Following recommendations of Bell "how to teach special relativity", we give an introduction into the theories of relativity which introduces not only the spacetime interpretation. It starts, instead, with the Lorentz ether interpretation of special relativity as well as a generalization of the Lorentz ether interpretation to the Einstein equations of general relativity in harmonic coordinates. The impossibility to identify, by observation, absolute time in the Lorentz ether, and the preferred background coordinates in its generalization to relativistic gravity is, then, used to introduce and justify the spacetime interpretation.

The differences between the ether interpretations and the spacetime interpretations are considered and discussed. While for special relativity, the only difference seems to be the possibility to derive Bell's inequality, much more serious differences appear in relativistic gravity. While the spacetime interpretation is background-independent, the ether interpretation requires a fixed Newtonian background. This excludes solutions incompatible with this background, in particular solution with non-trivial topology as well as with causal loops. The existence of a background also leads to different notions of completeness and homogeneity. The quantization problems depend on the existence of a background too.

The final decision which interpretation is preferable is left to the reader.

An “ether-based” introduction into the theories of relativity

I. Schmelzer

August 25, 2018

Chapter 1

Introduction

The aim of this text is to give an introduction into special and general relativity which gives not only the standard spacetime interpretation of relativty, as proposed by Minkowski for special relativity and developed into the curved spatime interpretation of general relativity by Einstein. It also gives an introduction into the Lorentz ether interpretation of special relativity, which can be seen as the original interpretation of the theory developed by Lorentz, Poincare and Einstein. The failure of the Lorentz ether to extend to relativistic gravity was, of course, a strong argument against the Lorentz ether, and has probably played a decisive role in the rejection of the ether concept. But it was only a historical accident.

Today it is known that there exists an extension of the Lorentz ether interpretation to the Einstein equations of general relativity in harmonic gauge. This fact alone does not mean that this interpretation has any advantages in comparison with the spacetime interpretation. But even for those who reject the ether completely it makes sense not to use a historical accident which is unfortunate for the ether interpretation, but, instead, to present their arguments against the best imaginable defense of an ether interpretation.

Moreover, to have several interpretations is a value in itself, because it allows, in a much more objective way, to distinguish those parts of the theory which are really physical from those which are metaphysical. Indeed, the criterion to distinguish the two is simple: The really physical results should be the same in all interpretations. Instead, those statements which differ in different interpretations, or make sense only in one of the interpretations, have to be classified as metaphysical. Having only a single interpretation, it would be much more difficult to make this distinction.

The ability to identify correctly the metaphysical aspects of the spacetime interpretation may become essential for the development of quantum gravity or a theory of everything, because the metaphysical elements can be given up without any problem, while the physical parts have to be recovered accurately even if the more fundamental theory uses completely different metaphysics.

Chapter 2

Special Relativity

2.1 How to construct Doppler-shifted solutions of a wave equation

The wave equation is the equation

$$\Box u(\vec{x},t) = \left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right) u(\vec{x},t) = 0$$

The operator ∇^2 is the Laplace operator, which acts on the spatial variables, so that in the three-dimensional case we have

$$\nabla^2 = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right).$$

The constant c is the speed of the wave. In relativity, c is the speed of light waves, but such a wave equation can be used also to describe sound waves, or waves on the surface of water.

Let's assume that we have found a particular solution of this equation $u_0(x,t)$. Then there exists a surprisingly simple method to construct other solutions of the same wave equation. We can choose an arbitrary parameter $|v| < c$, and define the following coordinates:

$$t' = \gamma\left(t - \frac{v}{c^2}\,x\right), \quad x' = \gamma(x - v\,t), \quad y' = y, \quad z' = z, \tag{2.1}$$

with $\gamma = \frac{1}{\sqrt{1-v^2/c^2}}$.

