{\|x_i\|}{c},$$

hence the total proper time along the piecewise worldline is

$$\tau = \sum_{i=1}^{N} \tau_i = \frac{1}{c} \sum_{i=1}^{N} \|x_i\|$$

- Applying repeatedly the reverse triangle inequality, we obtain

$$\frac{\|x\|}{c} = \frac{1}{c} \left\| \sum_{i=1}^{N} x_i \right\| \geqslant \frac{1}{c} \sum_{i=1}^{N} \|x_i\| = \tau.$$

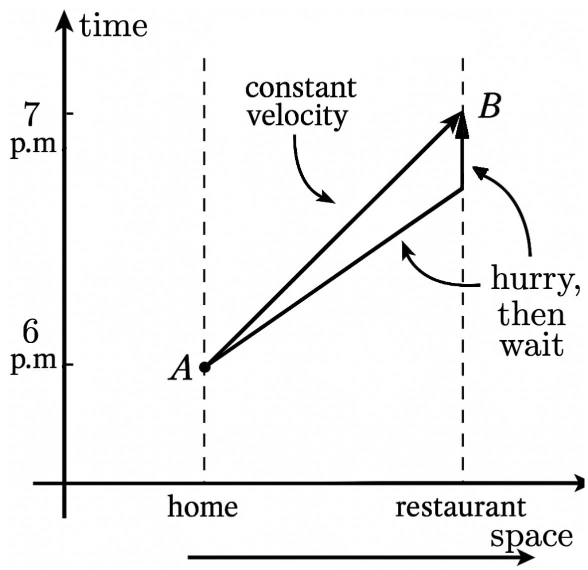
- But for the inertial (straight) motion from O to x we have

$$\tau_{\text{in}} = \frac{\|x\|}{c} \implies \tau_{\text{in}} \geqslant \tau,$$

so the inertial (straight) worldline joining O to e maximizes the proper time $\square$

- The same result holds for any smooth future-directed time-like curve γ: intuitively, one approximates γ by the piecewise inertial tangential worldlines and passes to the limit, so that no non-inertial curve can yield a larger proper time than the inertial one.

- Example (Sean Carroll): imagine you want to meet a friend at your favorite restaurant at 7 p.m. leaving home at 6 p.m.
  - you travel the straight line Home → Restaurant with the the correct constant speed v that allows you to arrive there precisely at 7 p.m.
  - your friend travels with a faster constant speed, arrives sooner and then waits 7 p.m.

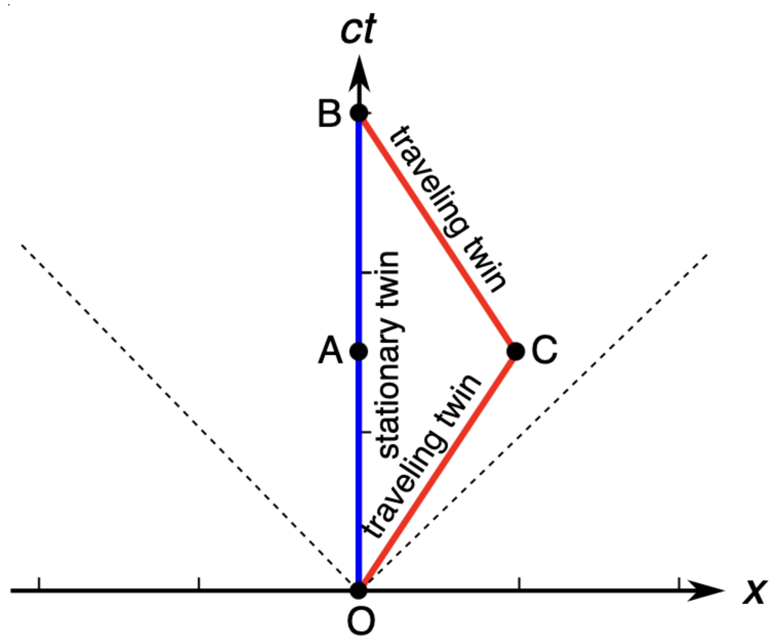


- This spacetime diagram uses the *spherical cow approximation*: assuming (incorrectly!) that your friend decelerates instantaneously so that we can ignore the fact that when the motion stops there is a deceleration which breaks inertial motion
- Your friend's clock will show that less time has passed because the clock traveled a longer path.

3.26 The twin 'paradox'

- The *twin 'paradox'* is a paradox in the sense of the Greek word $\pi\alpha\rho\alpha\delta o\xi o\zeta$: 'beyond common sense', but when we think relativistically, the paradox disappears

- Imagine two twins separated from birth: one remains on Earth, while the other travels on a spaceship at 86.6% of c (so that $\gamma = 2$) towards the Andromeda galaxy. At some point, the traveling twin returns to Earth and finds that the Earth-bound twin is older...



- The *apparent paradox* arises from the fact that, from each twin's perspective, the other is moving, so each should see the other's clock running slower!

- Even neglecting that the 'symmetry' is broken because in order to reunite and compare their ages, the traveling twin must either *reverse* course or move in a *non-inertial* (e.g., circular) trajectory, the twin, to come back, has traveled a longer path.

3.27 Time dilation of the mean lifetime of cosmic muons

- In 1936, Carl Anderson and Seth Neddermeyer (Caltech) discovered muons, the 'evil cousins' of electrons: unstable elementary particle with the same charge as the electron, but around 200 times heavier (famously Isaac Rabi asked: *who ordered that?!?*)

- The Earth is continuously bombarded by showers of extremely energetic particles, mainly protons, improperly called *cosmic rays* for historical reasons. When they hit the atmosphere, they produce other particles, among them is the muon (μ^-), which decays in a very short time into an electron (e^-), a muonic neutrino (ν_μ) and an electron antineutrino ($\bar{\nu}_e$):
$$\mu^- \to e^- + \nu_\mu + \bar{\nu}_e$$

- The decay law is exponential. If at a certain instant $t = 0$ there are n_0 muons, at a later time t the number n of muons is
$$n = n_0 e^{-t/\tau},$$
where τ is the **mean lifetime** of the muons: after a time interval τ the number of muons is reduced by a factor $e \simeq 2.718$, that is, slightly more than one third of the initial muons survive

- The mean lifetime of muons, measured in their rest frame, is $2.2\,\mu$s. If there were no time dilation effect, the mean lifetime at rest (denoted by τ_0) would be equal to that measured in a reference frame with respect to which they are moving (the so-called lifetime in flight, that we have denoted by τ)

- Muons from cosmic radiation moves at ultrarelativistic speed, typically $v = 0.995c$. In a time interval equal to their mean lifetime at rest, they travel a distance $d = v\tau_0$, that is 660 m. After $3\tau_0$ the distance traveled is three times larger, about 2000 m

- Hence, we'd expect to observe, at sea level, a fraction e^{-3}, that is about 5%, of the muons observed at an altitude of 2000 m

- The first experiment of this kind was performed by Bruno Rossi and David Hall in 1941 measuring muon flux between Echo Lake in Colorado and Denver (USA)

- A more precise experiment was performed in 1962 by D.H. Frisch and J.H. Smith at MIT using a scintillation detector, the muons of cosmic radiation were counted at the summit of Mount Washington, in New Hampshire (USA), at an altitude of 1910 m

- Frisch and Smith had the brilliant idea to film the experiment and distribute it all around the world. It became one of the most famous physics demonstration film ever made: https://www.youtube.com/watch?v=Fcyn9UEeWus

- The number of muons observed is on average about 560 per hour. Under the assumption that the muon flux is constant, repeating the count at 0 m altitude, one would expect to detect only about thirty particles. Instead, more than 400 are observed

- This result is perfectly explained by time dilation: in the reference frame at rest with the Earth, in which the muons have velocity $v = 0.995c$, their mean lifetime is not $\tau_0 = 2.2\,\mu$s, but

$$\tau = \gamma\tau_0,$$

  and, since $\gamma \simeq 10$, it is $22\,\mu$s. In a time interval equal to τ, the muons travel more than 6000 m, and therefore it is natural that a large fraction of them survive after a height difference of 1900 m

- To cover a distance $d = 1900$ m at velocity $v = 0.995c$, the muons take a time

$$t = \frac{d}{v} = 6.4\,\mu\text{s}$$

- As said before, taking relativistic effects into account, their mean lifetime is $\tau = 22\,\mu$s and after $t = 6.4\,\mu$s, according to the decay law, the number of surviving muons is

$$n = n_0 e^{-t/\tau} = n_0 e^{-6.4/22} = 0.75\, n_0$$

- If 560 muons are counted at the summit of Mount Washington, after $6.4\,\mu$s and a height difference of 1900 m, that is at ground level, we should observe

$$0.75 \cdot 560 = 420$$

  muons, a value in excellent agreement with the experimentally determined one

- One may wonder what happens from the rest frame of the muon, where its internal clock measures a mean lifetime $\tau_0 = 2.2\,\mu s$ . . . to understand that we need to discuss length contraction. . .

3.28 Length contraction (from the viewpoint of an observer who sees an object moving)

- Also called *Lorentz-Fitzgerald contraction* because it was proposed by Fitzgerald and Lorentz before the development of Einstein's theory of special relativity in order to stubbornly save the idea of luminiferous aether against the negative result of the Michelson Morley experiment

- It is crucial to start by **defining the length** of an object in a given frame and to translate it in mathematical terms:

  length := spatial distance between two simultaneous events

- The **rest** (or proper) **length** is the length of an object measured in the co-moving inertial frame, i.e. where the object is seen at rest, hence the name

- As always, let's consider the standard configuration: two inertial frames $\mathscr{R}$ and $\mathscr{R}'$ moving along the x-coordinates with relative velocity v

- If a rod lying on the x-axis is at rest in $\mathscr{R}'$, then its extreme points define two simultaneous events $e_1' = (ct', x_1')$ and $e_2' = (ct', x_2')$, with *the same* t'

- The rest length of the rod in $\mathscr{R}'$ is then: $L_0 := |x_2' - x_1'|$

- The key to understand length contraction consists in noticing that the events $e_1' = (ct', x_1')$ and $e_2' = (ct', x_2')$, simultaneous in $\mathscr{R}'$, are *not simultaneous anymore in* $\mathscr{R}$. In fact, using the inverse Lorentz transformations we obtain

$$t_1 = \gamma\left(t' + \frac{vx_1'}{c^2}\right), \qquad t_2 = \gamma\left(t' + \frac{vx_2'}{c^2}\right),$$

  hence

$$t_2 - t_1 = \gamma\,\frac{v}{c^2}(x_2' - x_1') = \gamma\,\frac{v}{c^2}L_0 \neq 0$$

- To compute the length L in $\mathscr{R}$, one must impose that the events e_1 and e_2 are *simultaneous in* $\mathscr{R}$, i.e. $e_1 = (ct, x_1)$ and $e_2 = (ct, x_2)$ (notice *the same* t), with spatial coordinates x_1 and x_2 related to x_1' and x_2' by the Lorentz transformations

$$x_1' = \gamma(x_1 - vt), \qquad x_2' = \gamma(x_2 - vt),$$

- The relation between the length $L := |x_2 - x_1|$ measured in $\mathscr{R}$ and the rest length L_0 is then

$$L_0 = |x_2' - x_1'| = \gamma|x_2 - x_1| = \gamma L$$

- Therefore

$$\boxed{L = \frac{L_0}{\gamma}}$$

  i.e. the length of the rod measured in $\mathscr{R}$ is contracted w.r.t. the rest length

- One should firmly resist the temptation to think that the molecular structure of rod is modified! The contraction of the length of an object is not an intrinsic property, but a relation between measurements performed by observers in relative motion!

- Note that the transverse dimensions of a body, that is, those perpendicular to the direction of motion, do not change

- As time dilation, also length contraction is *intended from the point of view of the observer who sees the motion* and it is a symmetric effect for the two frames

- Turning back to **the journey of a muon**: each muon sees the surface of the Earth rushing upward toward it at a speed $v \approx 0.995c$, so, from its rest reference frame the atmosphere is contracted by the Lorentz factor $\gamma(v)$ and this is what allows it to reach the surface before decaying!

- Since simultaneity is crucial to define the length of an object, it is interesting to analyze the possibility to determine the a frame in which two events occur simultaneously...

3.29 Existence of a frame in which two space-like separated events occur simultaneously

- Let $x = (ct, \mathbf{x}) \in \mathbb{R}^{1,3}$ be a space-like vector, i.e. $q(x) = (ct)^2 - |\mathbf{x}|^2 < 0$.

