Special Relativity

Introduction

Time and space are two of the most basic concepts in describing the world around us. Even young children soon master the difference between ”here” and ”there”, and between ”now” and ”later” or ”earlier”. However, strictly speaking, the concepts of time and space are only derived notions, which we introduce as a theoretical construct to help order our observations of events in the world around us.

Because our intuitions about time and space are so strong, it is hard to imagine a description of the world without them, or to explain in detail how those constructs come about. It should not be surprising, therefore, that time and space were long thought to have a fixed structure, which has been formalised by Newton, Galileo and Leibniz in the early days of modern science. These notions are now called ”absolute time” and ”absolute space” (or ”absolute distance”). Immanuel Kant could not imagine how time and space could possibly have any other structure and he even went so far as to raise that structure to a fundamental ordering principle of human thought.

Interestingly, Newton himself noted that absolute time and distance were not entirely without problems. He illustrated this in the form of the ”rotating bucket” paradox. However, this paradox did not seem to admit any experimental investigation, so it remained unresolved for a long time and didn't diminish the impression that the notions of absolute time and space were correct.

It should be considered a major achievement of human thought that the notions of absolute time and space were found to be faulty and that they could be replaced by something better: the concept of spacetime. It was only possible after Maxwell completed the unification of the theories of electricity and magnetism into a single theory of electromagnetism, based on Faraday's concept of a field, which avoids action at a distance. Maxwell's theory brought to light a new problem with the notions of absolute time and space: the theory was not invariant under all Galilean transformations. Einstein[1] found that this discrepancy could be resolved in the most elegant way

[1]Albert Einstein won the 1921 Nobel Prize in physics “for his services to theoretical

by noting that the concepts of absolute time and space were not well motivated in an operational way. In his Special Theory of Relativity he replaced them by the new concept of spacetime.

Time and Space in Classical Mechanics

For comparison with Special Relativity, it will be useful to give a fairly detailed analysis of the structure of time and space in the classical mechanics of Newton. Let us denote by M the set of all possible events in the universe. Here an event can be identified uniquely by a time and place, indicating when and where the event happens. We usually encode the time and place into numbers, which we call coordinates.

To find out when an event $p \in M$ happens, we may use a clock: when p happens we read off the time, $t(p)$. In this way the clock defines a function $t : M \to \mathbb{R}$. To understand the notion of absolute time, we need to understand how much of the number $t(p)$ depends on the properties of the clock we use, and how much of it is independent of those properties.

Let us first agree to use an ideal clock: it runs for ever and is not influenced by anything happening around it. (Of course such ideal clocks don't really exist, but they help to clarify the notion of absolute time.) Ideal clocks go a long way to exhibit the structure of absolute time, but two ideal clocks may still give different readings for two reasons:

1. Two ideal clocks can disagree which events happen at $t = 0$.

2. Two ideal clocks can use different units of time.

In other words, given an ideal clock $t : M \to \mathbb{R}$, there is no special physical significance to events p with $t(p) = 0$, or on the time difference $t(p_1) - t(p_2)$ for any two events p_1, p_2. However, all ideal clocks do agree on the following:

Tenet 2.1 (Absolute Time) *In classical mechanics, for arbitrary events $p_1, p_2, p'_1, p'_2 \in M$, all ideal clocks $t : M \to \mathbb{R}$ agree on*

1. *whether $t(p_1) > t(p_2)$, $t(p_1) < t(p_2)$ or $t(p_1) = t(p_2)$,*

2. *the value of $\frac{t(p_1)-t(p_2)}{t(p'_1)-t(p'_2)}$ when $t(p'_1) \neq t(p'_2)$.*

---

physics, and especially for his discovery of the law of the photoelectric effect".

The first property defines a time orientation on M: the notions of before, after and simultaneous are well defined for all events, independent of the ideal clock. In the second property, the denominator $t(p_1') - t(p_2')$ essentially fixes the units of time. When the units of time are fixed, all ideal clocks agree on the time interval between any two events. Together these two properties fully characterise absolute time. (Note in particular that we can compare the duration of any two time intervals, regardless of when or where the events are located.)

Exercise 2.2 *Let $t : M \to \mathbb{R}$ and $t' : M \to \mathbb{R}$ be time coordinates defined by ideal clocks. Show that $t'(p) = rt(p) + t_0'$ for some $r > 0$ and $t_0' \in \mathbb{R}$.*

In an analogous way one may identify places in M using ideal measuring rods, which are infinitely extended and which are not influenced by any physical processes happening in the universe. These ideal measuring rods can be used to form an infinite grid, where the rods intersect at right angles. The grid defines coordinate maps $\mathbf{x} : M \to \mathbf{R}^3$, where $\mathbf{x} = (x^1, x^2, x^3)$ in components.

