Typeset in JHEP style September 11, 2014

The Principle of Relativity

Konsta Kurki

E-mail: konstakurki@gmail.com

www: theoreticalphysics.co

Abstract: The Principle of Relativity is discovered through a thought experiment. Lorentz boosts and basic features of Special Relativity are qualitatively discussed with the aid of Minkowski diagrams. Quantitative description of the Lorentz group is developed following Lie's approach. Particle action, energy and momentum are discussed and spacetime forms of the action principle and Noether's theorem developed.

Keywords: Principle of Relativity, Special Relativity, spacetime, Principle of Least Action, field theory, Noether's theorem

Contents

Acknowledgments

I thank Karri Kuivanen for proofreading the text and helping with figures, and Ville Kivioja for proofreading equations and suggesting improvements. Jarkko Saltiola took the picture of Figure 2.

1 More than one perspective

There exist a few more than one physicists, and further, many more other potential observers in the universe. A question arises: how does the personalities of Nature that she reveals to these different observers relate to each other? Obviously the perceptions of the observers themselves differ drastically. For example, an observer at the Equator finds the air that he breathes to be much hotter than someone at the Arctic Circle.

However, we know that much of these differencies can be reduced to different states of the environment. If the one at the Arctic Circle goes inside and heats his living room above, say the temperature of his body, the feeling of breathing no more differs so much from the one at the Equator.

1.1 Experiments around the universe

Let us build a couple of containers out of some very robust material. The walls of the containers must be thick and tight enough to not let anything go through them. This makes it difficult for the environment to affect the conditions in the containers.

We load each of the containers with an experimentalist, his assistant and equipment. The experimentalists of every container are given identical clocks, meter sticks and other reference units needed to make comparable measurements.

We then distribute the experimentalists in their containers around the universe: Imppu goes to sail on the ocean and Ilkka is dropped from a hot air balloon as is shown in Figures 1 and 2. Someone stays at the site of construction. One is put on a carousel, one on a spacestation and one is transported through time to the Stone Age. Some are loaded on trains, aeroplanes and spacecrafts, some sent to distant planets, stars and galaxies.

After the rumble of transport with spaceships and timemachines has abated, the experimentalists start experimenting in their containers. They conduct all kinds of experiments: swing pendulums, grow plants, burn stuff, and try whatever has come into their minds. Everybody carries out all the experiments. After they have been all performed, the containers are brought together and the experimentalists compare their results.

It is found that of all the containers that stayed in the vicinity of Earth, in the one at the construction site, in the one on the train and in all those that were moving with a uniform velocity relative to Earth's surface and were not horizontally tilted, the results of the experiments agree completely: pendulums swing with the same rate and so on. Imppu on her sailboat at the mercy of wind and other experimentalists in shaking, rotating or in any way accelerating containers find additional effects and do not get results in complete agreement.

In all these uniformly moving containers there is one particular direction, the up-down direction. The reason for that can be traced down to the fact that Earth

Figure 1. Imppu experimenting in her container on a sailboat. For the sake of clarity the walls of the container are removed from the picture.

pushes all these containers upwards. The results of some experiments in horizontally tilted containers differ because this direction appears different in those.

An even larger class of containers exists in which the results of the experiments agree perfectly. Ilkka, the one at the spacestation and every other experimentalist in a container that was left to fall freely with no rotational motion and with no external push from any direction find their results to be in complete agreement. These experimentalists do not observe any particular direction as those at Earth did.

The fact that experiments performed in all those freely falling nonrotating containers yield identical results suggests that on a fundamental level Nature is identical from the perspectives of the observers in those containers. This idea is called the Principle of Relativity.

We may now let the experimentalists pierce windows to their containers so that they can all observe the same system. The observations they make about the system are of course not identical: one sees the system from the left, one from the right. In the spirit of the Principle of Relativity, we expect Nature to obey the same laws in every container so the differences in the observations must lie in the description of the state and evolution of the system.

Figure 2. Ilkka falling from a hot air balloon. Most of the walls are removed from the picture.

1.2 The inertial frame

When an observer describes the evolution and state of a system, he needs to keep track of when things happen. This can be done by watching. The phenomenon that mediates sight, light, is very fast. According to perceptions from our everyday setting it may well be instantaneous: when a person moves her body in front of a mirror, she sees the move immediately despite the fact that light must travel to the mirror and back. So, if the assumption of instantaneous propagation of light can be made, we can identify the time of occurrence of an event with the time of its observation. We read this time from our clock and denote it by t.

The other thing to note is where things happen. This can be done by measuring the distance to the event in the back-front, left-right and down-up directions with our meter stick. The numbers obtained this way are called the Cartesian $\mathbf{x}^1$, $\mathbf{x}^2$ and $\mathbf{x}^3$ coordinates.

The Cartesian coordinates are not necessarily possible to determine directly, but we can for example record the direction where the considered event appears to happen and the apparent size, luminosity or other property of the event which we know depends on the distance to it. The Cartesian coordinates can be reasoned geometrically from these perceptions.

This way of determining the locations and times of occurrence of events is called the reference frame of the observer. If the observer is in a nonrotating, freely falling motion he is called an inertial observer and his reference frame an inertial frame. Usually the terms inertial frame and inertial observer are used interchangeably.

The time coordinate t and space coordinates $(\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3)$ are often advantageous to consider together as four coordinates $(t, \mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3) \doteq (x^0, x^1, x^2, x^3)$ of space and time or, as we often say, those of spacetime.

2 The group of coordinate transformations

To see how the description of the state and evolution of the system from the perspectives of different inertial observers relate to each other, we need to find the law of transformation that changes the spacetime coordinates between them. Mathematically the transformations form a group.

From the point of view of the observer transforming him looks exactly the same as transforming the system in the opposite way. Transforming the observer is called a passive transformation and transforming the system an active one. The difference is only in the direction in which the transformation is done; an active transformation is the inverse of the corresponding passive one. We speak about transforming the observer or the system depending on which is most clarifying, but in the mathematical formulation we adopt the active point of view since it makes some sign conventions simpler.

Often we write the coordinates by one component $x^\bullet$ at a time. Round colored symbols $\bullet$, ${\color{red}\bullet}$, ${\color{cyan}\bullet}$, ${\color{green}\bullet}$, ${\color{orange}\bullet}$ and ${\color{magenta}\bullet}$ are used to denote indices labeling the components of the coordinates and other quantitities we will meet.

Unless otherwise stated, we use the following notational conventions. Indices on objects denoted by symbols in italics take the values 0, 1, 2 and 3. On objects written in bold indices take only the values 1, 2 and 3. The whole set of the componets of a quantity is denoted by the symbol of the quantity without indices, for example $x = (x^0, x^1, x^2, x^3)$. The summation convention is used, which means that when the same index occurs twice in a term then a summation is performed over that index, for example $V_\bullet V^\bullet = V_0V^0 + V_1V^1 + V_2V^2 + V_3V^3$ and $\mathbf{x}^\bullet\mathbf{x}^\bullet = (\mathbf{x}^1)^2 + (\mathbf{x}^2)^2 + (\mathbf{x}^3)^2$.

2.1 Translations

Different inertial observers are located in different places in space and may be doing their observations at different times. Translation means changing the spatial location and transporting in time without rotating or changing velocity. Such an active transformation is evaluated simply by adding the changes $\Delta x^\bullet$ of the coordinates to

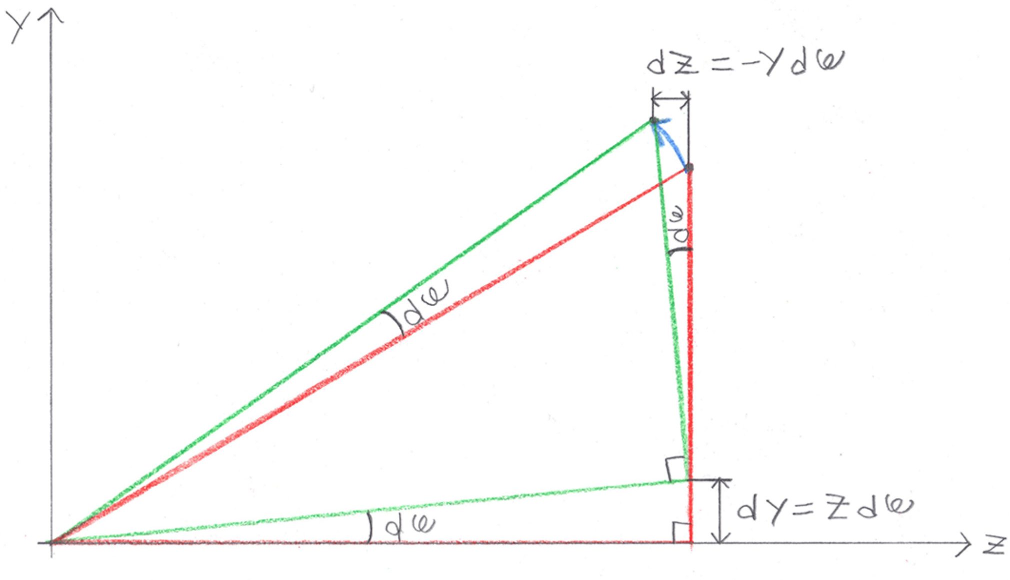

Figure 3. A small rotation.

the old ones:

$$x^\bullet \to x^\bullet + \Delta x^\bullet. \tag{2.1}$$

The order of successive translations clearly does not matter. It is said that the translations commute.

2.2 Rotations

Inertial observers may also be tilted with respect to each other. Rotation means changing the tilting angles. For simplicity only rotations around the origin are called pure rotations; if the system is not located at the origin, also its location changes in the rotation.

Rotations keep the spatial distance from the origin invariant. By pythagorean theorem the square of this distance is equal to $\mathbf{x}^\bullet\mathbf{x}^\bullet$. Time coordinate t remains untouched. It is also clear that space and spacetime volume elements $\mathrm{d}^3\mathbf{x} \doteq \prod_\bullet \mathrm{d}\mathbf{x}^\bullet$ and $\mathrm{d}^4x \doteq \mathrm{d}t\,\mathrm{d}^3\mathbf{x}$ are invariant.