Then, all we have to do is to replace in the solution $u_0(x,y,z,t)$ all the x,y,z,t by x',y',z',t', and use the formula above to obtain another, different function $u_v(x,y,z,t)$:

$$u_0(x,y,z,t) \rightarrow$$
$$u_v(x,y,z,t) = u_0(x'(x,y,z,t), y'(x,y,z,t), z'(x,y,z,t), t'(x,y,z,t))$$

Surprisingly, this function $u_v(x,y,z,t)$ is also a solution of the same wave equation. You can simply try it out:

$$\begin{aligned}
\frac{\partial}{\partial t}u_v &= \frac{\partial t'}{\partial t}\frac{\partial}{\partial t'}u_0 + \frac{\partial x'}{\partial t}\frac{\partial}{\partial x'}u_0 = \gamma\left(\frac{\partial}{\partial t'} - v\frac{\partial}{\partial x'}\right)u_0,\\
\frac{\partial^2}{\partial t^2}u_v &= \gamma^2\left(\frac{\partial^2}{\partial t'^2} - 2v\frac{\partial}{\partial t'}\frac{\partial}{\partial x'} + v^2\frac{\partial^2}{\partial x'^2}\right)u_0\\
\frac{\partial}{\partial x}u_v &= \frac{\partial t'}{\partial x}\frac{\partial}{\partial t'}u_0 + \frac{\partial x'}{\partial x}\frac{\partial}{\partial x'}u_0 = \gamma\left(-\frac{v}{c^2}\frac{\partial}{\partial t'} + \frac{\partial}{\partial x'}\right)u_0,\\
\frac{\partial^2}{\partial x^2}u_v &= \gamma^2\left(\frac{v^2}{c^4}\frac{\partial^2}{\partial t'^2} - 2\frac{v}{c^2}\frac{\partial}{\partial t'}\frac{\partial}{\partial x'} + \frac{\partial^2}{\partial x'^2}\right)u_0\\
\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2}\right)u_v &= \gamma^2\left(\left(1-\frac{v^2}{c^2}\right)\frac{1}{c^2}\frac{\partial^2}{\partial t'^2} + \frac{v^2-c^2}{c^2}\frac{\partial^2}{\partial x'^2}\right)u_0\\
&= \left(\frac{1}{c^2}\frac{\partial^2}{\partial t'^2} - \frac{\partial^2}{\partial x'^2}\right)u_0\\
\left(\frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)u_v &= \left(\frac{\partial^2}{\partial y'^2} + \frac{\partial^2}{\partial z'^2}\right)u_0\\
\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right)u_v &= \left(\frac{1}{c^2}\frac{\partial^2}{\partial t'^2} - \nabla'^2\right)u_0 = 0.
\end{aligned}$$

The transformation of the coordinates (2.1) which we have used here to create the new solution is, in the case of c being the speed of light, named Lorentz transformation. For other wave equations, like sound waves or water waves, the name “Lorentz transformation” is not used, and the Lorentz transformations are seldom used.

But, nonetheless, the mathematics of the Lorentz transformation works in the same way for these equations too. The new solutions which can be obtained by this method are also well-known, they are the Doppler-shifted solutions. They have a well-defined physical meaning: If the source of the wave is, in the original solution, at rest, then the Doppler-shifted solution is the solution where the same source is moving with the velocity v.

Note that both solutions are physically different. The source is, in the first solution, at rest, while it moves in the second solution. And if the first solution has spherical symmetry, the second solution does no longer have such a symmetry: If the train moves in your direction, you hear a different sound than if the train moves away. Despite these differences, some properties remain unchanged: namely the speed of the wave remains the same.

2.2 Some general denotations

In a self-consistent textbook, all the denotations used should be defined. In relativity, especially in general relativity, a lot of mathematical formulas are used which do not have their origin in relativity. Relativity is simply using a general mathematical formalism, which has been developed for differential geometry in general. So, these are formulas which have useful applications already in standard three-dimensional Euclidean geometry. They are, in particular, useful if one wants to handle curved coordinates, like polar or spherical coordinates. The appropriate place to learn them would be some mathematical book about differential geometry.

2.2.1 Mathematical denotations to handle coordinate transformations

It is helpful to introduce a lot of mathematical conventions, conventions which allow to write memorize quite short formulas for otherwise quite long and boring sets of equations. Let's start with some pure mathematical conventions, used in differential geometry for computations with arbitrary numbers of coordinates. What we really need here are the three coordinates which denote a point in usual three-dimensional space and the four coordinates which denote an event – a point in space and a moment in time. But there are many "higher-dimensional spaces" in physics, like the so-called "configuration space": If we want to describe the configuration of N point particles, you have to define $3N$ coordinates – three for each particle, and this is what a mathematician names a $3N$-dimensional space.