For a given absolute time coordinate t we denote by $M_{t=c}$ the subset of M of all events taking place at time $t = c$. All these simultaneous events make up the entire space at the moment $t = c$. The notion of absolute space in classical mechanics can then be formulated as follows:

Tenet 2.3 (Absolute Space) *In classical mechanics, for any $c \in \mathbb{R}$ and any events $p_1, p_2, p_1', p_2' \in M_{t=c}$, all ideal measuring rods agree on the following:*

1. *the value of $\frac{\|\mathbf{x}(p_1)-\mathbf{x}(p_2)\|}{\|\mathbf{x}(p_1')-\mathbf{x}(p_2')\|}$, when $\|\mathbf{x}(p_1') - \mathbf{x}(p_2')\| \neq 0$,*
2. *$M_{t=c}$ is a three-dimensional Euclidean space, i.e. we may construct a grid of ideal measuring rods such that the distance between p_1 and p_2 satisfies Pythagoras' Theorem*

$$\|\mathbf{x}(p_1) - \mathbf{x}(p_1)\| = \sqrt{\sum_{i=1}^{3}(x^i(p_1) - x^i(p_2))^2}.$$

Again the origin of the coordinates $\mathbf{x}$ does not have any physical significance, nor does the orientation. The first property means that all ideal grids agree on the (absolute) distances, up to a change of units. The second property fixes the global shape of space $M_{t=c}$.

Exercise 2.4 *For fixed $c \in \mathbb{R}$, let $\mathbf{x} : M_{t=c} \to \mathbb{R}^3$ and $\mathbf{x}' : M_{t=c} \to \mathbb{R}^3$ be bijections, such that*

$$\frac{\|\mathbf{x}'(p_1) - \mathbf{x}'(p_2)\|}{\|\mathbf{x}'(p_1') - \mathbf{x}'(p_2')\|} = \frac{\|\mathbf{x}(p_1) - \mathbf{x}(p_2)\|}{\|\mathbf{x}(p_1') - \mathbf{x}(p_2')\|}$$

for all $p_1, p_2, p_1', p_2' \in M_{t=c}$ with $\|\mathbf{x}'(p_1')) - \mathbf{x}'(p_2'))\| \neq 0$. Show that

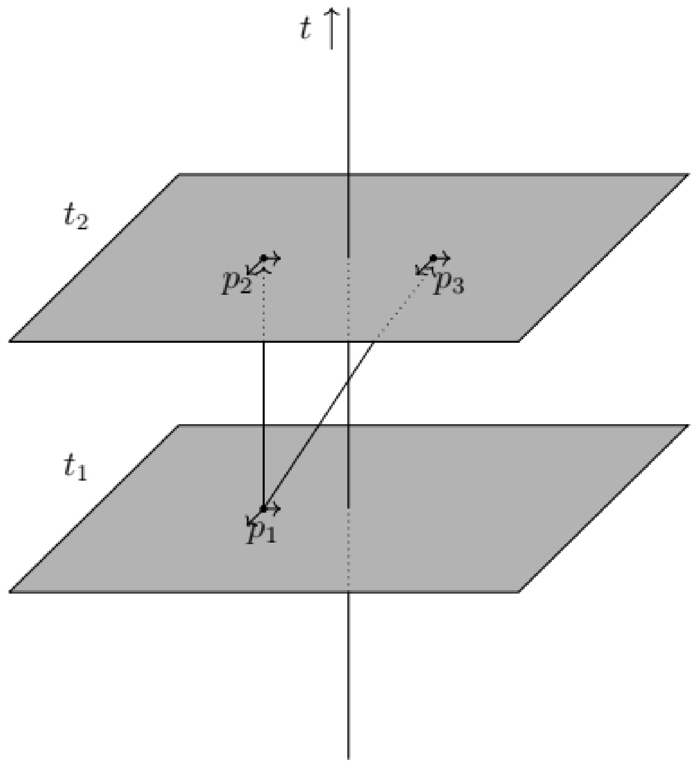
$$\mathbf{x}'(p) = rA \cdot \mathbf{x}(p) + \mathbf{x}_0'$$