Let us focus on the rotation in the $\mathbf{x}^2$–$\mathbf{x}^3$ plane in which the $\mathbf{x}^1$ coordinate does not change. We suppress the $\mathbf{x}^1$ coordinate and write $\mathbf{x} = (\mathbf{x}^2, \mathbf{x}^3)$, $y = \mathbf{x}^2$ and $z = \mathbf{x}^3$ for compactness. If we rotate only by a small amount $\mathrm{d}\omega$, then the change in y is the product of y and $\mathrm{d}\omega$ as we can see from Figure 3. The same Figure tells us that the change in z is equal to $-y\mathrm{d}\omega$. We have

$$y \to y + z\mathrm{d}\omega \qquad \text{and} \qquad z \to z - y\mathrm{d}\omega. \tag{2.2}$$

For this rotation it holds that

$$y^2 + z^2 \to (y + z\mathrm{d}\omega)^2 + (z - y\mathrm{d}\omega)^2 \tag{2.3}$$
$$= y^2 + 2yz\mathrm{d}\omega + z^2\mathrm{d}\omega^2 + z^2 - 2zy\mathrm{d}\omega + y^2\mathrm{d}\omega^2 \tag{2.4}$$
$$= y^2 + z^2 + (y^2 + z^2)\mathrm{d}\omega^2. \tag{2.5}$$

Since the rotation is small, $\mathrm{d}\omega^2$ is insignificant and can be dropped. This confirms that $\mathbf{x}^\bullet\mathbf{x}^\bullet$ is invariant.

The rotation can be written with matrices as

$$\mathbf{x} \to \mathbf{x} + \mathrm{d}\omega \begin{pmatrix} & + \\ - & \end{pmatrix} \mathbf{x} \tag{2.6}$$

or

$$\mathbf{x} \to \Lambda(\mathrm{d}\omega)\mathbf{x}, \qquad \Lambda(\mathrm{d}\omega) = 1 + \mathrm{d}\omega R, \qquad R = \begin{pmatrix} & + \\ - & \end{pmatrix}. \tag{2.7}$$

The matrix R is called the generator of rotation. Finite rotation can be acquired by doing many small rotations in succession. For finite ω we have

$$\Lambda(\omega) = \Lambda(\omega/N)^N = \left(1 + \frac{\omega R}{N}\right)^N \tag{2.8}$$

where N is large.

Using an elementary result in calculus,

$$\lim_{N\to\infty} \left(1 + \frac{x}{N}\right)^N = e^x, \tag{2.9}$$

we can write the finite rotation as

$$\Lambda(\omega) = e^{\omega R}. \tag{2.10}$$

Expanding the exponential as a power series we have

$$\Lambda(\omega) = \sum_{n=0}^{\infty} \frac{1}{n!}(\omega R)^n = 1\left(1 - \frac{\omega^2}{2!} + \frac{\omega^4}{4!} - \cdots\right) + R\left(\omega - \frac{\omega^3}{3!} + \frac{\omega^5}{5!} - \cdots\right). \tag{2.11}$$

The power series in the first and second brackets are the ones of cosine and sine, so we have

$$\Lambda(\omega) = \begin{pmatrix} \cos\omega & \sin\omega \\ -\sin\omega & \cos\omega \end{pmatrix}. \tag{2.12}$$

Graduationg to four dimensions and writing $x = (t, \mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3)$, the generators of rotations in the $\mathbf{x}^1$–$\mathbf{x}^2$, $\mathbf{x}^1$–$\mathbf{x}^3$ and $\mathbf{x}^2$–$\mathbf{x}^3$ planes are obviously

$$\mathbf{J}^{12} = \begin{pmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & & + & \\ \cdot & - & & \\ \cdot & & & \end{pmatrix}, \quad \mathbf{J}^{13} = \begin{pmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & & & + \\ \cdot & & & \\ \cdot & - & & \end{pmatrix} \quad \text{and} \quad \mathbf{J}^{23} = \begin{pmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & & & \\ \cdot & & & + \\ \cdot & & - & \end{pmatrix}. \tag{2.13}$$

There is no rotation in $\mathbf{x}^\bullet$–$\mathbf{x}^\bullet$ plane, so $\mathbf{J}^{\bullet\bullet} = 0$ (no sum over $\bullet$ here). Rotation in $\mathbf{x}^\bullet$–$\mathbf{x}^{\color{red}\bullet}$ plane is clearly the opposite of the one in $\mathbf{x}^{\color{red}\bullet}$–$\mathbf{x}^\bullet$, so for all $\bullet$ and ${\color{red}\bullet}$ it holds that $\mathbf{J}^{\bullet{\color{red}\bullet}} = -\mathbf{J}^{{\color{red}\bullet}\bullet}$. For the same reasons the corresponding rotation angles $\omega_{\bullet{\color{red}\bullet}}$ are antisymmetric, $\omega_{\bullet{\color{red}\bullet}} = -\omega_{{\color{red}\bullet}\bullet}$. As with $\mathbf{J}$, here indices on ω take only the values 0, 2 and 3.

Small rotation in an arbitrary direction can be carried out by rotating successively around the $\mathbf{x}^\bullet$–$\mathbf{x}^{\color{red}\bullet}$ planes. For instance, rotating first in the $\mathbf{x}^1$–$\mathbf{x}^2$ and then in the $\mathbf{x}^2$–$\mathbf{x}^3$ plane results in

$$
\begin{aligned}
\Lambda &= (1 + \mathrm{d}\omega_{23}\mathbf{J}^{23})(1 + \mathrm{d}\omega_{12}\mathbf{J}^{12}) && (2.14)\\
&= 1 + \mathrm{d}\omega_{23}\mathbf{J}^{23} + \mathrm{d}\omega_{12}\mathbf{J}^{12} + \mathrm{d}\omega_{23}\mathrm{d}\omega_{12}\mathbf{J}^{23}\mathbf{J}^{12}. && (2.15)
\end{aligned}
$$

Since the angles $\omega_{\bullet{\color{red}\bullet}}$ are small, their product is insignificant and the last term can be dropped. Carrying out the rotations in opposite order would result in exactly the same rotation, so for small rotations their order does not matter.

A general small rotation can be written as

$$
\begin{aligned}
\Lambda &= 1 + \mathrm{d}\omega_{12}\mathbf{J}^{12} + \mathrm{d}\omega_{13}\mathbf{J}^{13} + \mathrm{d}\omega_{23}\mathbf{J}^{23} && (2.16)\\
&= 1 + \frac{1}{2}(\mathrm{d}\omega_{12}\mathbf{J}^{12} + \mathrm{d}\omega_{13}\mathbf{J}^{13} + \mathrm{d}\omega_{23}\mathbf{J}^{23} + \mathrm{d}\omega_{21}\mathbf{J}^{21} + \mathrm{d}\omega_{31}\mathbf{J}^{31} + \mathrm{d}\omega_{32}\mathbf{J}^{32}) && (2.17)\\
&= 1 + \frac{1}{2}\mathrm{d}\omega_{\bullet{\color{red}\bullet}}\mathbf{J}^{\bullet{\color{red}\bullet}} && (2.18)
\end{aligned}
$$

and a finite one as

$$
\Lambda = \exp\left(\frac{1}{2}\omega_{\bullet{\color{red}\bullet}}\mathbf{J}^{\bullet{\color{red}\bullet}}\right). \tag{2.19}
$$

If a finite rotation is not performed around some coordinate axis but around some other direction, then $\omega_{\bullet{\color{red}\bullet}}$'s cannot straightforwardly interpreted as definite rotation angles.

Note also that if $AB \neq BA$, then in general

$$
e^A e^B \neq e^{A+B}. \tag{2.20}
$$

Finite rotations do not commute unless they are performed in the same plane.

2.3 Boosts

There is a third way in which freely falling nonrotating observers may differ from each other. When several observers jump off a roof at different times, they fall at different velocities. However, the difference in their velocities does not change. In the same fashion, when a spaceship is about to dock with a space station, it moves with a constant velocity with respect to the station unless it is using its steering engines. In general, freely falling observers move at constant velocity with respect to each other.

Transformations that change the velocity are called boosts. For simplicity, only such transformations that do not rotate the system and for which $(0,0,0,0) \to (0,0,0,0)$ are referred to as pure boosts.

If all the velocities in the problem including the boost velocity $\mathbf{v}$ are so small that the propagation of light can be considered instantaneous, then both observers, the boosted and the unboosted can identify the time of occurrence with the time of the observation. The observers agree about their time coordinates and the boost is merely a translation that depends linearly on time:

$$t \to t, \qquad \mathbf{x}^\bullet \to \mathbf{x}^\bullet + \mathbf{v}^\bullet t. \tag{2.21}$$

These transformations are called Galilean boosts. Like translations, the Galilean boosts commute.

Since the Galilean boosts leave the t coordinate untouched, any velocity $\dot{\mathbf{y}}$ get simply transformed by

$$\dot{\mathbf{y}}^\bullet = \frac{\mathrm{d}\mathbf{y}^\bullet}{\mathrm{d}t} \to \frac{\mathrm{d}(\mathbf{y}^\bullet + \mathbf{v}^\bullet t)}{\mathrm{d}t} = \frac{\mathrm{d}\mathbf{y}^\bullet}{\mathrm{d}t} + \mathbf{v}^\bullet = \dot{\mathbf{y}}^\bullet + \mathbf{v}^\bullet. \tag{2.22}$$

3 The speed of light

Careful experiments show that contrary to our guess based on everyday life, the propagation of light is rapid but not instantaneous. The speed by which light travels is finite, and is usually denoted by c. Mechanics that take into account the finity of the speed of light is called relativistic.

It is a striking experimental fact that whenever the speed of light is measured, the result is always the same provided there is no dense matter in the way retarding the propagation. For light our atmosphere is very close to vacuum. The speed of light does not depend on the velocity of the observer nor the velocity of the light source and is roughly a million times the speed of sound in the air surrounding us.

It has been thought that light propagates in so-called aether and its speed is c relative to it. The observer would measure always the same speed because he somehow drags the aether along him. However, despite efforts this aether has never been observed.

A simpler account is to propose that the speed of light is an absolute quantity independent of the motion of the observer, aether or anything else. This path leads to a simple, beautiful and most importantly, correct picture of Nature. We will not consider the historical aether hypothesis further.

Since the speed is constant, we can still find the time coordinate of an event by sight. We just have to subtract from the moment of arrival of the light ray its time of flight.

The units of time and length can be chosen in any way wished. At this point it is advantageous to choose them so that the speed of light is equal to one, i.e. in a unity time light travels a unity length. Everyday velocities are thus very close to zero. This convention will make equations look very much cleaner. In other kind of systems of units there would be factors of c in the equations, which can be located by analyzing the dimensionality of the equations.

As we saw, in the Galilean boosts the velocities are transformed by just addition. This implies that the speed of light changes, so it is evident that Galilean boosts need some modification. The time coordinate must also get transformed.