The Lorentz transformation (2.1) is an example of a coordinate transformation. So, let's consider some purely mathematical rules how to handle coordinate transformations. Coordinates are functions on the space, usually denoted x^i, with an upper index. To fix a point in an n-dimensional space one needs n coordinates. A transformation of one system of coordinates into another one needs, therefore, n functions $x^{i'}(x) = x^{i'}(x^1, \ldots, x^n)$ which express the new coordinates $x^{i'}$ as functions of the old coordinates x^i. [1]

Often enough, one also needs the inverse transformation $x^i(x') = x^i(x^{1'}, \ldots, x^{n'})$, the old coordinates expressed as functions of the new coordinates.

How the various mathematical objects have to be transformed depends on these objects. So, for a point, which is defined by the coordinates x^i, one needs the $x^{i'}(x)$ to compute the new coordinates. Instead, for a function $f(x)$ defined for all points of the space, one needs the inverse transformation $x^i(x')$ to compute the new function values $f(x(x'))$. [2]

There are not only points and functions one has to transform, there are also other objects, like vector fields. If there is some flow, say, of water, then we have in every point of the flow a velocity, which defines a field of velocities $v^i(x)$. The component v^i is the component of the vector in the direction of the coordinate x^i, thus if the coordinates will be changed, we have to change the $v^i(x)$ to some $v^{i'}(x')$ too. The rules are the following:

$$v^{i'}(x'(x)) = \sum_{i=1}^{n} \frac{\partial x^{i'}(x)}{\partial x^i} v^i(x), \qquad v^i(x(x')) = \sum_{i'=1}^{n} \frac{\partial x^i(x')}{\partial x^{i'}} v^{i'}(x').$$

One should not forget that the two matrices $\frac{\partial x^{\mu'}}{\partial x^\mu}$ and $\frac{\partial x^\mu}{\partial x^{\mu'}}$ are different. To compute the first, we need the new coordinates x' as functions of the old coordinates x, so that $x^{\mu'} = x^{\mu'}(x)$. Then we can compute the Jacobi matrix of all the partial derivatives $\frac{\partial x^{\mu'}(x)}{\partial x^\mu}$. For $\frac{\partial x^\mu}{\partial x^{\mu'}}$ we have to compute the inverse coordinate transformation: We need the old coordinates x expressed as functions

[1] We use here indices $i', j', \ldots$ for everything related with new coordinates instead of $i, j, \ldots$ used for the old coordinates. Another nice convention if there are many different indices – one always knows which indices are related to new coordinates, and which to old ones.

[2] Mathematicians distinguish these two types of objects by different names. If we have a map $f : X \to Y$, then some object is named "covariant" if every such object on X uniquely defines one on Y, and "contravariant" if every such object on Y uniquely defines one on X. So, points are covariant, functions are contravariant.

of the new coordinates x', that means $x^\mu = x^\mu(x')$, and use these functions to compute the Jacobi matrix $\frac{\partial x^\mu}{\partial x^{\mu'}}$. It is important not to mingle these two cases.

How to remember such a formula, with all the different upper and lower indices? Fortunately the mathematicians have invented some simple rules so that one can check if all the indices are correct: *The indices which are part of some summation should be the same on both sides, and those which are upper on the left have to be upper in the right too. And if there is a summation, the index should appear exactly twice, once as an upper index and once as a lower index.*

Let's see how this applies to another interesting object, a covector, which has a lower index. The classical example of a covector is the gradient ∇f of a function. This covector has the partial derivatives of the function as its components. They are denoted $\frac{\partial}{\partial x^i} f(x)$ or shorter (mathematicians are lazy) $\partial_i f$[3]. The rules how to define which rule has to be used for the transformation is the same as before, but the result is now the following (check it!):

$$\partial_{i'} f(x') = \sum_{i=1}^{n} \frac{\partial x^i(x')}{\partial x^{i'}} \partial_i f(x(x')), \qquad \partial_i f(x) = \sum_{i'=1}^{n} \frac{\partial x^{i'}(x)}{\partial x^i} \partial_{i'} f(x'(x)).$$

Another covector field important in physics is the momentum $p_i(x)$. The reason is that its origin is also a partial derivative, namely of the Lagrangian: $p_i = \frac{\partial}{\partial v^i} L(x^i, v^i)$.