for some $\mathbf{x}_0' \in \mathbb{R}^3$ and some orthogonal matrix A and some $r > 0$. (Hint: first identify $\mathbf{x}_0'$ and $r > 0$. Then consider the map $\psi : \mathbb{R}^3 \to \mathbb{R}^3 : X \mapsto r^{-1}(\mathbf{x}' \circ \mathbf{x}^{-1}(X) - \mathbf{x}_0')$. Using the real polarisation identity for the standard inner product, $2\langle X, Y\rangle = \|X\|^2 + \|Y\|^2 - \|X - Y\|^2$ for all $X, Y \in \mathbb{R}^3$, show that ψ must preserve the inner product, because it preserves all distances. Use this to prove that ψ is linear and to find A.)

We can use an ideal clock and an orthogonal grid of ideal measuring rods to define a coordinate system $(t, \mathbf{x}) : M \to \mathbb{R}^4$. In this way, each event in M is assigned a unique set of coordinates. Coordinates which are obtained in this way are called *classical inertial coordinates*. The following is a central tenet of classical mechanics:

Tenet 2.5 (Galilean Relativity Principle) *In all classical inertial coordinate systems, the physics of classical mechanics is described by Newton's laws in terms of absolute time and distance.*

In other words, the classical inertial coordinate systems form a preferred class of coordinate systems and Newton's laws are invariant under the change of coordinates from one classical inertial coordinate system to another. However, within classical mechanics, no classical inertial coordinate system is preferred over any other, simply because one cannot distinguish them by any experimental procedure satisfying Newton's laws. E.g., (approximate) classical inertial coordinates attached to the Earth are no better or worse than the (approximate) classical inertial coordinates attached to a train moving with a constant velocity in a fixed direction. Our natural intuition to favour the first coordinate system, which is more familiar, cannot be justified by classical mechanics.



Because classical inertial coordinate systems are defined by ideal clocks and rods, they are insensitive to any outside forces, so (by Newton's laws) they have zero acceleration with respect to each other. If we express the inertial coordinates $(t', \mathbf{x}')$ in the same physical units as the inertial coordinates $(t, \mathbf{x})$ (which may require a rescaling), then one can show that they must be related by a *Galilean transformation*:

$$t'(p) = t(p) + t'_0, \qquad \mathbf{x}'(p) = A \cdot \mathbf{x}(p) - t(p)\mathbf{v} + \mathbf{x}'_0 \tag{1}$$

for an orthogonal matrix A and a vector $\mathbf{v} \in \mathbb{R}^3$. Conversely any Galilean transformation applied to a classical inertial coordinate system yields a new classical inertial coordinates system.

Exercise 2.6 *Show that two classical inertial coordinate systems which employ the same units are related by a Galilean transformation with uniquely determined A, $\mathbf{v}$, $\mathbf{x}'_0$ and t'_0. (Hint: use the previous exercises and the fact that a trajectory in M which is linear in one classical inertial coordinate system is linear in every other classical inertial coordinate system, due to Newton's laws.)*

Exercise 2.7 *Show that the Galilean transformations form a group (i.e. the composition and the inverses of such transformations are Galilean transformations).*

Exercise 2.8 *Let $(t, \mathbf{x})$ be classical inertial coordinates and let p_1 and p_2 be the events such that $\mathbf{x}(p_1) = \mathbf{x}(p_2)$, but $t(p_2) > t(p_1)$. Does it make sense to say that the event p_2 happens at the same place as p_1 but at a later time? (I.e. would all inertial coordinate systems agree that p_2 is at the same place as p_1?) Do points of space preserve their identity in the course of time?*

Example 2.9 (Rotating bucket paradox) *Let $(t, \mathbf{x})$ be a classical inertial coordinate system and consider new coordinates $(t', \mathbf{x}')$ which rotate around a fixed axis with a constant angular velocity, e.g.:*

$$t'(p) = t(p), \qquad \mathbf{x}'(p) = R_{t(p)} \cdot \mathbf{x}(p),$$

where $R_{t(p)}$ is a rotation around the x^3-axis over an angle $t(p)$. Note that $(t', \mathbf{x}')$ is not a classical inertial coordinate system. Expressing physical laws in the coordinates $(t', \mathbf{x}')$ leads to different formulae than in the coordinates $(t, \mathbf{x})$ (including a centrifugal force).