- We want to determine an inertial frame in which the two events $e_1, e_2 \in \mathcal{M}$ associated with the displacement vector $x = e_2 - e_1$ occur simultaneously, i.e. such that the temporal component of the transformed vector vanishes

- We consider a Lorentz boost along the direction of $\mathbf{x}$, with velocity $\mathbf{v} = \lambda\,\mathbf{x}$, with $\lambda > 0$ (with physical dimension T^{-1}) to be determined in such a way that $t' = 0$

- From the general boost formulas (section 3.10) we have

$$ct' = \gamma\left(ct - \frac{\mathbf{v}\cdot\mathbf{x}}{c}\right), \qquad \mathbf{x}' = \mathbf{x} + \frac{\gamma - 1}{|\mathbf{v}|^2}(\mathbf{v}\cdot\mathbf{x})\mathbf{v} - \gamma t\,\mathbf{v}$$

- Since $\mathbf{v} = \lambda\,\mathbf{x}$, we have

$$\mathbf{v}\cdot\mathbf{x} = \lambda|\mathbf{x}|^2,$$

hence

$$ct' = \gamma\left(ct - \frac{\lambda|\mathbf{x}|^2}{c}\right) = 0 \iff ct = \frac{\lambda|\mathbf{x}|^2}{c} \iff \lambda = \frac{c^2 t}{|\mathbf{x}|^2}$$

- Therefore

$$\mathbf{v} = \frac{c^2 t}{|\mathbf{x}|^2}\,\mathbf{x}.$$

- Since $q(x) < 0$, we have $|\mathbf{x}|^2 > c^2 t^2$, hence

$$|\mathbf{v}|^2 = \frac{c^4 t^2}{|\mathbf{x}|^4}|\mathbf{x}|^2 = \frac{c^4 t^2}{|\mathbf{x}|^2} < c^2 \iff |\mathbf{v}| < c,$$

so the boost is admissible

- In the corresponding inertial frame one has $ct' = 0$, hence the separation between the two events is purely spatial:

$$x' = (0, \mathbf{x}')$$

- Using the invariance of the spacetime interval we have

$$q(x) = q(x') \underset{(ct'=0)}{=} -|\mathbf{x}'|^2,$$

so

$$|\mathbf{x}'| = \sqrt{-q(x)}$$

- In conclusion, $\mathcal{B}(\mathbf{v})$ boosts to an inertial frame in which the events occur simultaneously, but at different spatial positions separated by the distance $\sqrt{-q(x)}$.

3.30 Relativity and philosophy: the reality of the future

- The theory of relativity, as is easy to imagine given its results, gave rise to an intense debate not only in the scientific field (at least during the first two decades after its formulation, that is, until the progressively accumulated experimental verifications dispelled any doubt about its validity), but also in the philosophical domain

- Even those most reluctant to attribute philosophical significance to the consequences of science had to acknowledge that the discussion of certain important problems traditionally belonging to philosophy (the concepts of space, time, and the universe) can't ignore the worldview provided by relativity. Authors such as Reichenbach, Grünbaum, and Putnam deserve credit for having assigned to relativity the role of an indispensable reference point in philosophical reflections on space and time

- As an example of how the theory of relativity can be used to clarify (or even, according to some, resolve) certain philosophical problems, we summarize here an argument by **Hilary Putnam**[6] (1926–2016) concerning the reality of the future. The question is the following: *can future events be considered real events?* Putnam argues that relativity provides a precise and positive answer to this question

- Putnam argument is founded on considering reality as a *relational* concept: in that view, when we claim that an event is real, we mean that it is real w.r.t. another event

- Let us introduce a binary relation δ between ordered pairs of events, called the *relation of reality*. Given two events A and B, if

$$A\,\delta\,B,$$

  we say that the event A is *real* (or *determined*) w.r.t. the event B

- To show that relativity implies the reality of the future, two (reasonably acceptable) assumptions are needed:

  1. Simultaneous events (in a given inertial frame) are real with respect to each other:

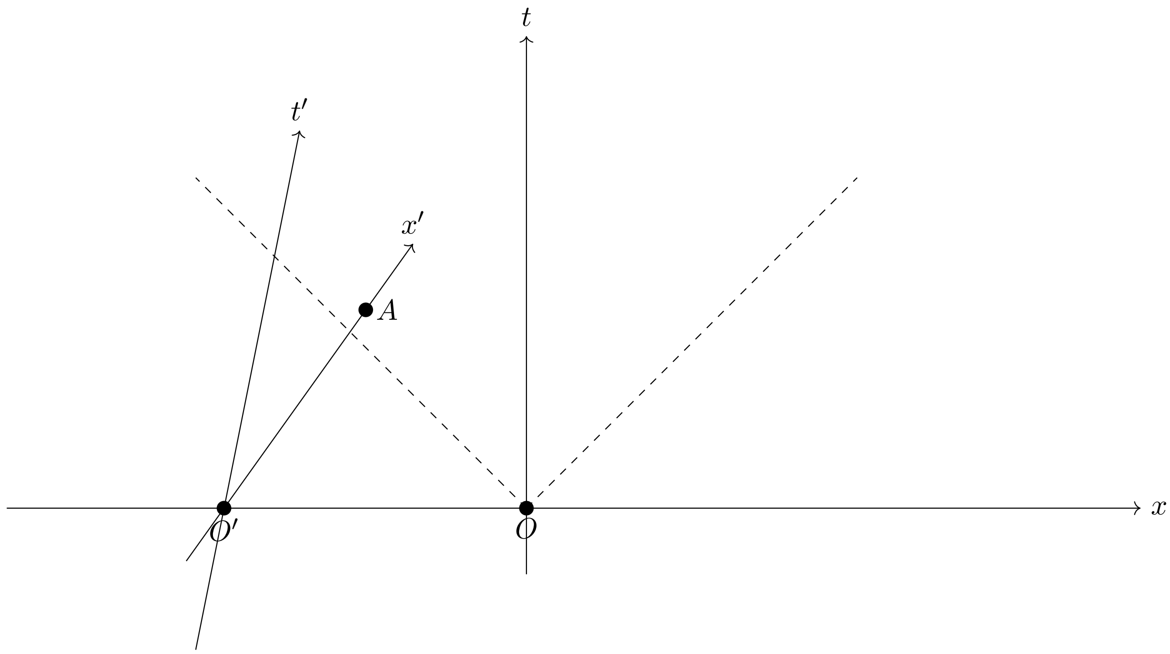
  $$A \text{ and } B \text{ simultaneous} \implies A\,\delta\,B$$

  2. The relation δ is transitive:

  $$A\,\delta\,B, \quad B\,\delta\,C \implies A\,\delta\,C$$

- Let O be the origin event of an inertial frame $\mathscr{R}$, and let A be an event in the (absolute) future of O. Consider another inertial frame $\mathscr{R}'$ such that its origin event O' lies on the simultaneity hyperplane of O in $\mathscr{R}$, and such that A lies on the simultaneity hyperplane of O' in $\mathscr{R}'$ (see figure below). We have just seen how to construct such a frame…

[6] *Time and physical geometry*, Journal of Philosophy (1967).



- A and O' are simultaneous in $\mathscr{R}'$ and O' and O are simultaneous in $\mathscr{R}$, so by assumption 1, we have

$$A\,\delta\,O', \qquad O'\,\delta\,O$$

and by transitivity (assumption 2), it follows that

$$A\,\delta\,O.$$

- Since A is an arbitrary event in the future of O, one concludes that all future events are real with respect to O. As O itself is arbitrary, one may conclude that all future events are real just as much as past and present events

- Notice that in the non-relativistic case, this construction is not possible: according to the classical worldview, the future is completely undetermined

- The conclusion drawn by Putnam is radical: the problem of the reality and determinateness of future events is resolved by physics rather than by philosophy

- Regardless of the degree to which one accepts this claim, the essential point is that relativity provides not only experimentally testable predictions, but also a perspective on reality that cannot be ignored in philosophical inquiry.

Poincaré transformation in Putnam's argument

- We want to construct a Poincaré transformation such that:
  - O' is simultaneous with O in $\mathscr{R}$,
  - A is simultaneous with O' in $\mathscr{R}'$

  The first condition implies

$$O' = (0, a), \qquad a \in \mathbb{R}$$

- Consider the displacement vector between O' and A:

$$x = A - O' = (ct_A,\, x_A - a)$$

  We require A and O' to be simultaneous in $\mathscr{R}'$. By Section 3.29, this is possible whenever x is space-like:

$$q(x) = c^2 t_A^2 - (x_A - a)^2 < 0 \iff |x_A - a| > ct_A$$

- Since A lies in the future lightcone of O, we have $|x_A| < ct_A$. Hence we can choose a such that

$$x_A - a > ct_A \quad \iff \quad a < x_A - ct_A$$

- Applying the result of Section 3.29 to the vector $x = (ct_A, x_A - a)$, we obtain the speed of the boost along the x-direction:

$$v = \frac{c^2 t_A}{x_A - a},$$

  which, by construction, is less than c, so the boost is admissible, and in the boosted inertial frame $\mathscr{R}'$ one has

$$t'_A = t'_{O'}$$

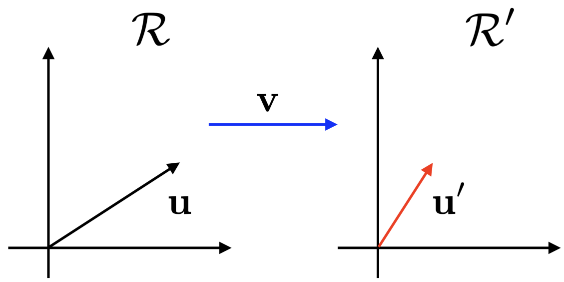
- In conclusion, fixing any $a < x_A - ct_A$ as shift vector and setting the speed of the boost along x as $v = (c^2 t_A)/(x_A - a)$, we obtain a Poincaré transformation such that

$$O \,\delta_{\mathscr{R}}\, O', \qquad O' \,\delta_{\mathscr{R}'}\, A,$$

  which is the geometric configuration required in Putnam's argument.

3.31 Relativistic addition of velocities

- Consider a particle, or a photon, moving with uniform velocity $\mathbf{u}$ in the inertial frame $\mathscr{R}$ and another inertial frames $\mathscr{R}'$ moving with relative velocity $\mathbf{v}$ along x, i.e. $\mathbf{v} = (v, 0, 0)$



- What is the velocity $\mathbf{u}'$ in $\mathscr{R}'$ as a function of $\mathbf{u}$ and $\mathbf{v}$?

- If we choose the origin of $\mathscr{R}$ on the worldline of the particle, then:

$$\mathbf{u} = \begin{pmatrix} u_x \\ u_y \\ u_z \end{pmatrix}, \quad \mathbf{u}' = \begin{pmatrix} u'_x \\ u'_y \\ u'_z \end{pmatrix} \implies \mathbf{x}(t) = \begin{pmatrix} x = u_x t \\ y = u_y t \\ z = u_z t \end{pmatrix}, \quad \mathbf{x}'(t) = \begin{pmatrix} x' = u'_x t \\ y' = u'_y t \\ z' = u'_z t \end{pmatrix}, \quad \gamma_v = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$$

applying the x-boost:

$$\begin{cases} t' = \gamma_v \left(t - \dfrac{v}{c^2} x \right) = \gamma_v t \left(1 - \dfrac{v}{c^2} u_x \right) \implies t = \dfrac{t'}{\gamma_v \left(1 - \frac{v}{c^2} u_x\right)} \\ x' = \gamma_v (x - vt) = \gamma_v (u_x - v) t = \dfrac{(u_x - v) t'}{1 - \dfrac{v}{c^2} u_x} \\ y' = u_y t = \dfrac{u_y t'}{\gamma_v \left(1 - \dfrac{v}{c^2} u_x \right)} \\ z' = u_z t = \dfrac{u_z t'}{\gamma_v \left(1 - \dfrac{v}{c^2} u_x \right)} \end{cases}$$

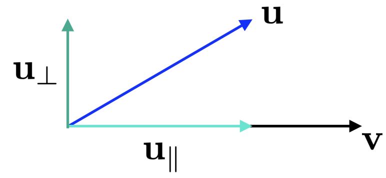
- Since $u'_x = x'/t'$ etc., we obtain the **relativistic addition of velocities**:

$$\begin{cases} v \oplus u_x := u'_x = \dfrac{u_x - v}{1 - \dfrac{v}{c^2} u_x} \\ v \oplus u_y := u'_y = \dfrac{1}{\gamma_v} \dfrac{u_y}{1 - \dfrac{v}{c^2} u_x} \\ v \oplus u_z := u'_z = \dfrac{1}{\gamma_v} \dfrac{u_z}{1 - \dfrac{v}{c^2} u_x} \end{cases} \tag{3.12}$$