3.1 Einstein's train

A train moves on rails. Suddenly a light flashes in the middle of one of the cars. After a tiny while the light hits the back and front doors of the car. We denote these events by A and B.

Imppu on the train analyzes the situation. The distances from the source to the front and back doors are equal, and since light travels with the same speed in both directions, both trips take the same time and the arrival at the front door gets the same value for time coordinate as the arrival to the back door. Imppu concludes that A and B are simultaneous.

llkka standing beside the track analyzes the same situation. From his perspective both doors of the car are moving forward. When the flash of light arrives at the front door the door has moved away from the source and the distance that the light pulse has travelled is more than half of the length of the car. On the other hand, at the moment of the arrival to the back the back door has moved towards the source, and the distance travelled by the light pulse is smaller than half of the length of the car. Since light again moves with the same speed in both directions, it reaches the back before the front. Ilkka concludes that A and B are not simultaneous.

It is illustrative to draw so-called Minkowski diagrams about the situation. In a Minkowski diagram time increases upwards and light rays are drawn in half right angles. Light rays coming to and escaping from a fixed point form a so-called light cone.

Figure 4 shows the Minkowski diagrams of the flash from the perspectives of Imppu and Ilkka. Only one space dimension which is denoted by z and $\tilde{z}$ is drawn. Light rays are drawn in blue and yellow and the lines of the moving front and back of the car in green and red. The light rays reflected from the doors, and the second reflections, the events C and D, are also drawn in the diagrams. We see that in Ilkka's frame the events of reflection form a square which in Imppu's frame turns into a rhombus with diagonals at the angles of light rays.

Generally, any shape is stretched in the direction of the light ray going to right and shrinked in the other. The area and so the spacetime volume element d^4x stay untouched, because the inverse transformation is only a boost in the opposite direction and changes the area with the same coefficient. Transforming and then doing its inverse is doing nothing, so the coefficient must be one.

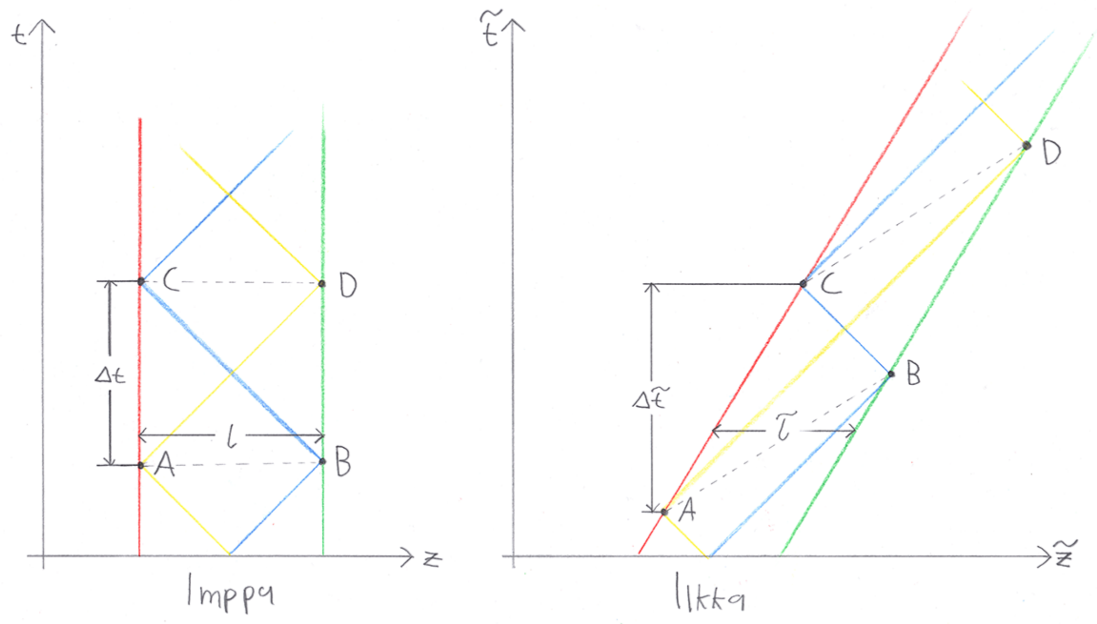

Figure 4. Minkowski diagrams of Einstein's train.

Generally, any shape is stretched in the direction of the light ray going to right and shrinked in the other. The inverse boost is just a boost of the same amount in the opposite direction and changes the areas in the diagram by the same coefficient. Since transfoming and then doing its inverse is doing nothing, coefficient must be one. Therefore the area and spacetime volume element d^4x stay untouched in the boost.

That kind of boosts are called Lorentz boosts. When the stretch is nearly infinite, every trajectory start to look like a light ray. This corresponds to a boost with a velocity near 1. It does not make sense to talk about boost velocities of one or higher.

3.2 Length contraction and time dilation

Consider the time interval between the events A and C. It can be easily seen from Figure 4 that the interval is longer in the frame of Ilkka. We have $\Delta t < \Delta\tilde{t}$. When the velocity of the train approaches one, the time interval measured by Ilkka increases to infinity.

This is an example of a general phenomenon of time dilation. Every physical process slows down at high velocities, freezing totally at the speed of light.

A related phenomenon occurs with distance. From the same Figure we see that the distance between the doors of the car is smaller in the boosted frame. It holds that $l > \tilde{l}$. When accelerated into high velocities, any physical body gets shrinked in the direction of the velocity, its length closing to zero when the velocity closes one. That is called length contraction.

3.3 Causality

Take two events and draw a straight line between them. If the line in the Minkowski diagram is exactly in the direction of a light ray, then no matter what kind of boost is performed, the line stays in the direction of light, and the sign of Δt, the difference of the time coordinates of the events, stays unchanged. The separation Δx between such events is called lightlike.

If the direction of the line is closer to the vertical than to the horizontal, then it stays as such in every boost. The sign of Δt stays the same. With a suitable boost the line can be made exactly vertical. This corresponds to boosting to a frame in which the events happen in the same spatial location. Separations between such events are called timelike.

Lastly, if the direction of the line is closer to the horizontal than to the vertical, it again stays as such in every boost. The line cannot be made vertical by boosting, so there is no frame in which the events happen in the same spatial location. The coordinate difference in time Δt can be made to vanish or change sign, so there is a frame where they happen at the same time and frames where they happen in different temporal order than in the original frame. That kind of separations are said to be spacelike.

For events with timelike and lightlike separations, the difference Δt in time coordinate preserves its sign in every boost. The temporal ordering of the events is independent of the observer. There is no problem of the events being in causal connection with each other.

For events separated in spacelike fashion, the sign of the time coordinate and thus their temporal order depends on the observer. Since it is pretty absurd that from the perspectives of half of the observers the effect could occur before the cause, we conclude that events with spacelike separation cannot be in causal connection with each others. The region of spacetime that can be causally connected to an event lies in its light cone as can be seen from the Minkowski diagram.

The line that the life of a body draws in spacetime is called the worldline of the body. Motion with a speed faster than one results in a worldline closer to vertical than horizontal in the Minkowski diagram. Successive events picked up from that kind of worldline have a spacelike separation, and thus cannot be in causal connection. Therefore information cannot propagate with speed greater than one.

This fact implies that if a body is pushed, then the push propagates to the other side of the body maximally at speed one, so it takes some time for the other side to react to the push and the body is not rigid. Strictly rigid bodies cannot exist in Nature.

3.4 Lorentz boosts

Let us focus on boosting a small amount $\mathrm{d}\omega$ in the $\mathbf{x}^3$ direction, suppress the $\mathbf{x}^1$ and $\mathbf{x}^2$ coordinates and write $x = (t, \mathbf{x}^3)$ and $z = \mathbf{x}^3$.

The boost appears in the most simple form if the left and right going light rays starting from the origin are used as the coordinate axes. Let's call these coordinates the c_- and c_+ lightcone coordinates. Note that they are not inertial coordinates of any observer. In terms of z and t these are

$$c_- = \frac{1}{2}(z - t) \qquad \text{and} \qquad c_+ = \frac{1}{2}(z + t). \tag{3.1}$$

If the area is to be unchanged, the stretch in the c_- direction must be the opposite but equal in size than in the c_+ direction, so $\mathrm{d}c_- = -c_-\mathrm{d}\omega$ and $\mathrm{d}c_+ = c_+\mathrm{d}\omega$.

For z and t this implies that $\mathrm{d}t = z\mathrm{d}\omega$ and $\mathrm{d}z = t\mathrm{d}\omega$ or

$$t \to t + z\mathrm{d}\omega \qquad \text{and} \qquad z \to z + t\mathrm{d}\omega. \tag{3.2}$$

We see that when $t \gg z$, which corresponds to locations of slowly moving bodies, $z\mathrm{d}\omega$ is insignificant and the small Lorentz boost reduces to a Galilean boost with $\mathbf{v} = \mathrm{d}\omega$.

The small Lorentz boost can be written with matrices as

$$x \to \Lambda x, \qquad \Lambda = 1 + \mathrm{d}\omega B, \qquad B = \begin{pmatrix} & + \\ + & \end{pmatrix}. \tag{3.3}$$

The only difference to rotation is that both entries in the generator are positive. This causes the boost to keep the difference rather than the sum of the squares of the coordinates invariant:

$$t^2 - z^2 \to (t + z\mathrm{d}\omega)^2 - (z + t\mathrm{d}\omega)^2 \tag{3.4}$$
$$= t^2 + 2tz\mathrm{d}\omega + z^2\mathrm{d}\omega^2 - z^2 - 2tz\mathrm{d}\omega - t^2\mathrm{d}\omega^2 \tag{3.5}$$
$$= t^2 - z^2. \tag{3.6}$$

We proceed as with rotations: a finite boost is

$$\Lambda = e^{\omega B} = \sum_{n=0}^{\infty} \frac{1}{n!}(\omega B)^n = 1 \sum_{n\,\text{even}} \frac{1}{n!}\omega^n + B \sum_{n\,\text{odd}} \frac{1}{n!}\omega^n. \tag{3.7}$$

The power series

$$\sum_{n\,\text{even}} \frac{1}{n!}\omega^n = \frac{1}{2}\sum_{n=0}^{\infty} \frac{1}{n!}(\omega^n + (-\omega)^n) = \frac{1}{2}(e^{\omega} + e^{-\omega}) \doteq \cosh\omega \tag{3.8}$$

and

$$\sum_{n\,\text{odd}} \frac{1}{n!}\omega^n = \frac{1}{2}\sum_{n=0}^{\infty} \frac{1}{n!}(\omega^n - (-\omega)^n) = \frac{1}{2}(e^{\omega} - e^{-\omega}) \doteq \sinh\omega \tag{3.9}$$

are called the hyperbolic cosine and sine due to their similarities with ordinary trigonometric functions. The finite boost becomes

$$\Lambda = \begin{pmatrix} \cosh\omega & \sinh\omega \\ \sinh\omega & \cosh\omega \end{pmatrix} \tag{3.10}$$

or

$$t \to t\cosh\omega + z\sinh\omega \qquad \text{and} \qquad z \to t\sinh\omega + z\cosh\omega. \tag{3.11}$$

Velocity of a body at rest in the spatial origin of the original frame transforms by

$$\frac{z=0}{t} \to \frac{t\sinh\omega}{t\cosh\omega} = \tanh\omega \doteq \mathbf{v}. \tag{3.12}$$

The velocity $\mathbf{v}$ is the boost velocity.