Newton already noted that a bucket of water in an otherwise empty universe, which stands still w.r.t. the coordinates $(t', \mathbf{x}')$, rotates w.r.t. $(t, \mathbf{x})$. The water in the bucket experiences a centrifugal force, which causes it to rise against the walls of the bucket. Without rotation, the water would remain flat. Note that the difference cannot be caused by any other matter in the universe, so it must be due to rotations w.r.t. empty space. This effect is surprising, because a rotation w.r.t. empty space is difficult to see. This problem is called "Newton's rotating bucket paradox".

Electromagnetism and Poincaré Invariance

A first modification of our understanding of time and space was initiated when Maxwell completed the unification of the theories of electricity and magnetism. The theory of electromagnetism concerns an electric field $\mathbf{E}$ and a magnetic field $\mathbf{B}$ on the spacetime M, which satisfy Maxwell's equations of motion. In vacuum these equations take the form

$$\begin{aligned} \nabla \times \mathbf{E} = -\partial_t \mathbf{B}, \qquad \nabla \cdot \mathbf{B} = 0 \\ \nabla \cdot \mathbf{E} = 0, \qquad \nabla \times \mathbf{B} = \frac{1}{c^2}\partial_t \mathbf{E}. \end{aligned} \tag{2}$$

expressed in inertial coordinates $(t, \mathbf{x})$ fixed to Earth.

One of the nice aspects of Maxwell's theory of electromagnetism is that it is a field theory[2], which implements the intuition that physical influences should propagate from point to neighbouring point. Indeed, Maxwell's equations imply that all components of $\mathbf{E}$ and $\mathbf{B}$ satisfy the wave equation:

$$\Box \mathbf{E} := -\frac{1}{c^2}\partial_t^2 \mathbf{E} + \Delta \mathbf{E} = 0, \tag{3}$$

where $\Delta := \sum_{i=1}^{3} \partial_{x^i}^2$ is the Laplace operator and $\Box$ is called the d'Alembert operator. This equation ensures that disturbances in the (components of the) fields can propagate no faster than a certain maximum speed c, the speed of light in vacuum.

Maxwell's equations have another property, however, which seems very puzzling in comparison to classical mechanics: the equations are not invariant under all Galilean transformations. Although there is no problem with translations or rotations in space, the equations are not invariant under uniform motion. In particular, the speed of light in vacuum follows from Maxwell's equations, so if these equations were invariant, all inertial coordinate systems would have to agree on this speed. However, for a coordinate system moving fast in the opposite direction of a light ray, one might expect the (relative) speed of that light ray to be bigger!

Exercise 3.1 *Consider a Galilean transformation to a coordinate system $(t', \mathbf{x}')$ which moves at a constant speed v in the x^1 direction. Show that the wave equation (3) transforms to the equation*

$$\Box' \mathbf{E} = v^2 \partial_{x'}^2 \mathbf{E} + 2v\partial_{t'}\partial_{x'}\mathbf{E},$$

where $\Box'$ is defined in a similar way as $\Box$, but with respect to the new coordinates. Show in a similar way that Maxwell's equations transform to

$$\nabla' \times \mathbf{E} = -\partial_{t'}\mathbf{B} - v\partial_{x'}\mathbf{B}, \qquad \nabla' \cdot \mathbf{B} = 0$$
$$\nabla' \cdot \mathbf{E} = 0, \qquad \nabla' \times \mathbf{B} = \frac{1}{c^2}(\partial_{t'}\mathbf{E} + v\partial_{x'}\mathbf{E}).$$

so Maxwell's equations are not invariant under this Galilean transformation. (This conclusion remains true, even if we argue that the components of $\mathbf{E}$

[2]The concept of a field is due to Faraday.

and **B** *are components of vectors, so they should also transform under the change of coordinates. We will see in Section 9.1 that several components of the fields should be unaffected by the change of coordinates, but the wave equation for those coordinates is not invariant.)*

The mathematical modifications needed to reconcile electromagnetism with absolute time and space were soon understood: Maxwell's equations are invariant under *Poincaré transformations*:

$$\begin{aligned} t'(p) &= \gamma\left(t(p) - c^{-2}v(A\cdot\mathbf{x}(p))_{\parallel}\right) + t'_0, \\ \mathbf{x}'(p) &= (A\cdot\mathbf{x}(p))_{\perp} + \gamma\left((A\cdot\mathbf{x}(p))_{\parallel} - t(p)v\right)\mathbf{n} + \mathbf{x}'_0 \\ \gamma &:= (1 - c^{-2}\|\mathbf{v}\|^2)^{-\frac{1}{2}}, \end{aligned} \tag{4}$$

where A is an orthogonal matrix, $\mathbf{v} = v\mathbf{n}$ a velocity vector with speed $v = \|\mathbf{v}\| < c$, and $_{\parallel}$ and $_{\perp}$ denote the components of a vector parallel or perpendicular to the unit vector $\mathbf{n}$. When $\mathbf{v} = 0$ the two transformations coincide, but for uniform motion this is not so. When the inhomogeneous terms t'_0 and $\mathbf{x}'_0$ vanish, we speak of a *Lorentz transformation*. Lorentz transformations for which $A = I$ are called *boosts* with relative velocity $\mathbf{v} \neq 0$.