Before giving the formula for the more general case with a tensor field with a quite arbitrary number of upper and lower indices, let's mention here some additional simplifications of the formulas. First, whenever we have the same letter i used twice, for an upper as well as a lower index, we have to sum $\sum_{i=1}^n$ anyway. There is no need to write such a boring sum symbol if everybody knows anyway that it is required, so let's omit it[4]. Then, the context is usually quite clear, so that there usually is no need to write explicitly an (x) or (x') if something is a function of x or x'. With these simplifications, the formulas above look now the following way:

$$\partial_{i'} f = \frac{\partial x^i}{\partial x^{i'}} \partial_i f, \qquad \partial_i f = \frac{\partial x^{i'}}{\partial x^i} \partial_{i'} f.$$

So, now we are prepared to write down a general formula for coordinate transformations of an object named "tensor field". Tensor fields are fields which can have several upper or lower indices, say $T^{ijk}_{lm}(x)$. And a main point of using these many upper and lower indices is that they allow to memorize how they have to be transformed into new coordinates: In the new coordinates, an upper index i has to be replaced by a corresponding upper index in new coordinates, i', and one has to use the transformation matrix so that there will be one identical upper and lower index for summation. So, this gives:

$$T^{i'j'k'}_{l'm'} = \frac{\partial x^{i'}}{\partial x^i} \frac{\partial x^{j'}}{\partial x^j} \frac{\partial x^{k'}}{\partial x^k} \frac{\partial x^l}{\partial x^{l'}} \frac{\partial x^m}{\partial x^{m'}} T^{ijk}_{lm}.$$

[3]...or even shorter, given that physicists are even lazier, $f_{,i}$.

[4]There are, in fact, exceptions, cases where one index i is found two or even more times, without summation over that i. But these exceptions are so rare that in such exceptional cases one has to add the text "no summation over i", and nobody even proposed to replace such texts with some $\not\Sigma_i$ or so "no summation" symbol.

Beyond scalar fields (simply functions without indices), vector fields (one upper index), and covector fields (one lower index) the most important objects are operators (one upper and one lower index) and metrics (either two upper indices or two lower indices).

An operator a^i_j can be used to transform a vector field into another vector field: $v^i = a^i_j v^j$, and a covector field into another covector field: $p_j = a^i_j p_i$. The Jacobi matrix $\frac{\partial x^{i'}}{\partial x^i}$ is an operator we have already used for transformations of coordinates. Another operator which is often used is the trivial one, which leaves everything unchanged. Its components are defined by the Kronecker delta δ^i_j which is 1 for $i = j$ and 0 if $i \neq j$, and it follows from these definitions that $v^i = \delta^i_j v^j$ as well as $p_j = \delta^i_j p_i$. Such a symbol appears useful if one wants, for example, write down that the two Jacobi matrices $\frac{\partial x^{i'}}{\partial x^i}$ and $\frac{\partial x^i}{\partial x^{i'}}$ are inverse to each other, which can be written in the following form:

$$\delta^{i'}_{j'} = \frac{\partial x^j}{\partial x^{j'}} \frac{\partial x^{i'}}{\partial x^i} \delta^i_j.$$

The transformations of the coordinates define a group. A group is some set with a multiplication operation which is associative (that means $a \cdot (b \cdot c) = (a \cdot b) \cdot c$), has a unit e (that means, for all a we have $a \cdot e = e \cdot a = a$) and an inverse a^{-1} so that $a \cdot a^{-1} = a^{-1} \cdot a = e$. Transformations of coordinates are, in fact, one of the most important examples of groups. The multiplication is defined in a quite natural way: the product of two transformations is simply the result of successive performance of the two transformations, thus, the transformation $x^{i''}(x^{i'}(x^i))$. This makes it automatically associative. The unit is simply the trivial transformation $x^{i'} = \delta^{i'}_i x^i$ defined by the Kronecker delta. And that there exists an inverse transformation is, essentially, part of what makes it a transformation of coordinates.