Hyperbolic functions satisfy many identities similar to the ones of ordinary trigonometric functions. One of them is

$$\cosh^2\omega - \sinh^2\omega \equiv 1. \tag{3.13}$$

Using it and the definition of $\mathbf{v}$ we get

$$\cosh\omega = \frac{1}{\sqrt{1-\mathbf{v}^2}} \doteq \gamma(\mathbf{v}) \qquad \text{and} \qquad \sinh\omega = \frac{\mathbf{v}}{\sqrt{1-\mathbf{v}^2}} = \gamma(\mathbf{v})\mathbf{v}. \tag{3.14}$$

When the boost velocity $\mathbf{v}$ goes from zero to one, the function $\gamma(\mathbf{v})$ rises from one to infinity. If $\mathbf{v}$ is small, then $\gamma(\mathbf{v}) \approx 1 + \frac{1}{2}\mathbf{v}^2$. In terms of the boost velocity and $\gamma(\mathbf{v})$ the equation (3.11) becomes

$$t \to (t + \mathbf{v}z)\gamma(\mathbf{v}) \qquad \text{and} \qquad z \to (\mathbf{v}t + z)\gamma(\mathbf{v}). \tag{3.15}$$

Continuing the analogy, we think of the Lorentz boosts as kind of hyperbolic rotations in the t–$\mathbf{x}^\bullet$ planes. In four dimensions their generators are

$$J^{01} = \begin{pmatrix} \cdot & + & & \\ + & \cdot & \cdot & \cdot \\ & \cdot & \cdot & \cdot \\ & \cdot & \cdot & \cdot \end{pmatrix}, \quad J^{02} = \begin{pmatrix} \cdot & & + & \\ & \cdot & \cdot & \cdot \\ + & \cdot & \cdot & \cdot \\ & \cdot & \cdot & \cdot \end{pmatrix} \quad \text{and} \quad J^{03} = \begin{pmatrix} \cdot & & & + \\ & \cdot & \cdot & \cdot \\ & \cdot & \cdot & \cdot \\ + & \cdot & \cdot & \cdot \end{pmatrix}. \tag{3.16}$$

We set for the generators $J^{\bullet 0} = -J^{0\bullet}$ and for the boost parameters $\omega_{\bullet 0} = -\omega_{0\bullet}$. Small Lorentz boost in any direction can then be written as

$$\Lambda = 1 + \mathrm{d}\omega_{01}J^{01} + \mathrm{d}\omega_{02}J^{02} + \mathrm{d}\omega_{03}J^{03} \tag{3.17}$$

$$= 1 + \frac{1}{2}(\mathrm{d}\omega_{0\bullet}J^{0\bullet} + \mathrm{d}\omega_{\bullet 0}J^{\bullet 0}) \tag{3.18}$$

and a finite one as

$$\Lambda = \exp\left(\frac{1}{2}(\omega_{0\bullet}J^{0\bullet} + \omega_{\bullet 0}J^{\bullet 0})\right). \tag{3.19}$$

Finite Lorentz boosts do no commute unless they are performed along the same axis.

The group spanned by translations, rotations and Lorentz boosts is called the Poincaré group. In this unified notation an arbitrary Poincaré transformation is written as

$$x \to \Lambda x + \Delta x, \qquad \Lambda = \exp\left(\frac{1}{2}\omega_{\bullet{\color{red}\bullet}}J^{\bullet{\color{red}\bullet}}\right) \tag{3.20}$$

where $J^{\bullet{\color{red}\bullet}} = \mathbf{J}^{\bullet{\color{red}\bullet}}$ when $\bullet, {\color{red}\bullet} = 1, 2$ or 3.

4 Minkowski space

The spacetime we have studied here is called Minkowski space. Its three rotations, three boosts and three categories of directions have their roots in the subjective split of spacetime into space and time which is naturally performed by every observer. In ordinary four dimensional Euclidean space there would be only six ordinary rotations and no distinct categories of timelike, spacelike and lightlike directions.

4.1 Vectors

The coordinate difference $\Delta x = (\Delta x^0, \Delta x^1, \Delta x^2, \Delta x^3)$ of two events is in its essence an arrow of some length pointing to some direction in spacetime. We can add two such arrows by placing them one after the other without changing their directions. Mathematically this is accomplished by adding the numerical values of the coordinate differences. The arrow can also be multiplied by a real number, which changes the length of the arrow Mathematically the separation Δx is a vector.

The components $\Delta x^\bullet$ of the separation vector do not depend on the origin of the coordinate system, so in a translation the separation vector is invariant. In a Poincaré transformation it gets only rotated and boosted, $\Delta x \to \Lambda \Delta x$. The group spanned by rotations and Lorentz boosts is called the Lorentz group.

We call any object V which transforms in a Poincaré transformation like Δx, that is $V \to \Lambda V$, a vector. Every such object share the spirit of the separation as an arrow or some length pointing into some directions in spacetime. Vectors can be added to each other and multiplied by real numbers. Lorentz transformation is linear so a linear combination of vectors transforms like it should, that is, as a vector.

4.2 Norm and inner product

Rotations leave V^0 and $\mathbf{V}^\bullet\mathbf{V}^\bullet$, where $\mathbf{V} = (V^1, V^2, V^3)$, invariant and a Lorentz boost in the $\mathbf{x}^\bullet$ direction leave $(V^0)^2 - \mathbf{V}^\bullet\mathbf{V}^\bullet$ invariant. Therefore every Lorentz transformation as a product of boosts and rotations leave the quantity $(V^0)^2 - \mathbf{V}^\bullet \mathbf{V}^\bullet$, called the square of the norm, invariant.

The norm of Minkowski space differs from the one of Euclidean space by not being positive definite. In Minkowski space the square of the norm of a nonzero vector can be positive, zero or negative. From the Minkowski diagram we can easily see that for lightlike separation we have $\Delta t^2 = \Delta\mathbf{x}^\bullet \Delta\mathbf{x}^\bullet$, so the square of its norm, the so-called spacetime interval, is zero. For timelike and spacelike separations we have $\Delta t^2 > \Delta\mathbf{x}^\bullet \Delta\mathbf{x}^\bullet$ and $\Delta t^2 < \Delta\mathbf{x}^\bullet \Delta\mathbf{x}^\bullet$ so the square of the spacetime interval of events with spacelike separation is negative and the one of timelike is positive. We extend the concept of likeness to all vectors.

We introduce the notation $V_\bullet$ with lower index by defining $V_0 = V^0$ and $\mathbf{V}_\bullet = -\mathbf{V}^\bullet$. Then the invariant square of the norm can be simply written as $V_\bullet V^\bullet \doteq V^2$. This lowering of an index can be done in terms of the so-called metric

$$\eta_{\bullet{\color{red}\bullet}} = \begin{pmatrix} + & & & \\ & - & & \\ & & - & \\ & & & - \end{pmatrix} \tag{4.1}$$

by writing $V_\bullet = \eta_{\bullet{\color{red}\bullet}} V^{{\color{red}\bullet}}$ and $V^2 = \eta_{\bullet{\color{red}\bullet}} V^\bullet V^{{\color{red}\bullet}}$.

We also define the metric with upper indices by $V^\bullet = \eta^{\bullet{\color{red}\bullet}} V_{{\color{red}\bullet}}$. Since rising is the inverse operation of lowering, we must have

$$\eta^{\bullet{\color{red}\bullet}} \eta_{{\color{red}\bullet}{\color{cyan}\bullet}} \doteq \eta^{\bullet}{}_{{\color{cyan}\bullet}} = \begin{pmatrix} + & & & \\ & + & & \\ & & + & \\ & & & + \end{pmatrix}, \tag{4.2}$$

i.e, $\eta^\bullet{}_\bullet$ is the unit matrix. It happens to be that numerically $\eta_{\bullet{\color{red}\bullet}} = \eta^{\bullet{\color{red}\bullet}}$. We use η to lower and rise indices of any quantity, for example $\eta_{\bullet{\color{red}\bullet}} \Lambda^{{\color{red}\bullet}}{}_{{\color{cyan}\bullet}} \doteq \Lambda_{\bullet{\color{cyan}\bullet}}$.

With the metric the Lorentz group generators can be compactly written as

$$(J^{\bullet{\color{red}\bullet}})^{{\color{cyan}\bullet}}{}_{{\color{green}\bullet}} = \eta^{\bullet{\color{cyan}\bullet}} \eta^{{\color{red}\bullet}}{}_{{\color{green}\bullet}} - \eta^{{\color{red}\bullet}{\color{cyan}\bullet}} \eta^{\bullet}{}_{{\color{green}\bullet}} \tag{4.3}$$

as can be easily verified. Here ${}^{\bullet{\color{red}\bullet}}$ labels the generator and ${}^{{\color{cyan}\bullet}}{}_{{\color{green}\bullet}}$ its component.