Exercise 3.2 *Show that any Poincaré transformation can be written as a composition of an orthogonal transformation A of the spatial coordinates, a boost with relative velocity* $\mathbf{v}$ *and a translation in spacetime along* $(t'_0, \mathbf{x}'_0)$. *Show that A,* $\mathbf{v}$ *and* $(t'_0, \mathbf{x}'_0)$ *are uniquely determined.*

Exercise 3.3 *Show that the wave equation (3) is invariant under Poincaré transformations. (Hint: by the previous exercise it suffices to consider orthogonal transformations, boosts and translations.)*

Exercise 3.4 *Consider a boost with relative velocity* $\mathbf{v}$ *and find the* 4×4 *matrix L such that*

$$\begin{pmatrix} t' \\ \mathbf{x}' \end{pmatrix} = L \cdot \begin{pmatrix} t \\ \mathbf{x} \end{pmatrix}.$$

Consider the limit $c \to \infty$ with $c^{-1}x^0$ remaining finite and show that one recovers a Galilean transformation.

One may argue that electromagnetism enables us to do experiments which go beyond classical mechanics and which allow us to distinguish certain inertial reference frames from others. However, there are several objections to this argument. Firstly, by Galilean Invariance the speed of light should differ for classical inertial coordinates moving in different directions, but the famous Michelson-Morley experiment[3] failed to detect such differences. Secondly, classical mechanics describes the motion of rigid bodies whose internal forces and collision forces are predominantly electromagnetic. In a uniformly moving frame at velocity v, Maxwell's equations take a different form, allowing us to detect v. It would be very remarkable if in some miraculous way the effect of electromagnetism on rigid bodies should cancel out completely to ensure that the Galilean Relativity Principle is satisfied.

Einstein realised that the conflict of classical mechanics with electromagnetism can be resolved most elegantly by noting that the notions of absolute time and distance are not operationally justified and need modification. E.g., let $p_1, p_2 \in M$ be two events which are supposedly simultaneous, but separated by a large distance. To verify that they really are simultaneous, one would need to be able to communicate between the places where the events are taking place at arbitrarily high speeds. Certainly no electromagnetic communication channel could satisfy this, because electromagnetic signals travel at a finite speed of propagation.

In his Special theory of Relativity, Einstein took electromagnetism as more fundamental than the notions of absolute time and distance, and he declared the speed of light in vacuum to be a universal physical constant. This forced him to weaken the assumptions on the properties of time and space, leading to the new concept of spacetime, which we investigate next.

Spacetime in Special Relativity

Let x^μ be coordinates on M, where $\mu = 0, \ldots 3$ and $x^0 := ct$, and suppose that Newton's first law and Maxwell's equations are both valid in these coordinates. We call these coordinates (relativistic) *inertial coordinates*, or a (relativistic) *inertial frame*. (That such coordinates exist, to a high degree of accuracy, is known from experience.) Now let x'^μ be any new set of coordinates, obtained from x^μ by a Poincaré transformation. Then Maxwell's equations will also hold in the new coordinates, and so will many essential aspects of Newton's laws, because the change of coordinates is linear. In

[3] Albert Michelson won the 1907 Nobel Prize in physics "for his optical precision instruments and the spectroscopic and metrological investigations carried out with their aid".

particular, the straight paths of free particles will map to straight paths, which means that the new coordinates are also inertial coordinates. However, two sets of inertial coordinates may no longer agree on time differences between all events or on the question whether two events are simultaneous, let alone on the distance between simultaneous events, masses of particles or the strengths of forces. This means that Newtonian mechanics will need some modification.