- Galilean limit: $u'_x = u_x - v$, $u'_y = u_y$, $u'_z = u_z$

General formula for the relativistic addition of velocities

- In the standard configuration: $\mathbf{u} = (u_x, u_y, u_z)^T$ and $\mathbf{v} = (v, 0, 0)^T$, so $\mathbf{u} \cdot \mathbf{v} = u_x v$
- If $\mathbf{v}$ has a general orientation, then we decompose $\mathbf{u} = \mathbf{u}_{\parallel} + \mathbf{u}_{\perp}$



u_x becomes $\mathbf{u}_{\parallel}$ and $u_x v$ becomes $\mathbf{u} \cdot \mathbf{v}$. Thus, we arrive to the following equations for the relativistic addition of velocities:

$$\begin{cases} \mathbf{u}'_{\parallel} = \dfrac{\mathbf{u}_{\parallel} - \mathbf{v}}{1 - \dfrac{\mathbf{u}\cdot\mathbf{v}}{c^2}} \\ \mathbf{u}'_{\perp} = \dfrac{1}{\gamma_v} \dfrac{\mathbf{u}_{\perp}}{1 - \dfrac{\mathbf{u}\cdot\mathbf{v}}{c^2}} \end{cases} \tag{3.13}$$

so

$$\mathbf{v} \oplus \mathbf{u} = \mathbf{u}' = \mathbf{u}'_{\parallel} + \mathbf{u}'_{\perp} = \frac{1}{1 - \dfrac{\mathbf{u}\cdot\mathbf{v}}{c^2}} \left[\mathbf{u}_{\parallel} - \mathbf{v} + \frac{1}{\gamma_v} \mathbf{u}_{\perp} \right]$$

but

$$\mathbf{u}_{\parallel} = \frac{\mathbf{u}\cdot\mathbf{v}}{v^2} \mathbf{v}, \quad \mathbf{u}_{\perp} = \mathbf{u} - \mathbf{u}_{\parallel}$$

hence

$$\begin{aligned} \mathbf{u}_{\parallel} - \mathbf{v} + \frac{1}{\gamma_v}\mathbf{u}_{\perp} &= \mathbf{u}_{\parallel} - \mathbf{v} + \frac{1}{\gamma_v}(\mathbf{u} - \mathbf{u}_{\parallel}) = \frac{1}{\gamma_v}\mathbf{u} + \left(1 - \frac{1}{\gamma_v}\right)\mathbf{u}_{\parallel} - \mathbf{v} = \frac{1}{\gamma_v}\mathbf{u} + \left(1 - \frac{1}{\gamma_v}\right)\frac{\mathbf{u}\cdot\mathbf{v}}{v^2}\mathbf{v} - \mathbf{v} \\ &= \frac{1}{\gamma_v}\mathbf{u} - \left[1 - \left(1 - \frac{1}{\gamma_v}\right)\frac{\mathbf{u}\cdot\mathbf{v}}{v^2}\right]\mathbf{v} \end{aligned}$$

but

$$\begin{aligned} \frac{1}{v^2}\left(1 - \frac{1}{\gamma_v}\right) &= \frac{1}{v^2}\frac{\gamma_v - 1}{\gamma_v} = \frac{1}{v^2}\frac{\gamma_v^2 - 1}{\gamma_v(\gamma_v + 1)} = \frac{1}{v^2}\frac{\gamma_v^2\left(1 - \frac{1}{\gamma_v^2}\right)}{\gamma_v(\gamma_v + 1)} = \frac{1}{v^2}\frac{\gamma_v^2\frac{v^2}{c^2}}{\gamma_v(\gamma_v + 1)} \\ &= \frac{1}{c^2}\frac{\gamma_v}{1 + \gamma_v} \end{aligned}$$

so

$$\mathbf{v} \oplus \mathbf{u} = \mathbf{u}' = \frac{1}{1 - \dfrac{\mathbf{u}\cdot\mathbf{v}}{c^2}} \left[\frac{1}{\gamma_v}\mathbf{u} - \left(1 - \frac{\mathbf{u}\cdot\mathbf{v}}{c^2} \cdot \frac{\gamma_v}{1 + \gamma_v}\right)\mathbf{v} \right] \tag{3.14}$$

- The relativistic addition is **neither associative**, **nor commutative** unless the velocities are collinear, e.g. it can be proven that

$$\mathbf{u} \oplus \mathbf{v} = \frac{1}{1 - \dfrac{\mathbf{u}\cdot\mathbf{v}}{c^2}} \left[\frac{1}{\gamma_v}\mathbf{u} + \left(1 + \frac{\mathbf{u}\cdot\mathbf{v}}{c^2} \cdot \frac{\gamma_v}{1 + \gamma_v}\right)\mathbf{v} \right]. \tag{3.15}$$

3.32 Relativistic aberration of velocities

- Consider the situation depicted in the figure (John Walker). Without wind, the Earth surface is approximately at rest w.r.t. the rain, so it falls straight down from the sky, but when we move at 30 km/h the rain drops are tilted. This phenomenon is called (classical) aberration of velocities

- Now we derive the relativistic version of aberration to describes **how the direction of motion** of a particle (or light ray) **transforms between two inertial frames in relative motion** with velocity $\mathbf{v}$ in the Minkowski spacetime.

- The velocity of a particle or light ray seen from a reference frame not at rest with it is called **relative velocity**

- The analysis will be clearer if we start considering two special cases: $\mathbf{u} \parallel \mathbf{v}$ and $\mathbf{u} \perp \mathbf{v}$

Parallel motion

- We start with *parallel motion*: $\mathbf{u} \parallel \mathbf{v}$, as depicted in figure 3.1, i.e. $\mathbf{u} = \mathbf{u}_\parallel$, $\mathbf{u}_\perp = \mathbf{0}$ and using (3.13) we have that also $\mathbf{u}'_\perp = \mathbf{0}$, which means that the particle is seen to move along the same direction in both frames: the direction (not the magnitude) of relative velocity in $\mathscr{R}'$ is the same as the velocity in $\mathscr{R}$

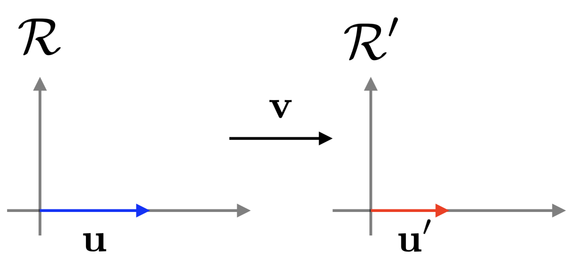


Figure 3.1: The special case of $\mathbf{u} = \mathbf{u}_\parallel$. Both $\mathbf{u}$ and $\mathbf{u}'$ are parallel to $\mathbf{v}$, no aberration occurs.

- Along the direction of motion, if $u = |\mathbf{u}|$ and $v = |\mathbf{v}|$ ($v < c$ as always because it's the relative frame speed), we have

$$u' = \frac{u - v}{1 - \dfrac{uv}{c^2}}$$

- If $u < c$, then

$$u' = c^2 \cdot \frac{u - v}{c^2 - uv} < c^2 \cdot \frac{c - v}{c\left(c - \frac{u}{c}v\right)} = c \cdot \frac{c - v}{c - \frac{u}{c}v} \underset{\frac{u}{c}<1}{<} c$$

so, no matter how u is close to c, $u' = v \oplus u$ will always be strictly less than c

- If $u = c$, then:

$$u' = \frac{c - v}{1 - \frac{v}{c}} = c \cdot \frac{c - v}{c - v} = c$$

so $c' = v \oplus c = c$, consistently with the invariance of light speed.

Orthogonal motion

- Consider now the other special case, that of orthogonal motion of the particle w.r.t. $\mathbf{v}$ as depicted in Figure 3.2 in which $\mathbf{u} = \mathbf{u}_\perp$ and $\mathbf{u}_\parallel = \mathbf{0}$

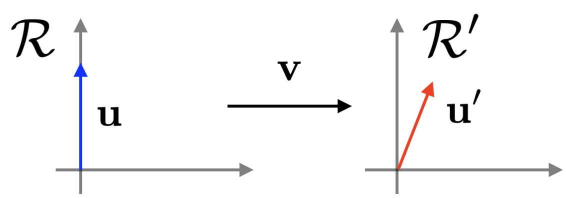


Figure 3.2: The special case of $\mathbf{u} = \mathbf{u}_\perp$. $\mathbf{u}'$ is not perpendicular to $\mathbf{v}$ anymore, thus showing aberration of velocities.

- Since now $\mathbf{u} \cdot \mathbf{v} = 0$, eqs. (3.13) give:

$$\begin{cases} \mathbf{u}'_\parallel = -\mathbf{v} \\ \mathbf{u}'_\perp = \dfrac{\mathbf{u}}{\gamma_v} \end{cases}$$

so, this time the particle doesn't move orthogonal to $\mathbf{v}$ also from the point of view of the frame $\mathscr{R}'$: both the magnitude and the direction of the relative velocity in $\mathscr{R}'$ are different than those of the velocity in $\mathscr{R}$, aberration occurred

- By the Pythagorean theorem

$$u' = \sqrt{|\mathbf{u}'|^2} = \sqrt{|\mathbf{u}'_\parallel|^2 + |\mathbf{u}'_\perp|^2} = \sqrt{v^2 + \frac{u^2}{\gamma_v^2}} \underset{\frac{1}{\gamma_v^2}=1-\frac{v^2}{c^2}}{=} \sqrt{u^2 + v^2 - \frac{u^2v^2}{c^2}},$$

the term in red corrects the Galilean addition of velocities and guarantees the preservation of light speed: if $u = c$, then $c' = v \oplus c = \sqrt{c^2 + v^2 - v^2} = c$.

The general configuration

- To analyze the general configuration, we need to come back to formula (3.14)

$$\mathbf{u}' = \frac{1}{1 - \dfrac{\mathbf{u}\cdot\mathbf{v}}{c^2}} \left[\frac{1}{\gamma_v}\mathbf{u} - \left(1 - \frac{\mathbf{u}\cdot\mathbf{v}}{c^2} \cdot \frac{\gamma_v}{1+\gamma_v}\right)\mathbf{v}\right],$$

  which shows that $\mathbf{u}' \in \operatorname{span}(\mathbf{u}, \mathbf{v})$, i.e. $\mathbf{u}'$ lies in the plane formed by $\mathbf{u}$ and $\mathbf{v}$

- $\vartheta, \vartheta' \in [0, \pi]$: angles between $\mathbf{u}$ and $\mathbf{v}$, and between $\mathbf{u}'$ and $\mathbf{v}$, then

$$\begin{cases} u_\parallel = u\cos\vartheta \\ u_\perp = u\sin\vartheta \end{cases}, \qquad \begin{cases} u'_\parallel = u'\cos\vartheta' \\ u'_\perp = u'\sin\vartheta' \end{cases},$$

  where $u_\parallel, u'_\parallel$ and $u_\perp, u'_\perp$ are the magnitudes of $\mathbf{u}_\parallel, \mathbf{u}'_\parallel$ and $\mathbf{u}_\perp, \mathbf{u}'_\perp$, respectively.

- From the equations

$$\mathbf{u}'_\parallel = \frac{\mathbf{u}_\parallel - \mathbf{v}}{1 - \dfrac{\mathbf{u}\cdot\mathbf{v}}{c^2}}, \qquad \mathbf{u}'_\perp = \frac{1}{\gamma_v}\frac{\mathbf{u}_\perp}{1 - \dfrac{\mathbf{u}\cdot\mathbf{v}}{c^2}},$$

  we obtain the magnitudes

$$u'_\parallel = \left|\mathbf{u}'_\parallel\right| = \frac{\left|\mathbf{u}_\parallel - \mathbf{v}\right|}{1 - \dfrac{\mathbf{u}\cdot\mathbf{v}}{c^2}} = \frac{\left|(u\cos\theta - v)\frac{\mathbf{v}}{v}\right|}{1 - \dfrac{\mathbf{u}\cdot\mathbf{v}}{c^2}} = \frac{u\cos\theta - v}{1 - \dfrac{\mathbf{u}\cdot\mathbf{v}}{c^2}}$$

$$u'_\perp = \left|\mathbf{u}'_\perp\right| = \frac{1}{\gamma_v}\frac{\left|\mathbf{u}_\perp\right|}{1 - \dfrac{\mathbf{u}\cdot\mathbf{v}}{c^2}} = \frac{1}{\gamma_v}\frac{u\sin\vartheta}{1 - \dfrac{\mathbf{u}\cdot\mathbf{v}}{c^2}}$$

  which lead to the **relativistic aberration formula for velocities**

$$\tan\vartheta' = \frac{u'_\perp}{u'_\parallel} = \frac{u\sin\theta}{\gamma_v(u\cos\theta - v)}, \qquad \tan\vartheta = \frac{u'\sin\theta}{\gamma_v(u'\cos\theta + v)}, \tag{3.16}$$

  unless $\vartheta = 0$, i.e. $\mathbf{u} = \mathbf{u}_\parallel$, we have $\tan\vartheta \neq \tan\vartheta'$, so $\vartheta \neq \vartheta'$ ($\vartheta, \vartheta' \in [0, \pi]$) and $\mathbf{u}$ and $\mathbf{u}'$ have different direction

- For light $u = c$, dividing numerator and denominator by c we find

$$\tan\vartheta' = \frac{\sin\vartheta}{\gamma_v(\cos\vartheta - \beta)} \qquad \text{(\textbf{relativistic aberration of light})}$$

- The relativistic aberration differs from the Galilean one only by the Lorentz factor at the denominator which *increases the aberration significantly as* $v \to c$

- For $v \to c$, velocities concentrates in a **narrow cone**, the closer v to c, the narrower the cone: in fact, $\gamma_v \to +\infty$ and $\tan\theta' \to 0$, so, from the perspective of $\mathscr{R}$, velocity directions in $\mathscr{R}'$ are distorted (crushed) toward $\mathbf{v}$.