2.2.2 The Euclidean metric

A metric is a symmetric tensor field with two lower indices: $g_{ij}(x) = g_{ji}(x)$. The probably most important application is the length of a curve defined by some trajectory $x^i(t)$:

$$l = \int_{t_0}^{t_1} \sqrt{g_{ij}(x(t)) \frac{dx^i}{dt} \frac{dx^j}{dt}} dt.$$

The simplest and most important example of a metric is the Euclidean metric, and it is defined by the same Kronecker delta, which is 1 for $i = j$ and 0 if $i \neq j$, except that it has now two lower indices: δ_{ij}. For the Euclidean metric, the formula for the length of a path simplifies to

$$l = \int_{t_0}^{t_1} \sqrt{\delta_{ij} \frac{dx^i}{dt} \frac{dx^j}{dt}} dt.$$

This would have been written, without all the conventions introduced above, as

$$l = \int_{t_0}^{t_1} \sqrt{\sum_{i=1}^{3} \left(\frac{dx^i}{dt} \right)^2} dt = \int_{t_0}^{t_1} |v| dt.$$

Together with the metric $g_{ij}(x)$ with lower indices, it is also useful to have the inverse matrix. It is denoted by using the same letter with upper indices: $g^{ij}(x)$. That the two matrices are inverse of each other can be written in the following way:

$$g^{ij}g_{jk} = \delta^i_k; \qquad g_{ij}g^{jk} = \delta^k_i.$$

A very important application of a metric is lowering and raising indices: If we have a vector field, with an upper index a^i, we can use the metric to define a corresponding covector field: $a_i(x) = g_{ij}(x)a^i(x)$, so that the index is now a lower index. The inverse operation is also possible, a lower index may be raised to an upper one, $a^i(x) = g^{ij}(x)a_j(x)$. For a general metric, the fields with upper and lower indices look quite different. But this simplifies in the case of the Euclidean metric: $a_i(x) = \delta_{ij}a^j(x)$ means that the fields with the upper and lower indices have the same components, thus, are essentially indistinguishable. So, in some sense, if there is an Euclidean metric, one should no longer bother about upper and lower indices? Not really. Even for the Euclidean metric, it makes sense to distinguish them, because this helps to use curved coordinates. The problem is that the Euclidean metric has the simple form $delta_{ij}$ only in Cartesian coordinates. In other, curved coordinates x' it will have a quite different form:

$$\delta_{i'j'} \neq g_{i'j'}(x') = \frac{\partial x^j}{\partial x^{j'}}\frac{\partial x^i}{\partial x^{i'}}\delta_{ij}.$$

So, if computed in the curved coordinates, the vector and covector variants of the same field look different again. The existence of a metric also allows to define a volume. In Cartesian coordinates the volume of a unit cube I^3 is defined by $V = \int_0^1\int_0^1\int_0^1 dx^1dx^2dx^3 = \int_{I^3} d^3x = 1$. In other coordinates, one has to introduce the determinant of the Jacobi matrix − the Jacobian − as a factor to the volume element: $d^nx = \left|\frac{\partial x^i}{\partial x^{i'}}\right| d^nx'$. Can we compute this factor if we know the metric? Yes, we can do this by taking the determinant of the equation above:

$$g = |g_{i'j'}(x')| = \left|\frac{\partial x^j}{\partial x^{j'}}\right|\left|\frac{\partial x^i}{\partial x^{i'}}\right||\delta_{ij}| = \left|\frac{\partial x^i}{\partial x^{i'}}\right|^2.$$

As a consequence, there is a simple coordinate-invariant form of the volume element: $\sqrt{g}d^nx$. In Cartesian coordinates, this is simply d^nx, and following the formula above, it transforms in the same way as required for the volume form.

2.2.3 The harmonic equation as an Euler-Lagrange equation

A function $\phi(x)$ is named harmonic if it is, in some sense, as smooth as possible. It appears that the most natural way to specify this is to minimize the square of the gradient $\nabla\phi(x)$, averaged over the volume. With the denotations we have introduced above, the gradient $\nabla\phi$ is a covector $\partial_i\phi$. This would be the integral

$$S = \frac{1}{2}\int |\nabla\phi(x)|^2 d^nx = \frac{1}{2}\int \delta^{ij}\partial_i\phi\partial_j\phi d^nx \tag{2.2}$$

But this is only valid in the Cartesian coordinates. In general coordinates, we would have to use the correspoding metric tensor. This gives $g^{ij}\partial_i\phi\partial_j\phi$. But this is nothing one can integrate as it is, we have to multiply it with the volume form $\sqrt{g}d^nx$. This gives:

$$S = \frac{1}{2}\int g^{ij}\partial_i\phi\partial_j\phi\ \sqrt{g}d^nx. \tag{2.3}$$

Let's look now at the equations for such minimum problems in field theories. As in general, what has to be minimized is named "action" and denoted with S. In field theory, it is an integral over the whole space of a function named Lagrange density and denoted $\mathcal{L}$. Let's consider the simplest case of a single field with at most first derivatives in the Lagrange density:

$$S = \int \mathcal{L}(\phi(x), \partial_i\phi(x), \partial_i\partial_j\phi(x), \ldots, x^i)d^nx.$$