Since the square of the norm is invariant, it looks like the transformation of the lower index undoes the transformation of the upper. This is indeed the case. The vector with lower index transforms by

$$V_\bullet = \eta_{\bullet{\color{red}\bullet}} V^{{\color{red}\bullet}} \to \eta_{\bullet{\color{red}\bullet}} \Lambda^{{\color{red}\bullet}}{}_{{\color{cyan}\bullet}} V^{{\color{cyan}\bullet}} = \eta_{\bullet{\color{red}\bullet}} \Lambda^{{\color{red}\bullet}}{}_{{\color{cyan}\bullet}} \eta^{{\color{cyan}\bullet}{\color{green}\bullet}} V_{{\color{green}\bullet}} = \Lambda_\bullet{}^{{\color{red}\bullet}} V_{{\color{red}\bullet}}. \tag{4.4}$$

If the Lorentz tranformation is small, then it can be written as $\Lambda = 1 + J$ with $J = \frac{1}{2} \mathrm{d}\omega_{\bullet{\color{red}\bullet}} J^{\bullet{\color{red}\bullet}}$. The corresponding transformation for lower index is

$$\Lambda_\bullet{}^{{\color{red}\bullet}} = \eta_{\bullet{\color{cyan}\bullet}} \Lambda^{{\color{cyan}\bullet}}{}_{{\color{green}\bullet}} \eta^{{\color{green}\bullet}{\color{red}\bullet}} = \eta_{\bullet{\color{cyan}\bullet}} (\eta^{{\color{cyan}\bullet}}{}_{{\color{green}\bullet}} + J^{{\color{cyan}\bullet}}{}_{{\color{green}\bullet}}) \eta^{{\color{green}\bullet}{\color{red}\bullet}} = \eta_\bullet{}^{{\color{red}\bullet}} + \eta_{\bullet{\color{cyan}\bullet}} J^{{\color{cyan}\bullet}}{}_{{\color{green}\bullet}} \eta^{{\color{green}\bullet}{\color{red}\bullet}} = \eta_\bullet{}^{{\color{red}\bullet}} + J_\bullet{}^{{\color{red}\bullet}}. \tag{4.5}$$

By using (4.3) it is straightforward to show that $J_{\bullet}{}^{\color{red}\bullet} = -J^{\color{red}\bullet}{}_{\bullet}$. Therefore

$$\Lambda^{\bullet}{}_{\color{red}\bullet}\Lambda_{\color{green}\bullet}{}^{\color{red}\bullet} = (\eta^{\bullet}_{\color{red}\bullet} + J^{\bullet}{}_{\color{red}\bullet})(\eta^{\color{red}\bullet}_{\color{green}\bullet} - J^{\color{red}\bullet}{}_{\color{green}\bullet}) = \eta^{\bullet}_{\color{green}\bullet} + J^{\bullet}{}_{\color{green}\bullet} - J^{\bullet}{}_{\color{green}\bullet} + J^{\bullet}{}_{\color{red}\bullet}J^{\color{red}\bullet}{}_{\color{green}\bullet} = \eta^{\bullet}_{\color{green}\bullet} \tag{4.6}$$

so $\Lambda_{\bullet}{}^{\color{red}\bullet}$ is the inverse of $\Lambda^{\color{red}\bullet}{}_{\bullet}$. Since finite Lorentz transformations are products of small ones, this holds also for finite transformations. This implies that the quantity $V_{\bullet}W^{\bullet}$, called the inner product of two vectors, transforms by

$$V_{\bullet}W^{\bullet} \to \Lambda_{\bullet}{}^{\color{red}\bullet}\Lambda^{\bullet}{}_{\color{cyan}\bullet}V_{\color{red}\bullet}W^{\color{cyan}\bullet} = \eta^{\color{red}\bullet}_{\color{cyan}\bullet}V_{\color{red}\bullet}W^{\color{cyan}\bullet} = V_{\bullet}W^{\bullet}. \tag{4.7}$$

The inner product remains invariant in Lorentz transformations.

4.3 Proper time

The personal time felt by an inertial observer is measured by his own clock and time coordinate t_o of his own inertial frame. Observer in his own frame is always located at the origin, so for the separation $\mathrm{d}x_o$ of two closely successive events in his life it holds that $\mathrm{d}\mathbf{x}^{\bullet}_{o} = 0$. Therefore in his own frame

$$\mathrm{d}x_{\bullet o}\mathrm{d}x^{\bullet}_{o} = (\mathrm{d}t_o)^2 \doteq \mathrm{d}\tau^2 \tag{4.8}$$

and

$$t_o = \int \mathrm{d}t_o = \int \mathrm{d}\tau = \tau. \tag{4.9}$$

As a norm of a vector $\mathrm{d}\tau$ is invariant. It is the spacetime interval of two successive events in the observers life and τ, called the proper time of the observer, is the spacetime distance travelled by him. The personal time of an inertial observer is measured by the spacetime length of his worldline.

Even if the observer is not an inertial one but accelerates, for any short enough time interval he is in an approximately uniform motion. His personal time interval between two successive events in his life can still be reasonably measured by $\mathrm{d}\tau$ and his total lifetime by τ, the length of his now curved worldline. We call the length of the worldline of any observer, body or particle its proper time.

For τ we have

$$\tau = \int \mathrm{d}t\,\dot{\tau} \tag{4.10}$$

where

$$\dot{\tau} = \frac{\mathrm{d}\tau}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\sqrt{(\mathrm{d}t)^2 - \mathrm{d}\mathbf{x}^{\bullet}\mathrm{d}\mathbf{x}^{\bullet}} = \sqrt{1 - \dot{\mathbf{x}}^2} = \gamma(\dot{\mathbf{x}})^{-1} \tag{4.11}$$

and $\dot{\mathbf{x}}^2 = \dot{\mathbf{x}}^{\bullet}\dot{\mathbf{x}}^{\bullet}$. Note that the component number two of $\dot{\mathbf{x}}$ and $\dot{\mathbf{x}}^{\bullet}\dot{\mathbf{x}}^{\bullet}$ are both denoted by $\dot{\mathbf{x}}^2$. It is clear from context which one is meant.

Consider two events with timelike separation. We can always choose a frame in which they happen in the same spatial location. In that frame we have $\dot{\tau} = 1$ for a straight line between them. For any other line connecting the events we have $\dot{\tau} \leq 1$ so the proper time along them is smaller. In contrast to Euclidean space, in Minkowski space straight line maximizes the path length or proper time between two events.

Lightlike worldlines are of zero length and a spacelike worldline, which is a less useful concept, has an imaginary length.

If Ilkka takes a trip with a high-speed spacecraft around some near stars and Imppu stays on Earth, Imppu is in an inertial frame if the small acceleration of gravity is neglected, but Ilkka must change his velocity significantly to some day get back to Earth. His worldline is much more curved than Imppu's and so the proper time experienced by Imppu is greater than the one Ilkka experiences. When Ilkka and Imppu meet after the trip, they find that Imppu has aged significantly more.

4.4 Tensors

Consider a product of vectors, say $A^{\bullet}B^{{\color{red}\bullet}}C_{{\color{cyan}\bullet}}$. In a Poincaré transformation it transform by

$$A^{\bullet}B^{{\color{red}\bullet}}C_{{\color{cyan}\bullet}} \to \Lambda^{\bullet}{}_{{\color{green}\bullet}}\Lambda^{{\color{red}\bullet}}{}_{{\color{orange}\bullet}}\Lambda_{{\color{cyan}\bullet}}{}^{{\color{magenta}\bullet}}A^{{\color{green}\bullet}}B^{{\color{orange}\bullet}}C_{{\color{magenta}\bullet}}. \tag{4.12}$$

This is an example of a tensor. We call any quantity which transforms as if it were a product of vectors a tensor.

The number of indices of a tensor is called its rank. By definition, in a Poincaré transformation tensor gets one $\Lambda^{\bullet}{}_{{\color{red}\bullet}}$ for every upper and one $\Lambda_{\bullet}{}^{{\color{red}\bullet}}$ for every lower index. Vectors are tensors with one index and scalars are tensors with no indices at all. A scalar S transforms trivially, $S \to S$.

The metric $\eta_{\bullet{\color{red}\bullet}}$ is a tensor:

$$\eta_{\bullet{\color{red}\bullet}} \to \eta_{\bullet{\color{red}\bullet}} = \Lambda_{\bullet}{}^{{\color{cyan}\bullet}}\Lambda^{{\color{green}\bullet}}{}_{{\color{cyan}\bullet}}\eta_{{\color{red}\bullet}{\color{green}\bullet}} = \Lambda_{\bullet}{}^{{\color{cyan}\bullet}}\Lambda_{{\color{red}\bullet}{\color{cyan}\bullet}} = \Lambda_{\bullet}{}^{{\color{cyan}\bullet}}\Lambda_{{\color{red}\bullet}}{}^{{\color{green}\bullet}}\eta_{{\color{cyan}\bullet}{\color{green}\bullet}} \tag{4.13}$$

as is the metric $\eta^{\bullet{\color{red}\bullet}}$ with upper indices.

The product of two tensors is a tensor whose rank is the sum of the ranks of its factors. Also, since Lorentz transformation is linear, a linear combination of tensors which have the same number of upper and lower indices is a tensor.

We can set one upper and one lower index of a tensor equal and sum over it. That is called contraction. Since the transformation of the upper index is the inverse of the one of the lower index, these transformations cancel each other and the result is a tensor of lower rank. The inner product (4.7) of two vectors is an example of a contraction. Other examples are the metric with both indices $\eta^{\bullet}_{{\color{red}\bullet}} = \eta^{\bullet{\color{cyan}\bullet}}\eta_{{\color{red}\bullet}{\color{cyan}\bullet}}$ and any other raising or lowering of an index of a tensor.

The indices of a tensor can be thought as slots that can be connected to slots of other tensors by contraction. A lower slot is always connected with an upper one. Since this connection is just a sum of ordinary products of numbers, it is a linear operation, and so tensors can be thought of as linear maps from tensors to tensors.

Especially, if we connect every slot of a tensor to a vector, the result is a scalar. Mathematically tensors are often defined as multilinear maps from vectors to numbers. This idea makes it clear that actually it is not the tensor but its coordinate representation which transforms in the Poincaré transformation. The tensor itself is an invariant object.

Ordinary partial derivative

$$\frac{\partial}{\partial x^{\bullet}} \doteq \partial_{\bullet} \tag{4.14}$$

transforms in a Poincaré transformation by

$$\partial_{\bullet} \to (J^{-1})_{\bullet}{}^{{\color{red}\bullet}}\partial_{{\color{red}\bullet}} \tag{4.15}$$

where J is the Jacobian of the coordinate transformation. Since the Jacobian of a Poincaré transformation is just $\Lambda^{\bullet}{}_{{\color{red}\bullet}}$, then $(J^{-1})_{\bullet}{}^{{\color{red}\bullet}}$ is $\Lambda_{\bullet}{}^{{\color{red}\bullet}}$ and the partial derivative transforms as

$$\partial_{\bullet} \to \Lambda_{\bullet}{}^{{\color{red}\bullet}}\partial_{{\color{red}\bullet}}. \tag{4.16}$$

Partial derivative of a tensor, say $\partial_{\bullet}V_{{\color{red}\bullet}}$ transforms by

$$\partial_{\bullet}V_{{\color{red}\bullet}} \to \Lambda_{\bullet}{}^{{\color{cyan}\bullet}}\partial_{{\color{cyan}\bullet}}(\Lambda_{{\color{red}\bullet}}{}^{{\color{green}\bullet}}V_{{\color{green}\bullet}}) = \Lambda_{\bullet}{}^{{\color{cyan}\bullet}}\Lambda_{{\color{red}\bullet}}{}^{{\color{green}\bullet}}\partial_{{\color{cyan}\bullet}}V_{{\color{green}\bullet}} \tag{4.17}$$

so it is a tensor. When a tensor is partially differentiated, it gets one lower index more.