In order to preserve the equations of electromagnetism and (much of) classical mechanics, Einstein proposed to remove absolute time and distance from Galileo's Relativity Principle:

Tenet 4.1 (Special Relativity Principle) *In all inertial coordinate systems, the physics of classical mechanics and of electromagnetism are described respectively by (an adaptation of) Newton's laws and Maxwell's equations. In addition, nothing moves faster than c, the speed of light in vacuum.*

Despite the absence of absolute time and distances, there are still some statements that all inertial coordinate systems agree on. To formulate these, we introduce a standard basis on $\mathbb{R}^4$, e_μ with $\mu = 0, \ldots, 3$, so we may write $x = \sum_{\mu=0}^{3} x^\mu e_\mu$. In addition we introduce a bilinear symmetric form on $\mathbb{R}^4$,

$$\eta(x, y) := -x^0 y^0 + (x^1 y^1 + x^2 y^2 + x^3 y^3), \tag{5}$$

which is called the *Lorentz inner product* (or rather, pseudo-inner product). Note that $\eta(x, x)$ is not always non-negative, but η is non-degenerate: when $\eta(x, y) = 0$ for all y, then $x = 0$. We say that $x, y \in \mathbb{R}^4$ are orthogonal when $\eta(x, y) = 0$.

We define the *spacetime interval* (in appropriate units) by

$$\sigma(p_1, p_2) := \eta(x(p_1) - x(p_2), x(p_1) - x(p_2)). \tag{6}$$

When a physical signal travels from event p_1 to event p_2, then

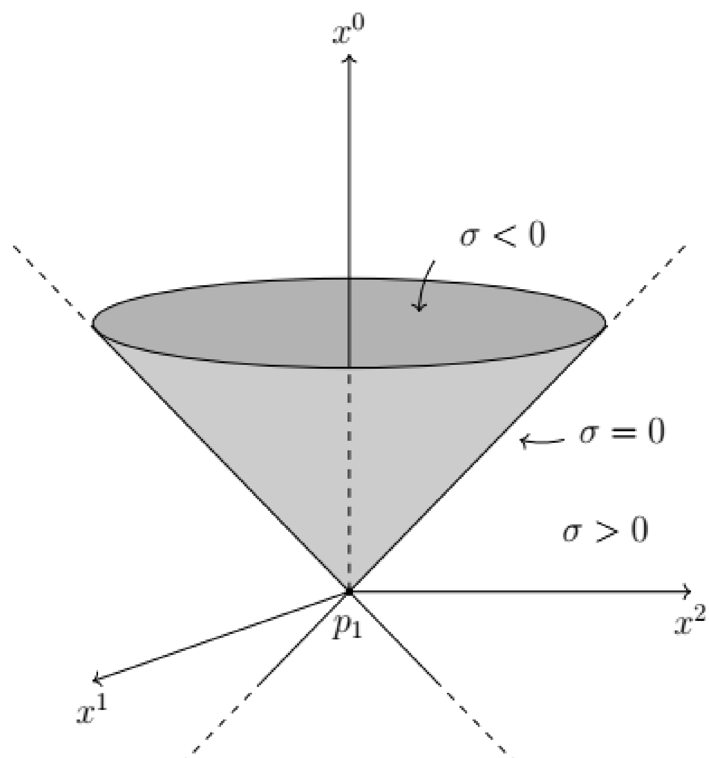
$$\begin{aligned} \sigma(p_1, p_2) = -(x^0(p_1) - x^0(p_2))^2 + \|\mathbf{x}(p_1) - \mathbf{x}(p_2)\|^2 &\leq 0, \\ x^0(p_2) - x^0(p_1) &\geq 0, \end{aligned}$$

because the signal cannot travel faster than the speed of light c. We have equality in the first line exactly when the signal travels at the speed of light and along a straight line.

Definition 4.2 *Two events* p_1, p_2 *are called* time-like*, resp.* light-like*, resp.* space-like related *when* $\sigma(p_1, p_2) < 0$*, resp.* $\sigma(p_1, p_2) = 0$*, resp.* $\sigma(p_1, p_2) > 0$*. Two events are called* causally related *when they are not space-like related* $(\sigma(p_1, p_2) \leq 0)$*.*

When p_1, p_2 *are causally related,* p_1 *lies* to the future*, resp.* past *of* p_2 *when* $x^0(p_1) \geq x^0(p_2)$*, resp.* $x^0(p_1) \leq x^0(p_2)$*.*

Note that when p_1 and p_2 are causally related and $x^0(p_1) = x^0(p_2)$, then $p_1 = p_2$.