In general, there can be many different fields, with a lot of tensor indices, and in principle with even higher than second partial derivatives. But even second derivatives are rare exceptions, usually there will be only first derivatives, and to handle many fields with indices is straightforward – doing the same for all fields and all their indices. For the simple case above, the resulting Euler-Langrange equations are

$$\frac{\delta S}{\delta\phi} = \frac{\partial\mathcal{L}}{\partial\phi(x)} - \partial_i\frac{\partial\mathcal{L}}{\partial(\partial_i\phi(x))} + \partial_i\partial_j\frac{\partial\mathcal{L}}{\partial(\partial_i\partial_j\phi(x))} - \ldots = 0. \tag{2.4}$$

If we apply this formula to (2.3), we obtain the so-called harmonic equation:

$$\partial_i\left(g^{ij}\sqrt{g}\partial_j\phi\right) = 0.$$

In the simplest case of the Euclidean metric in Cartesian coordinates, this is simply the Laplace equation:

$$\delta^{ij}\partial_i\partial_j\phi = \nabla^2\phi = \Delta\phi = 0.$$

2.2.4 What leaves the Laplace equation invariant?

The trick we have applied to the wave equation can be equally applied to the Laplace equation: If we have a solution $\Delta\phi(x) = 0$, we can try to find different solutions of this equation by using different coordinates $x'(x)$ and try to find out if $\phi'(x) = \phi(x'(x))$ is also a solution, so that $\Delta\phi'(x) = 0$ too. With the mathematics found above, we already know that $\phi'(x)$ is the same function in the other coordinates, and is a solution of the corresponding harmonic equation $\partial_{i'}\left(g^{i'j'}(x)\sqrt{g'}\partial_{j'}\phi'\right) = 0$ in these coordinates. So all what is necessary is that

$$\Delta = \partial_{i'}g^{i'j'}(x)\sqrt{g'}\partial_{j'}.$$

This would be the case if $g^{i'j'}(x) = \delta^{ij}$. That means, for an affine coordinate transformation $x^{i'} = \omega^{i'}_i x^i + t^{i'}$ we obtain the condition

$$\delta^{i'j'} = \omega^{i'}_i\omega^{j'}_j\delta^{ij}.$$

The set of all coordinate transformations which leave the Laplace equation invariant is also a group. Indeed, we already know that all coordinate transformations form a group. So, all we have to check is if the subset of those which leave the Laplace equation invariant form a subgroup. But if two transformations leave the Laplace equation invariant, then their combination will leave it invariant too. The trivial transformation leaves it invariant, and if a transformation leaves it invariant, the inverse transformation will leave it invariant too. This group is the Euclidean symmetry group, denoted $E(n)$ for the n-dimensional space, that means, in the three-dimensional space it is $E(3)$. It contains translations, rotations, and inversion. The most important subgroup is that of the rotations, $SO(n)$, resp. $SO(3)$ for the three-dimensional space.

This result is, of course, nothing unexpected: Nobody would be surprised that rotations and translations of a solution of such a symmetric equation like $\Delta\phi(x) = 0$ gives another solution of the same equation $\Delta\phi(x) = 0$.

2.3 Relativistic denotations

Up to now, we have introduced mathematics which have nothing to do with relativity – formalas useful for handling arbitrary coordinates in general, but also in the simply case of three-dimensional Euclidean geometry. Let's introduce now some specific denotations which are used in relativity.

These denotations have a purely mathematical character. From the mathematical point of view, there is no difference between the three coordinates one uses to define the position of a point in space and the time coordinate used to fix a moment of time: They are all simply real numbers used to identify events.