If a tensor is symmetric in two lower or two upper indices, say $T_{\bullet{\color{red}\bullet}} = T_{{\color{red}\bullet}\bullet}$, then this property is preserved in a Poincaré transformation:

$$T_{\bullet{\color{red}\bullet}} \to \Lambda_{\bullet}{}^{{\color{cyan}\bullet}}\Lambda_{{\color{red}\bullet}}{}^{{\color{green}\bullet}}T_{{\color{cyan}\bullet}{\color{green}\bullet}} = \Lambda_{\bullet}{}^{{\color{cyan}\bullet}}\Lambda_{{\color{red}\bullet}}{}^{{\color{green}\bullet}}T_{{\color{green}\bullet}{\color{cyan}\bullet}} = \Lambda_{{\color{red}\bullet}}{}^{{\color{green}\bullet}}\Lambda_{\bullet}{}^{{\color{cyan}\bullet}}T_{{\color{green}\bullet}{\color{cyan}\bullet}}. \tag{4.18}$$

Same is true for an antisymmetric $T_{\bullet{\color{red}\bullet}} = -T_{{\color{red}\bullet}\bullet}$ tensor.

Since d^4x is invariant in translations, rotations and boosts, it is invariant under the whole Poincaré group and an integral of a tensor over some spacetime region is a tensor.

Mathematically the Principle of Relativity means that the form of the equations describing Nature must be invariant in Poincaré transformations. It is said that such equations are covariant. If an equation is formed out of tensors, its both sides get transformed in the same way in a Poincaré transformation and the same equation holds in every coordinate system. Thus tensor equations are covariant and suitable for describing fundamental laws of Nature.

5 Matter in spacetime

The first object living in spacetime that comes to mind is a freely falling body. We will treat the body as a point particle with no intrinsic structure, its spatial coordinates $\mathbf{x}(t)$ as its only degrees of freedom. We will use the Principle of Least Action to determine the dynamics of the particle.

According to the Principle of Relativity, the equation governing the motion of the particle, $\delta S = 0$, must be covariant in the Poincaré transformation. If the action is a scalar, then its stationary points are unchanged by the transformation and the equation is covariant.

The length of the worldline of the particle, the proper time it experiences during the motion, is screaming to be used in the action. It is the only scalar easily formable of the motion of the particle with no other possibilities in sight. Since in Minkowski space there is no shortest path between two events but only the longest one, we take the action to be minus the worldline length to make it possible to have a minimum value. We have

$$S = -\int \mathrm{d}\tau. \tag{5.1}$$

Since a straight line maximizes the proper time between two events, the equation of motion is an equation of a straight line in spacetime.

The action can be written in terms of the time coordinate t and a Lagrangian L as

$$S = \int \mathrm{d}t\, L, \qquad L = -\dot{\tau}. \tag{5.2}$$

From (4.11) we see that the Lagrangian is $L = -\gamma(\dot{\mathbf{x}})^{-1}$. Note that it is not a scalar. Expanding it in $\dot{\mathbf{x}}$ we have

$$L = -1 + \frac{1}{2}\dot{\mathbf{x}}^2 + \cdots . \tag{5.3}$$

The first term is a constant and does not affect the equation of motion in a way or another. At slow velocities only the second term is relevant, so the Lagrangian of a nonrelativistic particle is $\frac{1}{2}\dot{\mathbf{x}}^2$.

5.1 Energy and momentum

The momentum of the particle is

$$\mathbf{p}^\bullet = \frac{\partial L}{\partial \dot{\mathbf{x}}^\bullet} = -\frac{\partial}{\partial \dot{\mathbf{x}}^\bullet}\gamma(\dot{\mathbf{x}})^{-1} = \gamma(\dot{\mathbf{x}})\dot{\mathbf{x}}^\bullet = \frac{\mathrm{d}t}{\mathrm{d}\tau}\frac{\mathrm{d}\mathbf{x}^\bullet}{\mathrm{d}t} = \frac{\mathrm{d}\mathbf{x}^\bullet}{\mathrm{d}\tau}. \tag{5.4}$$

When velocity reaches one, $\gamma(\dot{\mathbf{x}})$ goes to infinity and the momentum diverges. This illustrates the impossibility of superluminal motion. At slow velocities $\gamma(\dot{\mathbf{x}}) \approx 1$ so the momentum of a nonrelativistic particle is $\mathbf{p}^\bullet = \dot{\mathbf{x}}^\bullet$.

The energy of the particle is

$$E = \mathbf{p}^\bullet \dot{\mathbf{x}}^\bullet - L = \gamma(\dot{\mathbf{x}})\dot{\mathbf{x}}^2 + \gamma(\dot{\mathbf{x}})^{-1} = \gamma(\dot{\mathbf{x}}) = \frac{\mathrm{d}t}{\mathrm{d}\tau} \doteq p^0. \tag{5.5}$$

At velocities close to one the energy diverges. Even when the particle is at rest, its energy is one.

Energy and momentum can be combined into quantity

$$p^\bullet = \frac{\mathrm{d}x^\bullet}{\mathrm{d}\tau} \tag{5.6}$$

called four-momentum. It is a vector since $\mathrm{d}x^\bullet$ is a vector and $\mathrm{d}\tau$ is a scalar. It holds that

$$p_\bullet p^\bullet = \frac{\mathrm{d}x_\bullet}{\mathrm{d}\tau}\frac{\mathrm{d}x^\bullet}{\mathrm{d}\tau} = \frac{\mathrm{d}\tau^2}{\mathrm{d}\tau^2} = 1 \tag{5.7}$$

so

$$E^2 = \mathbf{p}^2 + 1 \tag{5.8}$$

where $\mathbf{p}^2 \doteq \mathbf{p}^\bullet \mathbf{p}^\bullet$.

The particle action is Poincaré invariant, so Noether's theorem tells that energy and momentum are conserved. These results are useful to think of as the conservation of the four-momentum caused by the spacetime translation invariance.

When the particle is playing with other physical entities, we need to quantify how dominant its term is in the total action. We do this my multiplying the action of the particle by a number m called the mass of the particle. The four-momentum becomes

$$p^\bullet = m\frac{\mathrm{d}x^\bullet}{\mathrm{d}\tau} \tag{5.9}$$

and (5.8) becomes

$$E^2 = \mathbf{p}^2 + m^2. \tag{5.10}$$

The mass of the particle is sometimes called its rest energy.

5.2 Reality of the rest energy

From (5.4) and (5.5) we see that $\mathbf{p}^\bullet = m\gamma(\dot{\mathbf{x}})\dot{\mathbf{x}}^\bullet = E\dot{\mathbf{x}}^\bullet$. The momentum component $p^\bullet$ is dependent not only on $\dot{\mathbf{x}}^\bullet$ but also on $\dot{\mathbf{x}}^\bullet \dot{\mathbf{x}}^\bullet$.

If say $\dot{\mathbf{x}}^1$ is high but $\dot{\mathbf{x}}^2$ and $\dot{\mathbf{x}}^3$ are low, then $E > m$ is dominated by $\dot{\mathbf{x}}^1$ and is approximately independent of $\dot{\mathbf{x}}^2$ and $\dot{\mathbf{x}}^3$. In the directions $\mathbf{x}^2$ and $\mathbf{x}^3$ the energy E acts as an effective mass greater than m. The greater $\dot{\mathbf{x}}^1$ is, the harder it is to change $\dot{\mathbf{x}}^2$ and $\dot{\mathbf{x}}^3$.

Consider a body consisting of particles moving in a disorganized fashion with four-momentums p_a. The four-momentum of the body as a whole is simply the sum

$$p^\bullet = \sum_a p_a^\bullet. \tag{5.11}$$

As a sum of vectors it is a vector.

The body can be considered to be at rest in frames in which the total momentum $\mathbf{p}^\bullet$ is zero. We denote by M the energy p^0 in these frames. Since the constituent particles move, clearly $M > \sum_a m_a$.

Since in these frames the body is at rest, its spatial coordinates are constant and its proper time runs like t. The four-momentum $p^\bullet$ can be written as

$$p^\bullet = M\frac{\mathrm{d}x^\bullet}{\mathrm{d}\tau}. \tag{5.12}$$

This is the four-momentum of a particle of mass M. If we now put the body in motion, its momentum gets Lorentz boosted and becomes

$$\mathbf{p}^\bullet = M\gamma(\dot{\mathbf{x}})\dot{\mathbf{x}}^\bullet. \tag{5.13}$$

as is seen from (5.4). The internal kinetic energy of the constituent particles contributes to the mass of the body, so we see that mass really is the energy of a body at rest. There is no reason to make a fundamental distinction between mass and energy.

5.3 Interactions

In reality particles are not free. They are acted on by forces. As is known from nonrelativistic particle mechanics, a force acting on a particle can be described by a potential term $V = V(\mathbf{x}(t))$ in the Lagrangian. The action corresponding to this is

$$S = \int \mathrm{d}t(-m\gamma(\dot{\mathbf{x}})^{-1} + V) = \int -\mathrm{d}\tau\, m + \mathrm{d}t\, V. \tag{5.14}$$

However, this action is not Poincaré invariant.

When particles move with nonrelativistic velocities, it holds that $\mathrm{d}t \gg \mathrm{d}\mathbf{x}^\bullet$. We can very well promote V to a time component A_0 of a vector A and $\mathrm{d}t\, V$ to scalar $\mathrm{d}x^\bullet A_\bullet$. If A is somewhat timelike and the motion of the particle nonrelativistic, then the new interaction term reduces to the old one. The action becomes

$$S = \int -\mathrm{d}\tau\, m + \mathrm{d}x^\bullet A_\bullet \tag{5.15}$$

which is manifestly Poincaré invariant.

When there are many particles at play, the action involves the sum of the actions of every individual particle:

$$S = \sum_a \int -\mathrm{d}\tau_a\, m_a + \mathrm{d}x_a^\bullet A_{\bullet a}. \tag{5.16}$$

Usually the forces are due to interactions between the particles. This means that the potentials A_a are somehow generated by the particles.

Since information cannot propagate instaneously, the particle cannot get immediately influenced by the motion of other particles. The potential A_a can depend only on the location and time coordinate of the corresponding particle. The influence of other particles must somehow come through the time dependence.