3.33 Relativistic jets and the headlight effect

- The relativistic aberration of velocities in observable as **jets** in high-energy physics, see Figure 3.3 (from Steane page 24)

- It is also one of the components of the complicated Blandford–Znajek process that explains the astrophysical jets coming from the accretion disk of super massive black holes or from two merging neutron stars (gamma-ray bursts), see Figure 3.4 (NASA).

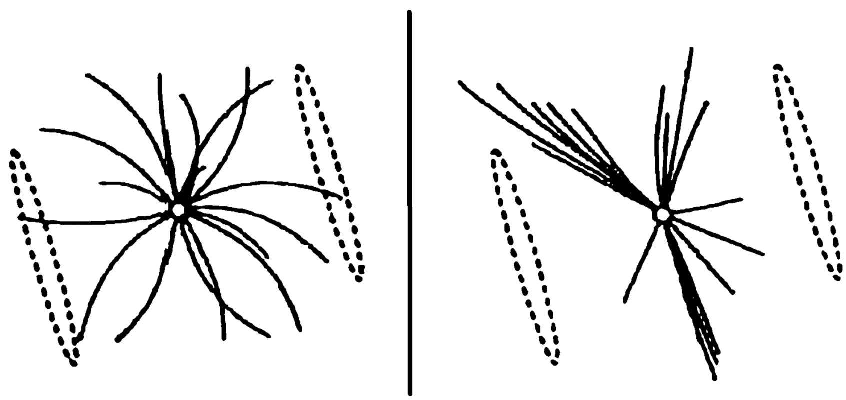


Figure 3.3: *Left*: non-relativistic decay: isotropic emission in the lab frame. *Right*: relativistic decay: emission angles are aberrated into a forward cone (jets) in the lab frame. The dashed ellipses indicate part of the cylindrical detector.

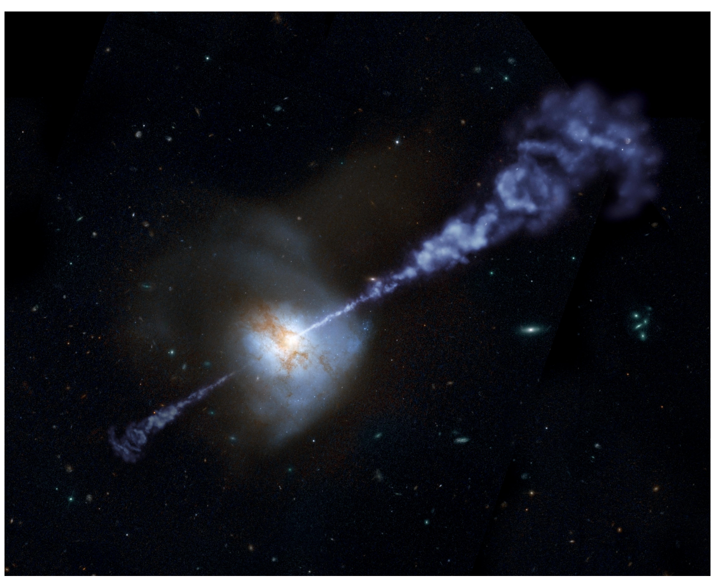


Figure 3.4: An astronomical jet.

- The relativistic aberration of light is responsible for the so-called headlight effect, depicted in Figure 3.4 (Wikipedia)

- The figure illustrates the relativistic headlight effect: as $\beta = v/c$ increases, an initially isotropic emission of light becomes progressively concentrated into a narrow forward cone along the direction of motion

- This angular compression is accompanied by a *Doppler* blue-shift in the forward direction and a red-shift in the backward one.

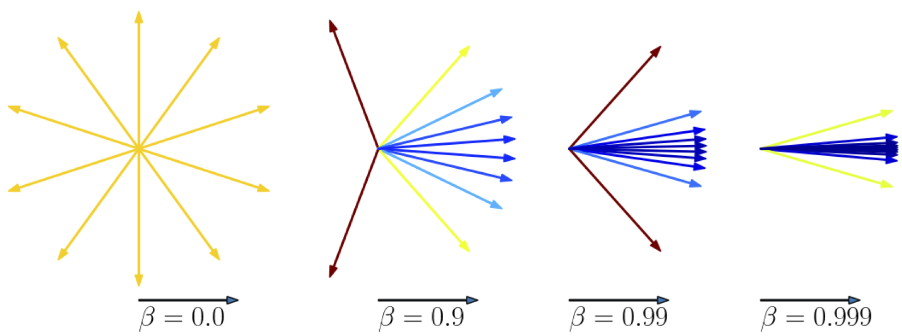


Figure 3.5: The headlight effect, depicted together with the Doppler effect.

3.34 The group of boosts along the same axis

- We can use the relativistic addition law of velocities to show that boosts along the same axis form a group w.r.t. composition

- To prove it, suppose we have three inertial reference frames $\mathscr{R}$, $\mathscr{R}'$ and $\mathscr{R}''$, in uniform relative motion along the x axis. $\mathscr{R}''$ moves with velocity v_2 with respect to $\mathscr{R}'$, which in turn moves with velocity v_1 with respect to $\mathscr{R}$

- The Lorentz transformations from $\mathscr{R}$ to $\mathscr{R}'$ are

$$x' = \gamma(v_1)(x - v_1 t), \qquad t' = \gamma(v_1)\left(t - \frac{v_1}{c^2}x\right)$$

while those from $\mathscr{R}'$ to $\mathscr{R}''$ are

$$x'' = \gamma(v_2)(x' - v_2 t'), \qquad t'' = \gamma(v_2)\left(t' - \frac{v_2}{c^2}x'\right)$$

so

$$\begin{aligned} x'' &= \gamma(v_2)\left[\gamma(v_1)(x - v_1 t) - v_2\gamma(v_1)\left(t - \frac{v_1}{c^2}x\right)\right] \\ &= \gamma(v_2)\gamma(v_1)\left[x - v_1 t - v_2 t + \frac{v_1 v_2}{c^2}x\right] \\ &= \gamma(v_2)\gamma(v_1)\left[\left(1 + \frac{v_1 v_2}{c^2}\right)x - (v_1 + v_2)t\right] \\ &= \gamma(v_2)\gamma(v_1)\left(1 + \frac{v_1 v_2}{c^2}\right)\left[x - \frac{v_1 + v_2}{1 + \frac{v_1 v_2}{c^2}}\,t\right] \end{aligned}$$

we recognize

$$v := \frac{v_1 + v_2}{1 + \frac{v_1 v_2}{c^2}}$$

to be the relativistic sum of v_1 and v_2 along the x-axis, so

$$x'' = \gamma(v_2)\gamma(v_1)\left(1 + \frac{v_1 v_2}{c^2}\right)(x - vt)$$

- Similarly,

$$\begin{aligned} t'' &= \gamma(v_2)\left[\gamma(v_1)\left(t - \frac{v_1}{c^2}x\right) - \frac{v_2}{c^2}\gamma(v_1)(x - v_1 t)\right] \\ &= \gamma(v_2)\gamma(v_1)\left[t - \frac{v_1}{c^2}x - \frac{v_2}{c^2}x + \frac{v_1 v_2}{c^2}t\right] \\ &= \gamma(v_2)\gamma(v_1)\left[\left(1 + \frac{v_1 v_2}{c^2}\right)t - \frac{v_1 + v_2}{c^2}x\right] \\ &= \gamma(v_2)\gamma(v_1)\left(1 + \frac{v_1 v_2}{c^2}\right)\left[t - \frac{1}{c^2}\frac{v_1 + v_2}{1 + \frac{v_1 v_2}{c^2}}\,x\right] \end{aligned}$$

hence

$$t'' = \gamma(v_2)\gamma(v_1)\left(1 + \frac{v_1 v_2}{c^2}\right)\left(t - \frac{v}{c^2}x\right)$$

and so the composed boosts transformation are

$$x'' = \gamma(v_2)\gamma(v_1)\left(1 + \frac{v_1 v_2}{c^2}\right)(x - vt)$$

$$t'' = \gamma(v_2)\gamma(v_1)\left(1 + \frac{v_1 v_2}{c^2}\right)\left(t - \frac{v}{c^2}x\right)$$

- With tedious computations, it can be proven that

$$\gamma(v) = \gamma(v_2)\gamma(v_1)\left(1 + \frac{v_1 v_2}{c^2}\right),$$

thanks to this, the previous expressions can be rewritten as

$$x'' = \gamma(v)(x - vt), \qquad t'' = \gamma(v)\left(t - \frac{v}{c^2}x\right)$$

and represent a boost with velocity $v = v_1 \oplus v_2$

- Therefore, composing two boosts with velocities v_1 and v_2 along the same axis, one obtains a boost of the same type with velocity v given by the relativistic addition of v_1 and v_2:

$$B(v_1) \cdot B(v_2) = B(v)$$

- By direct calculation, it can be verified that the composition law of boosts is associative

$$[B(v_1) \cdot B(v_2)] \cdot B(v_3) = B(v_1) \cdot [B(v_2) \cdot B(v_3)],$$

there exists the identity transformation (corresponding to $v = 0$)

$$B(0) = I$$

and, for every $B(v)$, there exists the inverse transformation $B^{-1}(v)$, given by

$$B^{-1}(v) = B(-v)$$

- To summarize, the set of boosts along the same axis is a group

- However, the set of boosts along arbitrary directions does not form a group because it is not closed under composition, as we'll show in section 3.37.

3.35 Connected components of the Lorentz group

- The full Lorentz group $O(1,3)$ has four connected components:
  - $SO^+(1,3) \equiv \mathscr{L}_+^\uparrow$: proper orthochronous
  - $SO^-(1,3)$: proper non-orthochronous
  - $O^+(1,3)$: improper orthochronous
  - $O^-(1,3)$: improper non-orthochronous
- The proper orthochronous transformations of $SO^+(1,3)$, already examined in Section 2.16, form the connected component of the identity. They are the very important Lorentz transformations in special relativity for three reasons
- First: **physical continuity**. Proper orthochronous transformations are those that can be obtained continuously from the identity by means of finite real parameters. This corresponds to the physical idea of passing from one inertial frame to another with subluminal velocities
- Second: **preservation of time orientation**: if a vector is future-directed, it remains so after the transformation
- Third: **preservation of spatial orientation** . "Proper" means determinant $+1$, hence no spatial reflections (parity) are introduced
- Finally, in relativistic field theory one requires the fundamental laws to be invariant at least under $SO^+(1,3)$; possible additional symmetries (parity P, time reversal T) are not mandatory and can in fact be violated
- In summary: proper orthochronous Lorentz transformations are the group of transformations between physically accessible inertial frames that respect causality in spacetime.

3.36 Decomposition of proper orthochronous Lorentz transformations

- Here we prove the possibility of writing in a unique way a generic proper orthochronous Lorentz transformation as the product of a Lorentz boost and a spatial rotation that fixes time, or the other way round

- We represent $\Lambda \in \mathscr{L}_+^\uparrow$ as follows:

$$\Lambda = \begin{pmatrix} \Lambda_{00} & \Lambda_{01} & \Lambda_{02} & \Lambda_{03} \\ \Lambda_{10} & \Lambda_{11} & \Lambda_{12} & \Lambda_{13} \\ \Lambda_{20} & \Lambda_{21} & \Lambda_{22} & \Lambda_{23} \\ \Lambda_{30} & \Lambda_{31} & \Lambda_{32} & \Lambda_{33} \end{pmatrix}$$

Theorem 3.36.1 *Every proper orthochronous Lorentz transformation $\Lambda \in \mathscr{L}_+^\uparrow$ can be uniquely decomposed as*

$$\Lambda = \mathcal{B}(\mathbf{v})\mathcal{R}, \tag{3.17}$$

where $B(\mathbf{v})$ represents a Lorentz boost in the $\mathbf{v}$-direction and $\mathcal{R}$ is an augmented spatial rotation that fixes the time coordinate and rotates the spatial ones, i.e.

$$\mathcal{R} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & R_{11} & R_{12} & R_{13} \\ 0 & R_{21} & R_{22} & R_{23} \\ 0 & R_{31} & R_{32} & R_{33} \end{pmatrix} \equiv \left(\begin{array}{c|c} 1 & \mathbf{0}^T \\ \hline \mathbf{0} & R \end{array}\right), \tag{3.18}$$

$\mathbf{v}$ *is fully determined by the first column of Λ as follows:*

$$\mathbf{v} = \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} = -\frac{c}{\Lambda_{00}} \begin{pmatrix} \Lambda_{10} \\ \Lambda_{20} \\ \Lambda_{30} \end{pmatrix}, \tag{3.19}$$

the Lorentz factor associated to $\mathbf{v}$ is the first matrix element of Λ

$$\gamma = \frac{1}{\sqrt{1 + \frac{|\mathbf{v}|^2}{c^2}}} = \Lambda_{00}, \tag{3.20}$$

finally $R \in \mathrm{SO}(3)$ and its entries depend on the matrix elements of Λ in this way:

$$R_{ij} = \Lambda_{ij} - \frac{\Lambda_{i0}\Lambda_{0j}}{\Lambda_{00} + 1}, \quad i, j = 1, 2, 3. \tag{3.21}$$

- *Proof.* First of all we notice that, thanks to theorem 2.16.1, if Λ is written as in (3.17) with $B(\mathbf{v})$ a boost and $R \in SO(3)$, then it is a proper orthochronous Lorentz transformation, as a composition of two such transformations[7]

[7]Notice that $\mathcal{R}$ leaves the temporal coordinate unchanged and preserves the norm of the spatial vector, thus it surely preserves the Minkowski norm.