The main difference is that they have different units. Spatial coordinates are defined in meters, temporal coordinates in seconds. But in relativity, there is a velocity which plays a central role, the speed of light c. It appears in almost every relativistic formula. To get rid of a lot of factors c, instead of the time coordinate t a new spatial coordinate $x^0 = ct$ is used.[5]

Many relativistic formulas contain the spatial coordinates and the time coordinate in a similar way. Given that spatial coordinates are (x^1, x^2, x^3), we can use for such formulas so-called "spacetime coordinates" (x^0, x^1, x^2, x^3). But this creates a problem with our simplifications above: If we, from time to time, want to use also formulas of the normal type, with summation only over the spatial coordinates, we would have to distinguish somehow between the usual sum $\sum_{i=1}^3$ over the three spatial coordinates and the "relativistic" sum $\sum_{i=0}^3$ over the four "spacetime coordinates". If we would write the sum in the complete form, this would be no problem – we would have to write down if the sum starts from 1 for spatial coordinates or from 0 for spacetime coordinates. But we have already made the convention to omit the sum sign. What to do now? The solution is an additional convention: One will use greek letters for spacetime indices, and latin letters for spatial indices. Typically, for spatial indices one uses i, j, k, l, and for spacetime indices the most popular choices are $\mu, \nu, \kappa, \lambda$. So, one can write a function which depends on all four coordinates as $f(x^\mu)$, or as $f(x^0, x^i)$.

Let's note here: These were only denotations. They are useful to write down formulas in the relativistic domain in an easier to remember, more compact way, that's all.

[5] In the early time, it was also common to use x^4 for this, sometimes also with an imaginary unit, $x^4 = ict$, but the modern conventions are quite clear, one has to use x^0 today, and without imaginary unit.

2.3.1 The wave equation

The harmonic equation was an equation for spatial coordinates only. Let's see now how to do similar mathematics for the wave equation. First, let's use the coordinate x^0 instead of t:

$$\Box = \frac{1}{c^2}\partial_t^2 - \Delta = \partial_0^2 - \Delta.$$

In this form, the only difference between this operator and the Laplace operator in four dimensions is that some signs are different:

$$\Box = \left(\frac{\partial}{\partial x^0}\right)^2 - \left(\frac{\partial}{\partial x^1}\right)^2 - \left(\frac{\partial}{\partial x^2}\right)^2 - \left(\frac{\partial}{\partial x^3}\right)^2.$$

This gives reasons to hope that the same mathematics we have developed in section 2.2.3 for the harmonic equation can be (somehow modified to get the signs correct) applied to the wave equation too. Let's see if we can rewrite the the Lagrangian (2.2) which gives the harmonic equation so that it gives, instead, the wave equation. Once the term for the time coordinate should get another sign, let's try giving another sign to a similar term for the time coordinate:

$$S = \frac{1}{2}\int \partial_0\phi\partial_0\phi - \delta^{ij}\partial_i\phi\partial_j\phi d^n x \tag{2.5}$$

And, indeed, the Euler-Lagrange equations (2.4) for the action (2.5) give, indeed, the wave equation $\Box\phi = 0$. But this is only correct in Cartesian coordinates. How we have to rewrite this is general, possibly curved, coordinates, which possibly even mix space and time coordinates? It appears that we can use the general rules developed in section 2.2.3, because these are rules which can be applied to even much more general tensor fields $t_{ij}(x)$, not only to δ_{ij}.

Let's start with defining the replacement for the δ_{ij}. We will name it $\eta_{\mu\nu}$, with the μ,ν following the convention above going from 0 to 3, and define it in the obvious way: 1 if $\mu = \nu = 0$, -1 if $\mu = \nu \neq 0$, and 0 if $\mu \neq \nu$. The inverse matrix will be denoted, similarly, $\eta^{\mu\nu}$, and its components are obviously defined in the same way.

This matrix is named "Minkowski metric". It is not a metric in the usual sense, because a metric should be positive definite: For an arbitrary non-zero vector a^i its length as defined by this metric – the scalar product with itself – should be positive, $\langle a,a\rangle = g_{ij}a^ia^j > 0$. For the Minkowski metric, this does not hold. Thus, it is not a metric in the usual meaning of the word.

One can reasonably name it a pseudometric. And to name it this way makes some sense, because beyond this important difference, a lot of other mathematical properties of metrics hold for such pseudometrics too. Let's, for example, take a look at how to rewrite the Lagrangian for the wave equation (2.5). It now looks like

$$S = \frac{1}{2}\int \eta^{\mu\nu}\partial_\mu\phi\partial_\nu\phi d^4x \tag{2.6}$$

How will this look like in other coordinates? As in the case of the Euclidean metric, in general coordinates it does not look that nice:

$$\eta_{\mu'\nu'} \neq \gamma_{\mu'\nu'}(x') = \frac{\partial x^\nu}{\partial x^{\nu'}}\frac{\partial x^\mu}{\partial x^{\mu'}}\eta_{\mu\nu}.$$