Tenet 4.3 (Causal Structure of Spacetime) *In Special Relativity, for any events* $p_1, p_2, p_1', p_2' \in M$*, all inertial coordinate systems agree on*

1. *whether* p_1 *and* p_2 *are time-like, light-like or space-like related,*

2. *whether* p_1 *lies to the future or past of* p_2*, when* p_1, p_2 *are causally related, and*

3. *the value of* $\frac{\sigma(p_1,p_2)}{\sigma(p_1',p_2')}$*, when the denominator is non-zero.*

The second statement means that all inertial coordinate systems agree on the time-ordering between any two causally related events. In the last statement, the denominator is once again used to fix units. Note that in Special Relativity it suffices to fix time units only, because corresponding spatial units are then obtained by multiplication with the universal constant c.

Exercise 4.4 *Show that Poincaré transformations preserve the spacetime interval σ and the causal structure of Tenet 4.3.*

Note that we have not defined simultaneity between events. In Special Relativity such a notion can only be defined in a coordinate dependent way.

One might worry that the choice of an inertial coordinate system x^μ is already more than can be justified in an operational way. E.g., if an observer remains at position $\mathbf{x} = 0$, how can he possibly ascribe coordinates to events with $\mathbf{x} \neq 0$, or synchronise watches at different places? The answers to such questions were elaborated by Einstein in an entirely operational way, using only light signals and their reflections. This means that inertial coordinates can be constructed by a procedure that is independent of the observer, and for any two observers the resulting sets of inertial coordinates are indeed related by a Poincaré transformation. For the details of these procedures we refer to interested reader to most text books on Special Relativity.

Mathematics of Minkowski Spacetime

Using inertial coordinates x^μ we identify M with $\mathbb{R}^4$ and the structure of spacetime is then entirely encoded in η. The pair $(\mathbb{R}^4, \eta)$ is called *Minkowski space* (or rather: *Minkowski spacetime*). We will now formulate the structure of M and of the Poincaré transformations systematically in this four-dimensional formulation, using the coordinates x^μ. Most of the results are formulated in terms of exercises.

The terminology for causal relations on spacetime M can also be used on $\mathbb{R}^4$, using the identification via inertial coordinates. E.g. we say that a vector $x \in \mathbb{R}^4$ is time-like when $\eta(x, x) < 0$. (This is equivalent to saying that the event in M corresponding to the coordinates x is time-like related to the event corresponding to the coordinates $0 \in \mathbb{R}^4$.) Similarly, x is future pointing when it is causal (i.e. $\eta(x, x) \leq 0$) and $x^0 \geq 0$. In addition we call the set of all light-light vectors the *light cone* and the set of all causal vectors the *causal cone*. These cones can be split up into the *forward* and *backward* cones (each containing the vector 0).

Exercise 5.1 *Give a sketch of $\mathbb{R}^4$, indicating the sets of space-like vectors and time-like vectors. Also indicate the forward and backward light cones and causal cones.*

Exercise 5.2 *Recall the standard basis e_μ of $\mathbb{R}^4$, such that $x = \sum_{\mu=0}^{3} x^\mu e_\mu$. Show that*

$$\eta(e_\mu, e_\nu) = \begin{cases} -1 & \mu = \nu = 0 \\ 1 & \mu = \nu \neq 0 \\ 0 & \mu \neq \nu \end{cases} .$$

Exercise 5.3 *Let $x \in \mathbb{R}^4$ be a time-like vector. Show that all vectors orthogonal to x (w.r.t. the Lorentz inner product η) are space-like or zero.*

Exercise 5.4 *Let $x, y \in \mathbb{R}^4$ be two future pointing causal vectors. Show that $\eta(x, y) \leq 0$, that $z := x + y$ is a future pointing causal vector and that $\eta(x, y) = 0$ if and only if z is light-like if and only if x and y are light-like and parallel to each other.*

Exercise 5.5 *Find all vectors in $\mathbb{R}^4$ which are orthogonal to themselves (w.r.t. the Lorentz inner product η).*

We will now focus on Poincaré transformations.

Exercise 5.6 *Show that any Poincaré transformation is of the form*

$$x'^\mu(p) = L \cdot x^\mu(p) + x_0'^\mu,$$

for a unique $x_0' \in \mathbb{R}^4$ and a unique 4×4-matrix L such that

$$\eta(L \cdot x, L \cdot y) = \eta(x, y)$$

for all $x, y \in \mathbb{R}^4$. Express L in terms of A and $\mathbf{v}$ of Ex. 3.2.