- Now observe that if we know how to write the velocity vector $\mathbf{v}$ and the Lorentz factor γ in terms of the entries of Λ, then by eq. (3.5), the boost matrix is completely specified in terms of the elements of Λ

- To this aim let us write explicitly:

$$
\begin{gathered}
\Lambda = \mathcal{B}(\mathbf{v})\mathcal{R} = \begin{pmatrix} \gamma & -\frac{\gamma}{c}\mathbf{v}^T \\ -\frac{\gamma}{c}\mathbf{v} & I_3 + \frac{\gamma^2}{c^2(1+\gamma)}\mathbf{v}\mathbf{v}^T \end{pmatrix} \left(\begin{array}{c|c} 1 & \mathbf{0}^T \\ \hline \mathbf{0} & R \end{array}\right) \\
\begin{pmatrix} \gamma & -\frac{\gamma}{c}\mathbf{v}^T R \\ -\frac{\gamma}{c}\mathbf{v} & R + \frac{\gamma^2}{c^2(\gamma+1)}\mathbf{v}\mathbf{v}^T R \end{pmatrix} = \begin{pmatrix} \gamma & -\frac{\gamma}{c}(R^T\mathbf{v})^T \\ -\frac{\gamma}{c}\mathbf{v} & R + \frac{\gamma^2}{c^2(\gamma+1)}\mathbf{v}(R^T\mathbf{v})^T \end{pmatrix}
\end{gathered} \tag{3.22}
$$

- The equality between the first column of Λ and the first column of the matrix on the right-hand-side, i.e.

$$
\begin{pmatrix} \Lambda_{00} \\ \Lambda_{01} \\ \Lambda_{02} \\ \Lambda_{03} \end{pmatrix} = \begin{pmatrix} \gamma \\ -\frac{\gamma}{c}v_x \\ -\frac{\gamma}{c}v_y \\ -\frac{\gamma}{c}v_z \end{pmatrix},
$$

  implies eqs. (3.20) and (3.19), hence $\mathcal{B}(\mathbf{v})$ is completely determined from Λ

- Thanks to this result, we can determine R in terms of Λ simply by observing that $\mathcal{R} = B(\mathbf{v})^{-1}\Lambda = B(-\mathbf{v})\Lambda$. We first notice that

$$
\frac{\gamma^2}{c^2(\gamma+1)}\mathbf{v}\mathbf{v}^T = \frac{\cancel{\Lambda_{00}^2}}{\cancel{c^2}(\Lambda_{00}+1)}\frac{\cancel{c^2}}{\cancel{\Lambda_{00}^2}}\begin{pmatrix} \Lambda_{10} \\ \Lambda_{20} \\ \Lambda_{30} \end{pmatrix}(\Lambda_{10}\ \Lambda_{20}\ \Lambda_{30}) = \frac{1}{\Lambda_{00}+1}\begin{pmatrix} \Lambda_{10}^2 & \Lambda_{10}\Lambda_{20} & \Lambda_{10}\Lambda_{30} \\ \Lambda_{10}\Lambda_{20} & \Lambda_{20}^2 & \Lambda_{20}\Lambda_{30} \\ \Lambda_{10}\Lambda_{30} & \Lambda_{20}\Lambda_{30} & \Lambda_{30}^2 \end{pmatrix}
$$

  and that

$$
\frac{\gamma}{c}\mathbf{v} = \frac{\cancel{\Lambda_{00}}}{\cancel{c}}\left(-\frac{\cancel{c}}{\cancel{\Lambda_{00}}}\right)\begin{pmatrix} \Lambda_{10} \\ \Lambda_{20} \\ \Lambda_{30} \end{pmatrix} = -\begin{pmatrix} \Lambda_{10} \\ \Lambda_{20} \\ \Lambda_{30} \end{pmatrix},
$$

  hence eq. (3.6) gives:

$$
B(-\mathbf{v}) = \begin{pmatrix} \Lambda_{00} & -\Lambda_{10} & -\Lambda_{20} & -\Lambda_{30} \\ -\Lambda_{10} & 1 + \frac{\Lambda_{10}^2}{\Lambda_{00}+1} & \frac{\Lambda_{10}\Lambda_{20}}{\Lambda_{00}+1} & \frac{\Lambda_{10}\Lambda_{30}}{\Lambda_{00}+1} \\ -\Lambda_{20} & \frac{\Lambda_{10}\Lambda_{20}}{\Lambda_{00}+1} & 1 + \frac{\Lambda_{20}^2}{\Lambda_{00}+1} & \frac{\Lambda_{20}\Lambda_{30}}{\Lambda_{00}+1} \\ -\Lambda_{30} & \frac{\Lambda_{10}\Lambda_{30}}{\Lambda_{00}+1} & \frac{\Lambda_{20}\Lambda_{30}}{\Lambda_{00}+1} & 1 + \frac{\Lambda_{30}^2}{\Lambda_{00}+1} \end{pmatrix}
$$

and so

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & R_{11} & R_{12} & R_{13} \\ 0 & R_{21} & R_{22} & R_{23} \\ 0 & R_{31} & R_{32} & R_{33} \end{pmatrix} = \begin{pmatrix} \Lambda_{00} & -\Lambda_{10} & -\Lambda_{20} & -\Lambda_{30} \\ -\Lambda_{10} & 1+\frac{\Lambda_{10}^2}{\Lambda_{00}+1} & \frac{\Lambda_{10}\Lambda_{20}}{\Lambda_{00}+1} & \frac{\Lambda_{10}\Lambda_{30}}{\Lambda_{00}+1} \\ -\Lambda_{20} & \frac{\Lambda_{10}\Lambda_{20}}{\Lambda_{00}+1} & 1+\frac{\Lambda_{20}^2}{\Lambda_{00}+1} & \frac{\Lambda_{20}\Lambda_{30}}{\Lambda_{00}+1} \\ -\Lambda_{30} & \frac{\Lambda_{10}\Lambda_{30}}{\Lambda_{00}+1} & \frac{\Lambda_{20}\Lambda_{30}}{\Lambda_{00}+1} & 1+\frac{\Lambda_{30}^2}{\Lambda_{00}+1} \end{pmatrix} \begin{pmatrix} \Lambda_{00} & \Lambda_{01} & \Lambda_{02} & \Lambda_{03} \\ \Lambda_{10} & \Lambda_{11} & \Lambda_{12} & \Lambda_{13} \\ \Lambda_{20} & \Lambda_{21} & \Lambda_{22} & \Lambda_{23} \\ \Lambda_{30} & \Lambda_{31} & \Lambda_{32} & \Lambda_{33} \end{pmatrix}$$

- The explicit computation of R_{11} gives:

$$\begin{aligned} R_{11} &= -\Lambda_{10}\Lambda_{01} + \left(1+\frac{\Lambda_{10}^2}{\Lambda_{00}+1}\right)\Lambda_{11} + \frac{\Lambda_{10}\Lambda_{20}}{\Lambda_{00}+1}\Lambda_{21} + \frac{\Lambda_{10}\Lambda_{30}}{\Lambda_{00}+1}\Lambda_{31} \\ &= \Lambda_{11} - \Lambda_{10}\Lambda_{01} + \frac{\Lambda_{10}}{\Lambda_{00}+1}\left(\Lambda_{10}\Lambda_{11} + \Lambda_{20}\Lambda_{21} + \Lambda_{30}\Lambda_{31}\right), \end{aligned}$$

but, using the fact that Λ is a Lorentzian matrix, all the columns are Lorentz-orthogonal, in particular the first two, so

$$\Lambda_{00}\Lambda_{01} - \Lambda_{10}\Lambda_{11} - \Lambda_{20}\Lambda_{21} - \Lambda_{30}\Lambda_{31} = 0 \iff \Lambda_{10}\Lambda_{11} + \Lambda_{20}\Lambda_{21} + \Lambda_{30}\Lambda_{31} = \Lambda_{00}\Lambda_{01},$$

hence

$$\begin{aligned} R_{11} &= \Lambda_{11} - \Lambda_{10}\Lambda_{01} + \frac{\Lambda_{10}}{\Lambda_{00}+1}\Lambda_{00}\Lambda_{01} = \Lambda_{11} + \frac{\cancel{-\Lambda_{10}\Lambda_{01}\Lambda_{00}} - \Lambda_{10}\Lambda_{01} + \cancel{\Lambda_{10}\Lambda_{00}\Lambda_{01}}}{\Lambda_{00}+1} \\ &= \Lambda_{11} - \frac{\Lambda_{10}\Lambda_{01}}{\Lambda_{00}+1}. \end{aligned}$$

- Analogous calculations for the other matrix elements of R give

$$R = \begin{pmatrix} \Lambda_{11} - \frac{\Lambda_{10}\Lambda_{01}}{\Lambda_{00}+1} & \Lambda_{12} - \frac{\Lambda_{10}\Lambda_{02}}{\Lambda_{00}+1} & \Lambda_{13} - \frac{\Lambda_{10}\Lambda_{03}}{\Lambda_{00}+1} \\ \Lambda_{21} - \frac{\Lambda_{20}\Lambda_{01}}{\Lambda_{00}+1} & \Lambda_{22} - \frac{\Lambda_{20}\Lambda_{02}}{\Lambda_{00}+1} & \Lambda_{23} - \frac{\Lambda_{20}\Lambda_{03}}{\Lambda_{00}+1} \\ \Lambda_{31} - \frac{\Lambda_{30}\Lambda_{01}}{\Lambda_{00}+1} & \Lambda_{32} - \frac{\Lambda_{30}\Lambda_{02}}{\Lambda_{00}+1} & \Lambda_{33} - \frac{\Lambda_{30}\Lambda_{03}}{\Lambda_{00}+1} \end{pmatrix},$$

thus confirming eq. (3.21).

- Having obtained the explicit form of R, it can be verified by direct computation, using the properties satisfied by the matrix elements of the Lorentzian matrix Λ, that R belongs to SO(3), in fact $\det(R) = +1$ and, if C_i, $i = 1, 2, 3$, are the columns of R, then $\|C_i\|^2 = 1$ and $\langle C_i, C_j\rangle = 0$ for $i \neq j \in 1, 2, 3$, thus they form an orthonormal basis of $\mathbb{R}^3$ and so $R \in \mathrm{SO}(3)$

- Finally, we prove the uniqueness of the decomposition: suppose that there exists another couple of Lorentz-boost $B(\mathbf{v}')$ and time-fixing spatial rotation $\mathcal{R}'$ such that

$$\Lambda = \mathcal{B}(\mathbf{v})\mathcal{R} = \mathcal{B}(\mathbf{v}')\mathcal{R}'$$

- Using eq. (3.22) we can write

$$\begin{pmatrix} \gamma & -\frac{\gamma}{c}(R^T\mathbf{v})^T \\ -\frac{\gamma}{c}\mathbf{v} & R+\frac{\gamma^2}{c^2(\gamma+1)}\mathbf{v}(R^T\mathbf{v})^T \end{pmatrix} = \begin{pmatrix} \gamma_{v'} & -\frac{\gamma_{v'}}{c}((R')^T\mathbf{v}')^T \\ -\frac{\gamma_{v'}}{c}\mathbf{v}' & R'+\frac{\gamma_{v'}^2}{c^2(\gamma_{v'}+1)}\mathbf{v}'((R')^T\mathbf{v}')^T \end{pmatrix},$$

  which implies that $\gamma = \gamma_{v'}$ and so, from the equality of the matrix element of position $(2,1)$, it follows that $\mathbf{v} = \mathbf{v}'$. Since a boost matrix is completely determined by the expression of the Lorentz factor and the velocity, we have that $\mathcal{B}(\mathbf{v}) = \mathcal{B}(\mathbf{v}')$

- Moreover, from the equality of the matrix element of position $(1,2)$ we get $R^T\mathbf{v} = (R')^T\mathbf{v}'$ and this, together with the equality

$$R + \frac{\gamma^2}{c^2(\gamma+1)}\mathbf{v}(R^T\mathbf{v})^T = R' + \frac{\gamma_{v'}^2}{c^2(\gamma_{v'}+1)}\mathbf{v}'((R')^T\mathbf{v}')^T$$

  implies immediately that $R = R'$ and so $\mathcal{R} = \mathcal{R}'$

- Thus, the decomposition $\Lambda = \mathcal{B}(\mathbf{v})\mathcal{R}$ is unique for all $\Lambda \in \mathscr{L}_+^\uparrow$. □

- An analogous result holds for the decomposition of a proper orthochronous Lorentz transformation with the boost applied first.