Since the particle senses the potential only in its immediate vicinity, we can use the same A for every particle, if A is thought to be a function of the whole spacetime. We set $A_a(x_a) = e_a A(x_a)$, where e_a is a number that quantifies the interaction strength of the particle. The potential $A = A(x)$ is called a field.

5.4 Covariant action principle

When a particle moves, eventually the effect of its motion must be felt by other particles. The effect propagates in the field like a wave propagates on the surface of water. The field possesses its own dynamical nature and thus needs its own term in the action.

Since the field exists in the same fashion at every point in spacetime, the action is clearly an integral over time and space. We write

$$S = \int \mathrm{d}^4x\, \mathcal{L} \tag{5.17}$$

or

$$S = \int \mathrm{d}t\, L, \qquad L = \int \mathrm{d}^3\mathbf{x}\, \mathcal{L}. \tag{5.18}$$

The function $\mathcal{L}$ is called Lagrangian density or sloppily just the Lagrangian. As with all Lagrangians of dynamical systems, we assume L to be a function of the configuration of the field, which is the values of $A(\mathbf{x})$ at every point of space at a constant t, and its time derivative. The dependence in configuration propably involves the amount of its wrinkliness, which is quantified by the spatial partial derivative. We take $\mathcal{L} = \mathcal{L}(A, \partial A)$.

The stationary points must be invariant in Poincaré transformation. If $\mathcal{L}$ is a scalar, then the action is a scalar and the stationary points remain invariant. This fact also demands that $\mathcal{L}$ must depend on both temporal and spatial derivatives. A scalar could not be formed from time derivative alone; Lorentz boosts mix temporal derivative with spatial ones and vice versa.

The familiar Principle of Least Action states that for a true motion and for a small variation δA it holds that

$$\delta \int_{t_a}^{t_b} \mathrm{d}t\, L = \int_R \mathrm{d}^4x\, \delta\mathcal{L} = 0 \tag{5.19}$$

where R is all the spacetime between t_a and t_b. Since the variation is small, the integrand $\delta\mathcal{L}$ is

$$\delta\mathcal{L} = \mathcal{L}(A+\delta A, \partial A + \delta(\partial A)) - \mathcal{L}(A, \partial A) \tag{5.20}$$

$$= \frac{\partial\mathcal{L}}{\partial A_\bullet}\delta A_\bullet + \frac{\partial\mathcal{L}}{\partial(\partial_{\color{red}\bullet} A_\bullet)}\delta(\partial_{\color{red}\bullet} A_\bullet). \tag{5.21}$$

Noting that

$$\delta(\partial_{\color{red}\bullet} A_\bullet) = \partial_{\color{red}\bullet}(A_\bullet + \delta A_\bullet) - \partial_{\color{red}\bullet} A_\bullet = \partial_{\color{red}\bullet}\delta A_\bullet \tag{5.22}$$

it becomes

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial A_\bullet}\delta A_\bullet + \frac{\partial\mathcal{L}}{\partial(\partial_{\color{red}\bullet} A_\bullet)}\partial_{\color{red}\bullet}(\delta A_\bullet) \tag{5.23}$$

$$= \frac{\partial\mathcal{L}}{\partial A_\bullet}\delta A_\bullet + \partial_{\color{red}\bullet}\left(\frac{\partial\mathcal{L}}{\partial(\partial_{\color{red}\bullet} A_\bullet)}\delta A_\bullet\right) - \partial_{\color{red}\bullet}\frac{\partial\mathcal{L}}{\partial(\partial_{\color{red}\bullet} A_\bullet)}\delta A_\bullet. \tag{5.24}$$

From here it is difficult to move on because the integration region is infinite in space. However, if the field is sealed in a sturdy box which forces the field to be zero at the walls, we can consider only the variations that vanish at the walls. Further, the rigid walls prevent the field inside the box from knowing about the field outside, so we can shrink R to include only the spacetime inside the box.

The walls cause the ripples in the field to be reflected, but if the ripples are initially contained in some region in space and the box is large enough, then due to the finite speed of propagation they do not have time to hit the walls. The field evolution is exactly the same as it would be without the walls. The box trick allows us to find the equation of motion of the real field pervading all space.

The middle term in (5.24) is a divergence and by Gauss's theorem it contributes to the action only at the boundary of R. The contribution is proportional to δA so it vanishes. We are left with

$$\delta\mathcal{L} = \left(\frac{\partial\mathcal{L}}{\partial A_\bullet} - \partial_{\color{red}\bullet}\frac{\partial\mathcal{L}}{\partial(\partial_{\color{red}\bullet} A_\bullet)}\right)\delta A_\bullet. \tag{5.25}$$

For the integral of this to vanish for any $\delta A_\bullet$ it must be that the term in parenthesis vanishes. We arrive at the Lagrange's equation of motion

$$\frac{\partial\mathcal{L}}{\partial A_\bullet} - \partial_{\color{red}\bullet}\frac{\partial\mathcal{L}}{\partial(\partial_{\color{red}\bullet} A_\bullet)} = 0 \tag{5.26}$$

for the field. Obviously this procedure can be carried out in any inertial frame and so the equation of motion is covariant.

The region R is a rectangular region in spacetime. However, as we saw in Section 3.11, it does not stay as such in a Lorentz boost but turns into a rhombus-like region. The concept of spacetime between two values of t in some spatial region is not a covariant concept.

But whatever the form of R is, Gauss's theorem tells us that the equation of motion is equivalent to the action having a stationary value with respect to variations that vanish at the boundary of R. This motivates us to formulate the Principle of Least Action in spacetime as

$$\delta \int_R \mathrm{d}^4 x\, \mathcal{L}(A, \partial A) = 0 \tag{5.27}$$

for any finite spacetime region R and variations δA which vanish at the boundary of R. This is completely covariant formulation of the action principle.

At this point a significant change in our viewpoint takes place. The old nonrelativistic intuition thinks the configuration of the field to be all the values $A_\bullet(\mathbf{x})$ of the field in all the space at a constant t and the evolution of this infnity of degrees of freedom happening in one dimensional time. This is not covariant thinking since hypersurface defined by constant t is not a hypersurface of constant t in boosted frames.

The relativistic treatment suggests to think that the configuration is only the four degrees of freedom $A_\bullet$ and the evolution of the field happens in four dimensional time or spacetime. From that perspective the Lagrangian density really is the Lagrangian.

5.5 Noether's theorem for fields in spacetime

If there is a way to transform the motions of the system in such a way that physical motions get transformed to physical motions, then the system is said to possess a symmetry.

Mathematically the symmetry transformation must take stationary points of the action to stationary points. This is accomplished if the Lagrangian $\mathcal{L}$ remains invariant or if it changes by a divergence, since by Gauss's theorem the divergence contributes to the action only at the boundary of the integration region and thus does not affect the equation of motion. We write $\delta\mathcal{L} = \partial_\bullet f^\bullet$ for some vector f.

On the other hand, if the symmetry is continuous, we can perform a small transformation $\delta A(x)$ and write

$$\delta\mathcal{L} = \frac{\partial \mathcal{L}}{\partial A_\bullet}\delta A_\bullet + \frac{\partial \mathcal{L}}{\partial(\partial_\bullet A_{\color{red}\bullet})}\delta(\partial_\bullet A_{\color{red}\bullet}) \tag{5.28}$$

$$= \frac{\partial \mathcal{L}}{\partial A_\bullet}\delta A_\bullet + \frac{\partial \mathcal{L}}{\partial(\partial_\bullet A_{\color{red}\bullet})}\partial_\bullet \delta A_{\color{red}\bullet}. \tag{5.29}$$

Using (5.26), the equation of motion for the field, $\delta\mathcal{L}$ becomes

$$\delta\mathcal{L} = \partial_\bullet \frac{\partial \mathcal{L}}{\partial(\partial_\bullet A_{\color{red}\bullet})}\delta A_{\color{red}\bullet} + \frac{\partial \mathcal{L}}{\partial(\partial_\bullet A_{\color{red}\bullet})}\partial_\bullet \delta A_{\color{red}\bullet} = \partial_\bullet \left(\frac{\partial \mathcal{L}}{\partial(\partial_\bullet A_{\color{red}\bullet})}\delta A_{\color{red}\bullet} \right). \tag{5.30}$$

In totality we have

$$\partial_\bullet I^\bullet = 0 \qquad \text{with} \qquad I^\bullet = \frac{\partial \mathcal{L}}{\partial(\partial_\bullet A_{\color{red}\bullet})}\delta A_{\color{red}\bullet} - f^\bullet. \tag{5.31}$$

Denoting transformed quantities with tildes, the quantity multiplying δA in I transforms by

$$\frac{\partial \mathcal{L}}{\partial(\partial_\bullet A_{\color{red}\bullet})} \to \frac{\partial \mathcal{L}}{\partial(\tilde{\partial}_\bullet \tilde{A}_{\color{red}\bullet})} \tag{5.32}$$

$$= \frac{\partial \mathcal{L}}{\partial(\partial_{\color{cyan}\bullet} A_{\color{green}\bullet})} \frac{\partial(\partial_{\color{cyan}\bullet} A_{\color{green}\bullet})}{\partial(\tilde{\partial}_\bullet \tilde{A}_{\color{red}\bullet})}. \tag{5.33}$$

We have $\tilde{A}_\bullet = \Lambda_{\color{red}\bullet}{}^{\bullet} A_\bullet$ and so $A_\bullet = \Lambda^{\color{red}\bullet}{}_\bullet \tilde{A}_{\color{red}\bullet}$. Same holds for ∂, so we can write

$$\frac{\partial \mathcal{L}}{\partial(\partial_\bullet A_{\color{red}\bullet})} \to \frac{\partial \mathcal{L}}{\partial(\partial_{\color{cyan}\bullet} A_{\color{green}\bullet})} \frac{\partial(\Lambda^{\color{orange}\bullet}{}_{\color{cyan}\bullet} \Lambda^{\color{magenta}\bullet}{}_{\color{green}\bullet} \tilde{\partial}_{\color{orange}\bullet} \tilde{A}_{\color{magenta}\bullet})}{\partial(\tilde{\partial}_\bullet \tilde{A}_{\color{red}\bullet})} \tag{5.34}$$

$$= \frac{\partial \mathcal{L}}{\partial(\partial_{\color{cyan}\bullet} A_{\color{green}\bullet})} \Lambda^{\color{orange}\bullet}{}_{\color{cyan}\bullet} \Lambda^{\color{magenta}\bullet}{}_{\color{green}\bullet} \frac{\partial(\tilde{\partial}_{\color{orange}\bullet} \tilde{A}_{\color{magenta}\bullet})}{\partial(\tilde{\partial}_\bullet \tilde{A}_{\color{red}\bullet})} \tag{5.35}$$

$$= \frac{\partial \mathcal{L}}{\partial(\partial_{\color{cyan}\bullet} A_{\color{green}\bullet})} \Lambda^{\bullet}{}_{\color{cyan}\bullet} \Lambda^{\color{red}\bullet}{}_{\color{green}\bullet}. \tag{5.36}$$

We thus find the transformation law of a tensor with two upper indices. As a mnemonic it can be thought that when something is divided by a quantity with indices, the positions of the indices are also 'divided' or inverted and the indices change their floor. The quantity I is a tensor contracted with a vector minus a vector, with indices properly placed, so it is a vector.