Exercise 5.7 *Show that the Poincaré transformations form a group.*

Exercise 5.8 *Show that two inertial coordinate systems which employ the same units are related by a Poincaré transformation.*

It is sometimes helpful to also introduce the standard Euclidean inner product $\langle x, y\rangle = \sum_{\mu=0}^{3} x^\mu y^\mu$. Any elements $x, y \in \mathbb{R}^4$ then have

$$\eta(x, y) = \langle x, \eta \cdot y\rangle$$

for the diagonal matrix

$$\eta := \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{7}$$

Note, however, that the standard basis or inner product have no special physical significance. (They are not invariant under Poincaré transformations.)

The 4×4-matrices satisfying $\eta(L \cdot x, L \cdot y) = \eta(x, y)$ form a (Lie-)group $\mathcal{L}$, called the *Lorentz group*. Equivalently, it consists of all matrices satisfying $L^T \cdot \eta \cdot L = \eta$. Any element $L \in \mathcal{L}$ has $\det L = \pm 1$ and, expressed in the basis e_μ, the entry $L_{00} \neq 0$, because

$$-L_{00}^2 + \sum_{i=1}^{3} L_{i0}^2 = \langle L \cdot e_0, \eta \cdot L \cdot e_0\rangle = \langle e_0, \eta \cdot e_0\rangle = -1.$$

We may therefore decompose $\mathcal{L}$ into the four disjoint subsets (which are not necessarily subgroups)

$$\begin{aligned} \mathcal{L}_+^\uparrow &= \{L \in \mathcal{L} |\ \det L = +1,\ L_{00} > 0\}, \\ \mathcal{L}_-^\uparrow &= \{L \in \mathcal{L} |\ \det L = -1,\ L_{00} > 0\}, \\ \mathcal{L}_+^\downarrow &= \{L \in \mathcal{L} |\ \det L = +1,\ L_{00} < 0\}, \\ \mathcal{L}_-^\downarrow &= \{L \in \mathcal{L} |\ \det L = -1,\ L_{00} < 0\}. \end{aligned}$$

One may show that all these subsets are connected. We can go from any of these subsets to any other by multiplication with one of the following special elements:

- parity operation ("spatial reflection"): $L = -\eta$,
- time reversal operation: $L = \eta$,
- spacetime reflection: $L = -I$.

Exercise 5.9 *Show that $L \in \mathcal{L}$ defines a Lorentz transformation (which preserves the causal structure of Tenet 4.3) if and only if L is in the* orthochronous Lorentz group

$$\mathcal{L}^{\uparrow} := \mathcal{L}_{+}^{\uparrow} \cup \mathcal{L}_{-}^{\uparrow}.$$

Other subgroups of $\mathcal{L}$ are the *proper Lorentz group* $\mathcal{L}_{+} = \{L \in \mathcal{L} | \det L = +1\}$ and the *proper, orthochronous Lorentz group* $\mathcal{L}_{+}^{\uparrow}$. (The latter would have been the group of interest if we had also introduced an orientation on Minkowski spacetime, in addition to the time orientation.)

Exercise 5.10 *Let $x \in \mathbb{R}^4$ be a future pointing time-like vector. Show that there is a matrix $L \in \mathcal{L}$ such that $L \cdot x = e_0$, the standard basis vector which is future pointing and time-like.*

In the remaining part of this section we give some exercises about boosts, which illustrate the sometimes counter-intuitive consequences of the absence of absolute time.

Exercise 5.11 *Show that a boost with relative velocity $\mathbf{v} = ve_1$ is given by the matrix $L \in \mathcal{L}$ such that*

$$L = \begin{pmatrix} \cosh(\theta) & -\sinh(\theta) & 0 & 0 \\ -\sinh(\theta) & \cosh(\theta) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

and $\mathbf{v} = c\tanh(\theta)e_1$.

The parameter θ is called the *rapidity* of the boost. Another often used notation is $\beta = c^{-1}\mathbf{v}$, or $\beta = c^{-1}\|\mathbf{v}\|$.

Exercise 5.12 *Consider two boosts L_i, $i = 1, 2$, in the same direction with rapidities θ_i. Show that $L_3 := L_1 \cdot L_2$ is another boost with rapidity $\theta_3 = \theta_1 + \theta_2$. Show that the corresponding velocities $v_i = c\tanh(\theta_i)$ satisfy the relativistic velocity addition theorem*

$$v_3 = \frac{v_1 + v_2}{1 + \frac{v_1 v_2}{c^2}}.$$