Theorem 3.36.2 *Every proper orthochronous Lorentz transformation $\Lambda \in \mathscr{L}_+^\uparrow$ can be uniquely decomposed as*

$$\Lambda = \mathcal{R}\mathcal{B}(\mathbf{u}), \tag{3.23}$$

where $\mathcal{R}$ is the same matrix as in theorem 3.36.1 and $\mathcal{B}(\mathbf{u})$ represents a Lorentz boost in the $\mathbf{u}$-direction. Moreover, the velocities $\mathbf{u}$ and $\mathbf{v}$ of theorem 3.36.1 are related as follows:

$$\mathbf{v} = R\mathbf{u} \iff \mathbf{u} = R^T\mathbf{v}.$$

Corollary 3.36.1 *If two vectors $\mathbf{v}, \mathbf{u} \in \mathbb{R}^3$ with identical magnitude are related by $\mathbf{v} = R\mathbf{u}$, with $R \in SO(3)$, then the boosts with velocities given by the previous vectors are related as follows:*

$$\mathcal{B}(\mathbf{u}) = \mathcal{R}^T\mathcal{B}(\mathbf{v})\mathcal{R} \iff \mathcal{B}(\mathbf{v}) = \mathcal{R}\mathcal{B}(\mathbf{u})\mathcal{R}^T,$$

where $\mathcal{R}$ is the extended spatial rotation associated to R.

Proof. From the previous theorem we know that there is a unique $\Lambda \in \mathscr{L}_+^\uparrow$ such that $\Lambda = \mathcal{B}(\mathbf{v})\mathcal{R} = \mathcal{R}\mathcal{B}(R^T\mathbf{v}) = \mathcal{R}\mathcal{B}(\mathbf{u})$, so $\mathcal{B}(\mathbf{v})\mathcal{R} = \mathcal{R}\mathcal{B}(\mathbf{u})$. Since $\mathcal{R}^{-1} = \mathcal{R}^T$, it follows that

$$\mathcal{B}(\mathbf{u}) = \mathcal{R}^T\mathcal{B}(\mathbf{v})\mathcal{R}.$$

□

- To determine the explicit formula for the angle and the axis of rotation corresponding to the matrix R as a function of the entries of Λ, we need to recall the famous Rodrigues rotation formula

Interlude: Rodrigues rotation formula

- The Rodrigues rotation formula establishes that if $\mathbf{v}$ is a vector in $\mathbb{R}^3$ and $\mathbf{n}$ is a unit vector describing an axis of rotation about which $\mathbf{v}$ rotates counterclockwise by an angle ϑ, then the rotated vector $\mathbf{v}_{\text{rot}}$ can be written as:

$$\boxed{\mathbf{v}_{\text{rot}} = \cos\vartheta\,\mathbf{v} + (1-\cos\vartheta)\,\langle\mathbf{n},\mathbf{v}\rangle\,\mathbf{n} + \sin\vartheta\,(\mathbf{n}\times\mathbf{v})} \tag{3.24}$$

- In order to understand this quite involved formula, let us first notice that, given a vector $\mathbf{v} = (x,y)^T \in \mathbb{R}^2$ and a rotation matrix $R_\vartheta \in \mathrm{SO}(2)$, then

$$\begin{aligned} R_\vartheta\mathbf{v} &= \begin{pmatrix} \cos\vartheta & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\vartheta\,x - \sin\vartheta\,y \\ \sin\vartheta\,x + \cos\vartheta\,y \end{pmatrix} = \cos\vartheta\begin{pmatrix} x \\ y \end{pmatrix} + \sin\vartheta\begin{pmatrix} -y \\ x \end{pmatrix} \\ &= \cos\vartheta\,\mathbf{v} + \sin\vartheta\,R_{\frac{\pi}{2}}\mathbf{v}, \end{aligned}$$

  so the 2D rotation of a vector $\mathbf{v}$ can be seen as a linear combination between $\mathbf{v}$ and the vector $R_{\frac{\pi}{2}}\mathbf{v}$, which is orthogonal to $\mathbf{v}$ and have the same length

- Let us now increase the dimension and consider two vectors $\mathbf{v},\mathbf{n} \in \mathbb{R}^3$, where $\mathbf{n}$ is a unit vector orthogonal to $\mathbf{v}$. Then, $\mathbf{n}\times\mathbf{v}$ is orthogonal to the plane defined by $\mathbf{n}$ and $\mathbf{v}$ and

$$|\mathbf{n}\times\mathbf{v}| = \cancelto{1}{|\mathbf{n}|}\,|\mathbf{v}|\,\cancelto{1}{\sin\frac{\pi}{2}} = |\mathbf{v}|,$$

  thus, thanks to the previous considerations, the 2D counterclockwise rotation by the angle ϑ of the 3D vector $\mathbf{v}$ *in the plane orthogonal to the axis defines by* $\boldsymbol{n}$ is:

$$\mathrm{Rot}(\mathbf{v}) = \cos\vartheta\,\mathbf{v} + \sin\vartheta\,(\mathbf{n}\times\mathbf{v}) \tag{3.25}$$

- Finally, let us consider the 3D counterclockwise rotation by the angle ϑ of the 3D vector $\mathbf{v}$ around the axis defined by a unit vector $\mathbf{n}$, this time in a generic position w.r.t. $\mathbf{v}$:

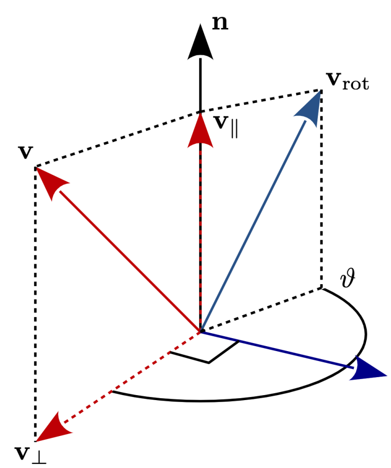


Figure 3.6: Interpretation of Rodrigues' rotation formula. Adapted from Wikipedia.

- Taking advantage of what we have just learned, a valid strategy to obtain the formula for the rotated vector $\mathbf{v}_{\text{rot}}$ is to decompose $\mathbf{v}$ into its parallel

$$\mathbf{v}_{\parallel} = \langle \mathbf{n}, \mathbf{v} \rangle \mathbf{n}$$

and orthogonal

$$\mathbf{v}_{\perp} = \mathbf{v} - \langle \mathbf{n}, \mathbf{v} \rangle \mathbf{n}$$

components w.r.t. $\mathbf{n}$, then perform the 2D rotation of the perpendicular component (as we have learned above) and finally add the result to $\mathbf{v}_{\parallel}$ to recover the correct height of $\mathbf{v}_{\text{rot}}$ as depicted in Figure 3.6

- Using eq. (3.25), the 2D rotation of $\mathbf{v}_{\perp}$ gives:

$$\begin{aligned}
\text{Rot}(\mathbf{v}_{\perp}) &= \cos\vartheta\, \mathbf{v}_{\perp} + \sin\vartheta\, (\mathbf{n} \times \mathbf{v}_{\perp}) = \cos\vartheta\, (\mathbf{v} - \langle \mathbf{n}, \mathbf{v} \rangle \mathbf{n}) + \sin\vartheta\, (\mathbf{n} \times (\mathbf{v} - \langle \mathbf{n}, \mathbf{v} \rangle \mathbf{n})) \\
&= \cos\vartheta\, \mathbf{v} - \cos\vartheta\, \langle \mathbf{n}, \mathbf{v} \rangle \mathbf{n} + \sin\vartheta\, (\mathbf{n} \times \mathbf{v}) - \sin\vartheta\, \langle \mathbf{n}, \mathbf{v} \rangle \cancelto{\mathbf{0}}{(\mathbf{n} \times \mathbf{n})} \\
&= \cos\vartheta\, \mathbf{v} - \cos\vartheta\, \langle \mathbf{n}, \mathbf{v} \rangle \mathbf{n} + \sin\vartheta\, (\mathbf{n} \times \mathbf{v}),
\end{aligned}$$

so, finally,

$$\mathbf{v}_{\text{rot}} = \mathbf{v}_{\parallel} + \text{Rot}(\mathbf{v}_{\perp}) = \cos\vartheta\, \mathbf{v} + (1 - \cos\vartheta)\, \langle \mathbf{n}, \mathbf{v} \rangle\, \mathbf{n} + \sin\vartheta\, (\mathbf{n} \times \mathbf{v}),$$

thus confirming the Rodrigues formula (3.24)

- From the Rodrigues rotation formula we can obtain a useful parameterization of a rotation matrix $R \in \text{SO}(3)$. In fact, by expanding the formula we obtain:

$$\begin{aligned}
\mathbf{v}_{\text{rot}} &= \cos\vartheta \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} + (1 - \cos\vartheta)(n_1 v_1 + n_2 v_2 + n_3 v_3) \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix} + \sin\vartheta \begin{pmatrix} n_2 v_3 - v_3 n_2 \\ n_3 v_1 - n_1 v_3 \\ n_1 v_2 - n_2 v_1 \end{pmatrix} \\
&= \begin{pmatrix} [\cos\vartheta + (1-\cos\vartheta) n_1^2] v_1 + [(1-\cos\vartheta) n_1 n_2 - \sin\vartheta n_3] v_2 + [(1-\cos\vartheta) n_1 n_3 + \sin\vartheta n_2] v_3 \\ [(1-\cos\vartheta) n_1 n_2 + \sin\vartheta n_3] v_1 + [\cos\vartheta + (1-\cos\vartheta) n_2^2] v_2 + [(1-\cos\vartheta) n_2 n_3 - \sin\vartheta n_1] v_3 \\ [(1-\cos\vartheta) n_1 n_3 - \sin\vartheta n_2] v_1 + [(1-\cos\vartheta) n_2 n_3 + \sin\vartheta n_1] v_2 + [\cos\vartheta + (1-\cos\vartheta) n_3^2] v_3 \end{pmatrix},
\end{aligned}$$

i.e. $\mathbf{v}_{\text{rot}} = R_{\vartheta,\mathbf{n}} \mathbf{v}$ with

$$R_{\vartheta,\mathbf{n}} = \begin{pmatrix} \cos\vartheta + (1-\cos\vartheta) n_1^2 & (1-\cos\vartheta) n_1 n_2 - \sin\vartheta n_3 & (1-\cos\vartheta) n_1 n_3 + \sin\vartheta n_2 \\ (1-\cos\vartheta) n_1 n_2 + \sin\vartheta n_3 & \cos\vartheta + (1-\cos\vartheta) n_2^2 & (1-\cos\vartheta) n_2 n_3 - \sin\vartheta n_1 \\ (1-\cos\vartheta) n_1 n_3 - \sin\vartheta n_2 & (1-\cos\vartheta) n_2 n_3 + \sin\vartheta n_1 & \cos\vartheta + (1-\cos\vartheta) n_3^2 \end{pmatrix},$$

called **Rodrigues rotation matrix**.

Coming back to the boost decomposition...