The condition $\partial_\bullet I^\bullet$ can be written as

$$\partial_0 I^0 = \partial_\bullet \mathbf{I}^\bullet \tag{5.37}$$

and be integrated over some spatial region R to get

$$\partial_0 \int_R \mathrm{d}^3x\, I^0 = \int_R \mathrm{d}^3x\, \partial_\bullet \mathbf{I}^\bullet. \tag{5.38}$$

By Gauss's theorem the right-hand side is the integral of the flux of $\mathbf{I}$ through the boundary of R. $\mathbf{I}$ quantifies how much of I^0 is flowing in or out of R. If R is so large that it contains every disturbation propagating in the field, then $I = 0$ at the boundary and it holds that

$$\partial_0 \int_R \mathrm{d}^3x\, I^0 = 0. \tag{5.39}$$

Quantity I is called a conserved current and I^0 a conserved charge or charge density. The notion that a continuous symmetry implies the existence of a conserverd current and charge is the spacetime from of Noether's theorem.

If the Lagrangian $\mathcal{L}$ describes some fundamental interaction of Nature, it is Poincaré invariant. Especially it is invariant in translations. If we make a small translation $x \to x + \epsilon$, then $\delta A_\bullet = \epsilon^{\color{red}\bullet} \partial_{\color{red}\bullet} A_\bullet$. As a scalar the Lagrangian transforms in the same way so

$$\delta \mathcal{L} = \epsilon^\bullet \partial_\bullet \mathcal{L} = \partial_\bullet (\epsilon^\bullet \mathcal{L}). \tag{5.40}$$

We have $f^\bullet = \epsilon^\bullet \mathcal{L} = \epsilon^{\color{red}\bullet} \eta^\bullet{}_{\color{red}\bullet} \mathcal{L}$ and

$$I^\bullet = \frac{\partial \mathcal{L}}{\partial(\partial_\bullet A_{\color{red}\bullet})} \epsilon^{\color{cyan}\bullet} \partial_{\color{cyan}\bullet} A_{\color{red}\bullet} - \epsilon^{\color{red}\bullet} \eta^\bullet{}_{\color{red}\bullet} \mathcal{L} = \epsilon^{\color{red}\bullet} T^\bullet{}_{\color{red}\bullet} \tag{5.41}$$

with

$$T^\bullet{}_{\color{red}\bullet} = \frac{\partial \mathcal{L}}{\partial(\partial_\bullet A_{\color{cyan}\bullet})} \partial_{\color{red}\bullet} A_{\color{cyan}\bullet} - \eta^\bullet{}_{\color{red}\bullet} \mathcal{L}. \tag{5.42}$$

The quantity T is formed appropriately out of tensors and thus is a tensor.

Now the conservation equation is

$$0 = \partial_\bullet I^\bullet = \epsilon^{\color{red}\bullet} \partial_\bullet T^\bullet{}_{\color{red}\bullet}. \tag{5.43}$$

Since ϵ is arbitrary, it must hold that $\partial_{\color{red}\bullet} T^{\color{red}\bullet}{}_\bullet = 0$. We thus have four conserved quantities.

We recognize the charge

$$\int \mathrm{d}^3x \, T^0{}_0 = \int \mathrm{d}^3x \, \frac{\partial \mathcal{L}}{\partial(\partial_0 A_\bullet)} \partial_0 A_\bullet - L \tag{5.44}$$

which correponds to time translation as the Hamiltonian of the field formed as a Legendre transformation of the Lagrangian. We interpret $T^0{}_0$ to be the energy density of the field.

The three conserved charges $\mathbf{T}^0{}_\bullet$ corresponding to space translations are

$$\mathbf{T}^0{}_\bullet = \frac{\partial \mathcal{L}}{\partial(\partial_0 A_{\color{red}\bullet})} \partial_\bullet A_{\color{red}\bullet}. \tag{5.45}$$

We interpret these as the momentum densities of the field. The quantity T, the conserved current arising from the spacetime translation invariance, is called the energy-momentum tensor of the field A.

6 Discussion

The theory of spacetime we have developed here is called the Special Theory of Relativity. It was developed in the beginning of the twentieth century by Hendrik Antoon Lorentz, Albert Einstein and Hermann Minkowski.

Special Relativity, which is based essentially on the Principle of Relativity and the notion that the speed of light does not depend on the observer, points out that space and time are not absolute concepts themselves. Only their union, which we call spacetime, enjoys physical objectivity independent of the observer.

The defining property of Euclidean space is the square of the distance $(\Delta\mathbf{x}^1)^2 + (\Delta\mathbf{x}^2)^2 + (\Delta\mathbf{x}^3)^2$. In the same fashion the Minkowski space is determined by the square of the spacetime interval $(\Delta t)^2 - (\Delta\mathbf{x}^1)^2 - (\Delta\mathbf{x}^2)^2 - (\Delta\mathbf{x}^3)^2$ which is captured by the metric η.

The concept of an absolute 'now' is not extendable outside a single spacetime point, since a hypersurface of a constant t is not a Poincaré invariant concept. Mass is merely the energy of a body at rest when considered as a whole. Many physical quantities, for example energy and momentum, are seen to be entangled in the same way as time and space, which is mathematically described by vector and tensor calculus on the Minkowski space.

6.1 Field theories

We discussed particles interacting through a field A in spacetime. The total action of such system is

$$S = \sum_a \int -\mathrm{d}\tau_a\, m_a + e_a \mathrm{d}x_a^{\bullet} A_{\bullet}(x_a) + \int \mathrm{d}^4x\, \mathcal{L}. \tag{6.1}$$

With a suitable $\mathcal{L}$ it is the action of the classical theory of electromagnetism.

The interaction part looks very satisfying, but the particle part is not that elegant. The particles with their properties m_a and e_a must be added by hand to the theory. Further, the proper time as the action of the particle is not compatible with the view of a system evolving in four rather than one dimensional time. Classical theories are generally not able to give understanding to the origin, properties or number of the particles.

A quantum treatment of the electromagnetic field reveals that the quanta of the field behave exactly like particles. On the other hand, the Schrödinger equation of a particle is a field equation for its propability density. These notions suggest that possibly every particle in Nature could be described by a field.

During the twentieth century quantum field theory was developed and it was found that this is the case. The standard model of particle physics, which was completed in the seventies, is a field theory which describes all the particles observed in Nature.

6.2 The Lorentz algebra

When seeking for field theories to describe elementary particles, the Principle of Relativity must be satisfied. The equations of the theory must be covariant in the Poincaré transformation.

Tensor fields, for example our A, are suitable for this task, but they are not the only possibilities. Consider a field ψ which transforms in a Poincare transformation by $\psi \to \Omega(\Lambda)\psi$. If two Poincaré transformations are performed, the field transforms by $\psi \to \Omega(\Lambda)\Omega(\tilde{\Lambda})\psi$. This must clearly be the same thing as if we performed the two transformations as one, when $\psi \to \Omega(\Lambda\tilde{\Lambda})\psi$. It must hold that

$$\Omega(\Lambda)\Omega(\tilde{\Lambda}) = \Omega(\Lambda\tilde{\Lambda}). \tag{6.2}$$

It is said that the transformations Ω form a representation of the Lorentz group. Tensors, symmetric tensors and antisymmetric tensors of different ranks are representations of the Lorentz group.

From (4.3) it is simple to see that the commutators $AB - BA \doteq [A, B]$ of the generators of the Lorentz group satisfy

$$[J^{\bullet{\color{red}\bullet}}, J^{{\color{cyan}\bullet}{\color{green}\bullet}}] = \eta^{{\color{red}\bullet}{\color{cyan}\bullet}} J^{\bullet{\color{green}\bullet}} - \eta^{\bullet{\color{cyan}\bullet}} J^{{\color{red}\bullet}{\color{green}\bullet}} - \eta^{{\color{red}\bullet}{\color{green}\bullet}} J^{\bullet{\color{cyan}\bullet}} + \eta^{\bullet{\color{green}\bullet}} J^{{\color{red}\bullet}{\color{cyan}\bullet}}. \tag{6.3}$$

This commutation relation is called the Lie algebra of the Lorentz group. Using the Baker–Campbell–Hausdorff formula

$$e^A e^B = \exp\left(A + B + \frac{1}{2}[A, B] + \cdots\right) \tag{6.4}$$

we see that the Lie algebra determines the algebra of the whole group. This means that a group being a representation of the Lorentz group is equivalent to it having the same Lie algebra. In addition to scalar and vector fields, the standard model of particle physics involves so-called spinor fields, which are kind of non-integer rank tensor fields.

6.3 General Theory of Relativity

When a ship sails towards the horizon, it looks that very far away it starts to sink under the water. However, if we go after the ship, we notice that in reality it did not sink. Rather the surface of water was curved like a sphere. Long expeditions made by Ferdinand Magellan and others show that we can go around Earth. Earth really is a sphere.

It is remarkable that Magellan did not fall away from Earth. He noted that many things, for example culture and weather differ around the world, he did not notice anything unusual in the freefall motion. It always happens towards the center of Earth. Further observations show that far away from Earth freefall motion happens always towards massive objects, for instance a planet or a galaxy.

This seems to be in conflict with the Principle of Relativity. The reference frames of observers falling at north and south poles are definitely not moving with a uniform velocity relative to each other.

However, within small distances this distortion of the picture of Special Relativity does not happen. We conclude that a reference frame can be an inertial one in a small but not large region of space, like a coordinate system on Earth can be Cartesian in only a small area. Like the surface of Earth, spacetime is curved.

The source of this curvature clearly seems to be mass, which is as we have seen equivalent to energy. Furthermore, energy is entangled to momentum. The source of the curvature thus is the energy-momentum tensor. The theory of curved spacetime and its dynamics is called the General Theory of Relativity.