- The trace of the Rodrigues rotation matrix $R_{\vartheta,\mathbf{n}}$ is

$$\mathrm{Tr}(R_{\vartheta,\mathbf{n}}) = R_{11} + R_{22} + R_{33} = 3\cos\vartheta + (1-\cos\vartheta)(n_1^2+n_2^2+n_3^2),$$

  but $\mathbf{n}$ is a unit vector, so $\mathrm{Tr}(R_{\vartheta,\mathbf{n}}) = 2\cos\vartheta + 1$, i.e.

$$\cos\vartheta = \frac{\mathrm{Tr}(R_{\vartheta,\mathbf{n}}) - 1}{2}$$

- Recall now from eq. (3.21) that the matrix elements of the rotation matrix appearing in the decomposition of the proper orthochronous Lorentz transformation are

$$R_{ij} = \Lambda_{ij} - \frac{\Lambda_{i0}\Lambda_{0j}}{\Lambda_{00}+1}, \quad i,j = 1,2,3,$$

  so we find

$$\boxed{\cos\vartheta = \frac{\Lambda_{11}+\Lambda_{22}+\Lambda_{33}-1}{2} - \frac{\Lambda_{10}\Lambda_{01}+\Lambda_{20}\Lambda_{02}+\Lambda_{30}\Lambda_{03}}{2(\Lambda_{00}+1)}}$$

- Then we observe that the Rodrigues matrix is such that

$$R_{32} - R_{23} = 2\sin\vartheta n_1,\ R_{13} - R_{31} = 2\sin\vartheta n_2,\ R_{21} - R_{12} = 2\sin\vartheta n_3,$$

  so

$$\sin\vartheta\mathbf{n} = \frac{1}{2}\begin{pmatrix} R_{32}-R_{23} \\ R_{13}-R_{31} \\ R_{21}-R_{12} \end{pmatrix},$$

  so, using again eq. (3.21),

$$\boxed{\sin\vartheta\,\mathbf{n} = \begin{pmatrix} \frac{\Lambda_{32}-\Lambda_{23}}{2} - \frac{\Lambda_{30}\Lambda_{02}-\Lambda_{20}\Lambda_{03}}{2(\Lambda_{00}+1)} \\ \\ \frac{\Lambda_{13}-\Lambda_{31}}{2} - \frac{\Lambda_{10}\Lambda_{03}-\Lambda_{30}\Lambda_{01}}{2(\Lambda_{00}+1)} \\ \\ \frac{\Lambda_{21}-\Lambda_{12}}{2} - \frac{\Lambda_{20}\Lambda_{01}-\Lambda_{10}\Lambda_{02}}{2(\Lambda_{00}+1)} \end{pmatrix}}$$

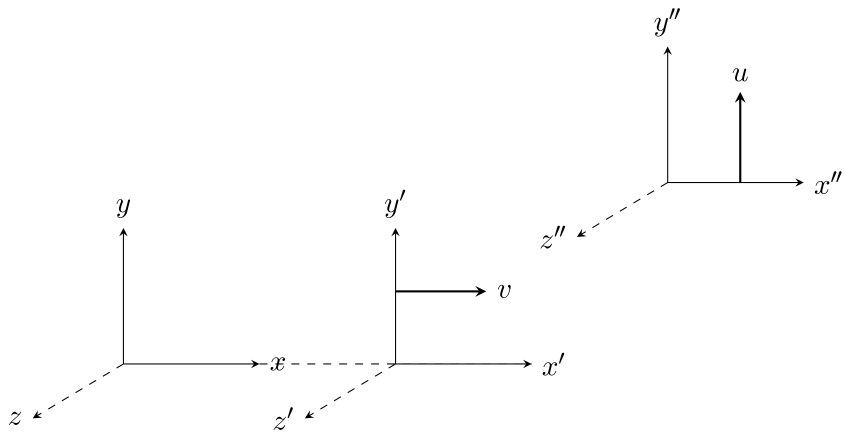
- The boxed formulae show how to obtain the rotation angle and axis of R appearing in the decomposition $\Lambda = \mathcal{B}(\mathbf{v})R$ from the matrix elements of Λ.

3.37 Composition of non-collinear boosts and Wigner rotation

- We have seen in 3.34 that boosts along the same axis form a group. This is no longer true if the boosts are performed along different directions. Indeed, composing two boosts of this type one does not obtain another boost but a boost and a rotation. This fact is at the basis of a phenomenon of great importance, the Thomas precession

- Let us consider three inertial reference frames $\mathscr{R}$, $\mathscr{R}'$ and $\mathscr{R}''$ in relative uniform motion. $\mathscr{R}'$ moves with respect to $\mathscr{R}$ with velocity v along the x axis $\mathscr{R}''$ moves with respect to $\mathscr{R}'$ with velocity u along the y' axis. We denote

$$\gamma_v = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}, \qquad \gamma_u = \frac{1}{\sqrt{1-\frac{u^2}{c^2}}}$$



- The Lorentz transformations from $\mathscr{R}$ to $\mathscr{R}'$ are

$$\begin{cases} x' = \gamma_v(x - vt) \\ y' = y \\ z' = z \\ t' = \gamma_v\left(t - \dfrac{vx}{c^2}\right) \end{cases}$$

- The Lorentz transformations from $\mathscr{R}'$ to $\mathscr{R}''$ are

$$\begin{cases} x'' = x' \\ y'' = \gamma_u(y' - ut') \\ z'' = z' \\ t'' = \gamma_u\left(t' - \dfrac{uy'}{c^2}\right) \end{cases}$$

- Combining these transformations one obtains

$$\begin{cases} x'' = \gamma_v(x - vt) \\ y'' = \gamma_u y - \gamma_u\gamma_v u\left(t - \dfrac{vx}{c^2}\right) \\ z'' = z \\ t'' = \gamma_u\gamma_v\left(t - \dfrac{vx}{c^2}\right) - \gamma_u\dfrac{uy}{c^2} \end{cases}$$

- We must now check whether these transformation laws correspond to a boost, that is whether they can be reduced to eqs. (3.3), (3.4), i.e.

$$\mathbf{x}' = \mathbf{x} + \left[(\gamma - 1)\frac{\mathbf{x}\cdot\mathbf{v}}{|\mathbf{v}|^2} - \gamma t\right]\mathbf{v}$$

$$ct' = \gamma\left(ct - \frac{\mathbf{x}\cdot\mathbf{v}}{c}\right)$$

with $\mathbf{x}'$ replaced by $\mathbf{x}''$, t' replaced by t'' and γ replaced by $\gamma(w) = (1 - w^2/c^2)^{-1/2}$, where $w = |\mathbf{w}|$ is the speed of $\mathscr{R}''$ with respect to $\mathscr{R}$

- $\mathbf{w}$ is obtained using (3.15), i.e.

$$\mathbf{u} \oplus \mathbf{v} = \frac{1}{1 - \dfrac{\mathbf{u}\cdot\mathbf{v}}{c^2}}\left[\frac{1}{\gamma}\mathbf{u} + \left(1 + \frac{\mathbf{u}\cdot\mathbf{v}}{c^2}\cdot\frac{\gamma}{1+\gamma}\right)\mathbf{v}\right]$$

with $\mathbf{u} = \mathbf{u}_\perp$ and $\gamma = \gamma_v$, which gives

$$\begin{cases} w_x = v \\ w_y = \dfrac{u}{\gamma_v} \\ w_z = 0 \end{cases}$$

and so

$$w^2 = v^2 + \frac{u^2}{\gamma_v^2}, \quad 1 - \frac{w^2}{c^2} = 1 - \frac{v^2}{c^2} - \frac{u^2}{\gamma_v^2 c^2},$$

using

$$\frac{1}{\gamma_v^2} = 1 - \frac{v^2}{c^2},$$

it follows that

$$1 - \frac{w^2}{c^2} = \left(1 - \frac{v^2}{c^2}\right)\left(1 - \frac{u^2}{c^2}\right) = \frac{1}{\gamma_v^2}\frac{1}{\gamma_u^2},$$

therefore,

$$\gamma(w) = \frac{1}{\sqrt{1 - \frac{w^2}{c^2}}} = \gamma_v\gamma_u$$

- We have

$$x'' = x + \left[(\gamma_w - 1)\frac{\mathbf{x}\cdot\mathbf{w}}{w^2} - \gamma_w t\right]w_x$$

with

$$\mathbf{x} \cdot \mathbf{w} = xw_x + yw_y + zw_z = vx + \frac{u}{\gamma_v} y$$

so

$$x'' = x + \frac{(\gamma_w - 1)v}{w^2} \left(vx + \frac{u}{\gamma_v} y \right) - \gamma_w vt, \tag{3.26}$$

which does not have the same form that we found by composing the boosts in the two orthogonal directions, i.e. $x'' = \gamma_v(x - vt)$, noticeably due to the lack of the y component

- It follows that the composition of non-collinear boosts is not a boost anymore!

- However, it is a proper orthochronous Lorentz transformation. Therefore, by Theorem 3.36.1, it can be uniquely decomposed as a rotation followed by a boost

- To find them, let's consider the Lorentz matrix Λ associated to the composition of the x-boost and the y-boost

$$\Lambda = \begin{pmatrix} \gamma_v\gamma_u & -\gamma_v\gamma_u\frac{v}{c^2} & -\gamma_u\frac{u}{c^2} & 0 \\ -\gamma_v v & \gamma_v & 0 & 0 \\ -\gamma_v\gamma_u u & \gamma_v\gamma_u\frac{uv}{c^2} & \gamma_u & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

- The boost velocity is obtained from the first column

$$\mathbf{w} = -\frac{c}{\Lambda_{00}} \begin{pmatrix} \Lambda_{10} \\ \Lambda_{20} \\ \Lambda_{30} \end{pmatrix} = \begin{pmatrix} v \\ \frac{u}{\gamma_v} \\ 0 \end{pmatrix}, \qquad \gamma(w) = \Lambda_{00} = \gamma_v\gamma_u$$

in agreement with what we've found

- The rotation is obtained from

$$R_{ij} = \Lambda_{ij} - \frac{\Lambda_{i0}\Lambda_{0j}}{\Lambda_{00} + 1}$$

which gives

$$R = \begin{pmatrix} \dfrac{\gamma_v + \gamma_u}{\gamma_v\gamma_u + 1} & -\dfrac{\gamma_v\gamma_u uv/c^2}{\gamma_v\gamma_u + 1} & 0 \\ \dfrac{\gamma_v\gamma_u uv/c^2}{\gamma_v\gamma_u + 1} & \dfrac{\gamma_v + \gamma_u}{\gamma_v\gamma_u + 1} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

- This is a rotation in the xy-plane, hence around the z-axis, with

$$\cos\Omega = \frac{\gamma_v + \gamma_u}{\gamma_v\gamma_u + 1}, \qquad \sin\Omega = \frac{\gamma_v\gamma_u uv/c^2}{\gamma_v\gamma_u + 1}$$

- Therefore, the final decomposition (which can be checked by direct computation) is

$$\mathcal{B}_{y'}(u)\,\mathcal{B}_x(v) = \mathcal{B}(\mathbf{w})\,\mathcal{R}_z(\Omega)$$

where $\mathcal{R}_z(\Omega)$ is the augmented matrix built from R that preserves time and

$$\mathbf{w} = \left(v, \frac{u}{\gamma_v}, 0\right) \quad \text{and} \quad \tan\Omega = \frac{\gamma_v\gamma_u uv/c^2}{\gamma_v + \gamma_u}$$

- The interpretation is the following: since composing non-collinear boosts doesn't create a boost, after boosting the frame $\mathscr{R}'$ parallel to $\mathscr{R}$ and then $\mathscr{R}''$ parallel to $\mathscr{R}'$, we produce a precession, i.e. $\mathscr{R}$ sees the axes of $\mathscr{R}''$ rotated. This is a purely relativistic effect

- Of course, there's nothing special about the x-axis and the y'-axis, and in fact it can be proven with an analogous strategy that the Lorentz transformation that describes the composition of two non-collinear boosts $\mathcal{B}(\mathbf{v}_1)$, $\mathcal{B}(\mathbf{v}_2)$ is written as the composition of a boost along a velocity vector $\mathbf{w}$ belonging to the plane generated by $\mathbf{v}_1$ and $\mathbf{v}_2$ and a rotation w.r.t. the axis orthogonal to this plane, i.e. the axis defined by $\mathbf{v}_1 \times \mathbf{v}_2$

- This has an important consequence: suppose a particle to which a vector quantity $\mathbf{A}$ is associated moves along a closed orbit in an inertial frame $\mathscr{R}$. The frame $\mathscr{R}'$ comoving with the particle is instant by instant rotated with respect to $\mathscr{R}$, so it experiences infinitesimal non-collinear boosts. At the completion of the orbit, the vector $\mathbf{A}$ undergoes a precession (or gyration), called **Thomas precession** which has important effect on the dynamics of the spin of electrons.

Chapter 4

Relativistic dynamics

4.1 4-vectors, manifestly covariant equations and Lorentz invariants

- **Kinematics**: description of motion without worrying about its causes

- **Dynamics**: study of the causes of motion and how interactions influence it

- In special relativity, physical laws must be valid in all inertial frames. Since these are related by Lorentz transformations, physical laws must have the same form if we transform the quantities that appear in them via Lorentz transformations

- The concept of 4-vector allows us to write physical laws with this property

- A **4-vector** is a vector $V \in \mathbb{R}^{1,3}$ whose components (V^0, V^1, V^2, V^3) transform under an inertial change of reference as follows

$$V' = \Lambda V \iff V'^{\mu} = \Lambda^{\mu}_{\nu} V^{\nu}, \qquad \mu, \nu = 0, \dots, 3$$

  where Λ^{μ}_{ν} are the entries of a *restricted*[1] *Lorentz transformations* $\Lambda \in \mathrm{SO}^+(1,3) \equiv \mathscr{L}_+^{\uparrow}$

- 4-vectors allow us to determine **covariant equations**, whose form remains unchanged in all inertial reference frames. In fact, consider two 4-vectors V and W, so that

$$V'^{\mu} = \Lambda^{\mu}{}_{\nu} V^{\nu} \quad \text{and} \quad W'^{\mu} = \Lambda^{\mu}{}_{\nu} W^{\nu},$$

  where Λ is the restricted Lorentz transformation from an inertial frame $\mathscr{R}$ to another inertial frame $\mathscr{R}'$

- Suppose that the equation $V^{\mu} = W^{\mu}$ holds in $\mathscr{R}$, then

$$V'^{\mu} = \Lambda^{\mu}{}_{\nu} V^{\nu} = \Lambda^{\mu}{}_{\nu} W^{\nu} = W'^{\mu},$$

  i.e. $V^{\mu} = W^{\mu}$ remains valid in every other reference frame $\mathscr{R}'$. The equation *co-varies* with the frame in such a way that its form remains intact

[1] transformations under $\mathrm{O}^+(1,3)$ define pseudo-4-vectors.

- Of course, we can rewrite $V^\mu = W^\mu$ as $V^\mu - W^\mu = 0$, or $Z^\mu = 0$, where $Z = V - W$, so also a law in which a 4-vector is equated to 0 is covariant

- Due to the simplicity in which covariance can be determined from the equality between 4-vectors, in the physical literature $V^\mu = W^\mu$ or $Z^\mu = 0$ are called **manifestly covariant** equations

- Manifestly covariant equations permits an intrinsic description of motion in special relativity and, more generally, appear as *equalities between 4-tensors*...

- A quantity that remains invariant under a Lorentz transformation is called a Lorentz invariant

- Since Lorentz transformations preserve the quadratic Minkowski form, for each 4-vector V, **the squared Minkowski norm**

$$q(V) = V \circ V = V^T \eta V$$

  is a Lorentz invariant, in fact:

$$V' \circ V' = (V')^T \eta V' = (\Lambda V)^T \eta \Lambda V = V^T (\Lambda^T \eta \Lambda) V = V^T \eta V = V \circ V.$$

4.2 The fundamental 4-vectors

4.2.1 4-displacement

- Recall that an event $e = (ct, x, y, z)^T$ is a point of the Minkowski space $\mathcal{M}$, but once a reference frame and an origin $O = (0, \mathbf{0})^T$ in $\mathcal{M}$ are selected, we obtain a 4-vector

$$X = e - O = (ct, x, y, z)^T$$

called either **4-displacement** or **coordinate 4-vector**

- The Lorentz invariant associated to an infinitesimal 4-displacement is obviously $\ldots ds^2$:

$$dX \circ dX = dX^T \eta dX = (dx^\mu)^T \eta_{\mu\nu} dx^\nu = ds^2$$

- $X(t)$: 4-displacement describing a given time-like worldline. Recall from eq. (3.11)

$$d\tau^2 = \frac{ds^2}{c^2}, \qquad d\tau = \frac{dt}{\gamma}$$

so, also **the proper time is a Lorentz-invariant** and we get an **intrinsic temporal description** of worldlines replacing the coordinate time t (which depends on the reference frame) with[2] the **proper time** $\tau = t/\gamma$.

4.2.2 4-velocity

- Using the 4-vector X and the Lorentz scalar τ we define the 4-velocity (4-tangent vector to the wordline)

$$U := \frac{dX}{d\tau} = \gamma\frac{dX}{dt} = \gamma\frac{d}{dt}\begin{pmatrix} ct \\ \mathbf{x} \end{pmatrix} = \begin{pmatrix} \gamma c \\ \gamma\mathbf{u} \end{pmatrix},$$

where $\mathbf{u} := d\mathbf{x}/dt$, $u := |\mathbf{u}| = (\mathbf{u}\cdot\mathbf{u})^{1/2}$ and $\gamma = \gamma(u)$

- U is a genuine 4-vector because $d\tau' = d\tau$ and, written $X = x^\mu$, there exists $\Lambda \in \mathscr{L}_+^\uparrow$ such that

$$dx'^\mu = \Lambda^\mu{}_\nu \, dx^\nu \implies U'^\mu = \frac{dx'^\mu}{d\tau'} = \frac{\Lambda^\mu{}_\nu \, dx^\nu}{d\tau} = \Lambda^\mu{}_\nu \, U^\nu$$

- The Lorentz invariant of the 4-velocity is… the only invariant square speed $\ldots c^2$:

$$U \circ U = U^T \eta \, U = \begin{pmatrix} \gamma c, & \gamma\mathbf{u}^T \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -I_3 \end{pmatrix} \begin{pmatrix} \gamma c \\ \gamma\mathbf{u} \end{pmatrix} = \gamma^2 \left(c^2 - \mathbf{u}\cdot\mathbf{u}\right) = \frac{c^2 - u^2}{1 - \frac{u^2}{c^2}} = c^2.$$

[2] in γ appears the velocity in the same reference frame of the coordinate time t.

4.2.3 4-acceleration

- **4-acceleration** (curvature 4-vector of the worldline):

$$A := \frac{dU}{d\tau} = \frac{d^2X}{d\tau^2} = \cdots = \gamma^2 \begin{pmatrix} \dfrac{\mathbf{u}\cdot\mathbf{a}}{c}\gamma^2 \\ \dfrac{\mathbf{u}\cdot\mathbf{a}}{c^2}\gamma^2\mathbf{u} + \mathbf{a} \end{pmatrix}, \quad (\text{with } \mathbf{a} := \frac{d\mathbf{u}}{dt})$$

- The Lorentz invariant for A is less straightforward than in the previous two cases

- The 4-velocity U and the 4-acceleration A are **Lorentz-orthogonal**:

$$U \circ A = 0$$

If $X(t)$ is the displacement along a time-like wordline, then U is a time-like 4-vector, hence, by section 2.10, **the 4-acceleration is a space-like 4-vector**

- The proof shows the **quite astonishing algebraic role** played by the Lorentz factor:

$$\begin{aligned} U \circ A &= U^0A^0 - \mathbf{U}\cdot\mathbf{A} \\ &= \gamma c\Big(\gamma^4\frac{\mathbf{u}\cdot\mathbf{a}}{c}\Big) - \gamma\mathbf{u}\cdot\Big(\gamma^4\frac{\mathbf{u}\cdot\mathbf{a}}{c^2}\mathbf{u} + \gamma^2\mathbf{a}\Big) \\ &= \gamma^5(\mathbf{u}\cdot\mathbf{a}) - \gamma^5\frac{u^2}{c^2}(\mathbf{u}\cdot\mathbf{a}) - \gamma^3(\mathbf{u}\cdot\mathbf{a}) \\ &= \gamma^3(\mathbf{u}\cdot\mathbf{a})\big[\gamma^2(1-u^2/c^2) - 1\big] = \gamma^3(\mathbf{u}\cdot\mathbf{a})\big[1-1\big] \\ &= 0. \end{aligned}$$

4.2.4 4-momentum

- The 4-momentum P replaces the 3-momentum $\mathbf{p}$ in Newtonian mechanics

- For a particle of mass $m > 0$ that moves along the curve $\mathbf{x}(t)$ *in space* with 3-velocity $\mathbf{u} = d\mathbf{x}/dt$, the 3-momentum is $\mathbf{p} = m\mathbf{u}$

- Given a time-like worldline $X(\tau)$ *in spacetime*, P must transform as a 4-vector and be collinear with the 4-velocity U

- The most natural candidate is:

$$P := mU = \begin{pmatrix} \gamma mc \\ \gamma m\mathbf{u} \end{pmatrix}$$

- The interpretation given by Einstein of P and the analysis of its consequences are extraordinary examples of his *unmatched physical insight*!

4.3 The 0-th component of the 4-momentum and the mass-energy equivalence relation

- Einstein noticed that the spatial components have the dimension of $\mathbf{p}$ and tend exactly to it in the Galilean limit ($\gamma \to 1$)

- The 0-th component: if γmc is further multiplied by c, it has the same dimension of an *energy*, so

$$P^0 = \frac{E}{c}$$

- To deduce E, Einstein considered the Galilean limit: for $u \ll c$, $\beta = u/c \to 0$ and the binomial expansion gives

$$\gamma = \frac{1}{\sqrt{1-\beta^2}} = \left(1-\beta^2\right)^{-1/2} = 1 + \frac{1}{2}\beta^2 + \frac{3}{8}\beta^4 + \mathcal{O}(\beta^6)$$

so

$$\gamma mc^2 = mc^2 + \frac{1}{2}mu^2 + \frac{3}{8}m\,\frac{u^4}{c^2} + \mathcal{O}\left(\frac{u^6}{c^4}\right) \approx mc^2 + \frac{1}{2}mu^2$$

- To summarize: the 0-th component of the 4-momentum is E/c, where E is the Newtonian kinetic energy of the particle plus an extra term independent of motion

$$P := mU = \begin{pmatrix} \gamma mc \\ \gamma m\mathbf{u} \end{pmatrix} = \begin{pmatrix} \frac{E}{c} \\ \gamma m\mathbf{u} \end{pmatrix},$$

where

$$E = mc^2 + \frac{1}{2}mu^2$$

is called the **relativistic energy** of the particle

- In the rest frame of the particle $u = 0$, so we find[3]

$$E_0 = mc^2 \qquad \textbf{mass-energy equivalence relation}$$

E_0: **rest energy**, intrinsic energy of a particle at rest just because it has mass m, called **rest mass**.

[3]The third **unification triumph** in history, after Newton and Maxwell (gravity and electromagnetism).

4.4 The energy–momentum relation

- The Lorentz-invariant scalar corresponding to the 4-momentum P is:

$$
\begin{aligned}
P \circ P = P^\mu \, \eta_{\mu\nu} \, P^\nu &= (P^0)^2 - \mathbf{P} \cdot \mathbf{P} = (\gamma mc)^2 - (\gamma m)^2 \mathbf{u} \cdot \mathbf{u} = \gamma^2 m^2 \, (c^2 - u^2) \\
&= m^2 \frac{c^2 - u^2}{1 - \frac{u^2}{c^2}} = m^2 c^2 \frac{c^2 - u^2}{c^2 - u^2} \\
&= m^2 c^2
\end{aligned}
$$

- Equivalently, writing $P = \left(\frac{E}{c}, \mathbf{p}\right)$, we have:

$$
P \circ P = \left(\frac{E}{c}\right)^2 - |\mathbf{p}|^2
$$

- Equating the two results we obtain:

$$
E^2 = (mc^2)^2 + c^2 |\mathbf{p}|^2,
$$

  called **energy–momentum relation**.

4.4.1 Consequence 1: massless particles must travel at light speed

- Dividing the three spatial vector components of P by E, we obtain the purely kinematic identity

$$
\frac{\mathbf{p}}{E} = \frac{\mathbf{u}}{c^2} \quad \Longleftrightarrow \quad \frac{|\mathbf{p}|}{E} = \frac{|\mathbf{u}|}{c^2}, \tag{4.1}
$$

  which holds for all particles, including massless ones

- For massless particles, the energy-momentum relation gives $E = c|\mathbf{p}|$, so eq. (4.1) becomes

$$
\frac{|\mathbf{p}|}{c|\mathbf{p}|} = \frac{|\mathbf{u}|}{c^2} \quad \Longleftrightarrow \quad |\mathbf{u}| = c \qquad (m = 0),
$$

  implying that **massless particles are constrained to move at speed** c**!**

- Besides **photons**, only **gravitons** and **gluons** (the particles that mediate gravity and the strong nuclear force between quarks inside protons and neutrons) 'are' predicted[4] to be massless.

[4] Gluons hide because of nuclear confinement, and gravitons have not been observed yet.

4.4.2 Consequence 2: the relativistic mass is not constant

- We can write

$$E = \gamma(u)mc^2 \equiv m_{\text{rel}}\, c^2,$$

  where the **relativistic mass** is

$$m_{\text{rel}} = m_{\text{rel}}(u) = \gamma(u)m = \frac{m}{\sqrt{1 - \frac{u^2}{c^2}}}$$

- Unlike the electric charge, the inertial mass of a particle, i.e. its resistance to be accelerated, is not constant anymore in special relativity, but it increases with its velocity, in particular

$$\lim_{u \nearrow c} m_{\text{rel}}(u) = +\infty,$$

  so the particle opposes an infinite resistance to be accelerated to the speed of light.

4.4.3 Consequence 3: particles with mass > 0 can't travel at light speed

- From an energetic perspective:

$$E(u) = \frac{mc^2}{\sqrt{1 - \frac{u^2}{c^2}}},$$

  accelerating means increasing u, and

$$\lim_{u \nearrow c} E(u) = +\infty,$$

  which means that **accelerating a particle of any mass $m > 0$ to c would require an infinite amount of energy**.