So, in general coordinates we have to use the metric $\gamma_{\mu\nu}(x)$ instead of the simple metric $\eta_{\mu\nu}$. And, in full analogy, we also have to multiply the volume element d^4x with the determinant $\left|\frac{\partial x^i}{\partial x^{i'}}\right|$. The way how to compute it if we have a pseudometric is also similar:

$$\gamma = |\gamma_{\mu'\nu'}(x')| = \left|\frac{\partial x^\nu}{\partial x^{\nu'}}\right|\left|\frac{\partial x^\mu}{\partial x^{\mu'}}\right| |\gamma_{\mu\nu}| = -\left|\frac{\partial x^\mu}{\partial x^{\mu'}}\right|^2.$$

The only difference is that the determinant γ of the Minkowski metric is negative, so that the Jacobian is now $\sqrt{-\gamma}$. So, the Lagrangian for the wave equation in general coordinates looks like

$$S = \frac{1}{2}\int \gamma^{\mu\nu}(x)\partial_\mu\phi(x)\partial_\nu\phi(x)\sqrt{-\gamma}d^4x. \tag{2.7}$$

The Euler-Lagrange equation in these general coordinates is, then,

$$\frac{\delta S}{\delta\phi}\Box\phi = \partial_\mu\left(\gamma^{\mu\nu}(x)\sqrt{-\gamma}\partial_\nu\phi\right) = 0. \tag{2.8}$$

2.4 The Lorentz group

Let's now take a look at transformations of coordinates which leave the form of the wave equation invariant. That means, transformations which have the very special property that

$$\eta^{\mu'\nu'} = \frac{\partial x^{\nu'}}{\partial x^\nu}\frac{\partial x^{\mu'}}{\partial x^\mu}\eta^{\mu\nu}.$$

Similar to the case of the Laplace operator, these transformations have the form of an affine transformation:

$$x^{\mu'} = \Lambda^{\mu'}_\mu x^\mu + t^{\mu'},$$

which combine an arbitrary shift defined by $t^{\mu'}$ with a linear transformation $\Lambda^{\mu'}_\mu$ which has to fulfill the following property:

$$\eta^{\mu'\nu'} = \Lambda^{\mu'}_\mu\Lambda^{\nu'}_\nu\eta^{\mu\nu}. \tag{2.9}$$

The transformations which have this property define a group.

2.4.1 An important difference: Active vs. passive coordinate transformations

The mathematics of coordinate transformations can be used in two from a physical point of view very different ways.

The first way is named **passive** or **alias** coordinate transformations. In this case, one and the same physical solution is described in different ways, using different coordinates. This is nothing but an application of pure mathematics, without any physical importance. If the mathematics are used correctly, it does not matter at all which system of coordinates you use to describe the solution. Your choice of coordinates is arbitrary. You can check this, by trying to compute something measurable using different coordinates. The final result should be the same. If not, you have made a mathematical error. But you are not obliged to do such things at all. You can, as well, choose one system of coordinates forever, and refuse even to look at any other one, but you are nonetheless able to compute everything physical using only that single system of coordinates.

The second way is named **active** or **alibi** coordinate transformation. In this case, the coordinates remain unchanged. What changes is the solution. This is what we have used here to obtain a new, Doppler-shifted solution of the wave equation out of a given one.

Note the difference: We could have used any other coordinate transformation to get another description of the same solution. If that other coordinate transformation would not be a Lorentz transformation, we would have to rewrite the same wave equation in these other coordinates. It would have been a different-looking equation, but the different-looking solution of the different-looking equation would have been simply another, equally valid, description of the same solution of the same equation.

Instead, a different coordinate transformation would not have allowed us to construct a new, different solution of the same equation. For this trick, it is essential that the transformed equation looks, by accident, like the original equation. In general, the tranformation would have given us some solution of some other equation, something in no way useful to study the solutions of the original wave equation. But in our case, the Lorentz transformation has given us another, new solution, with different properties, of the same wave equation.

Could we have used the Lorentz transformations to describe the same wave equation in other coordinates? Yes, of course. But these other coordinates would have been a quite unnatural, strange choice. Of course, one is free to use whatever coordinates one likes. But, in the case of water waves, would be the point of using a "time coordinate" which have nothing to do with real time?

So, once we have understood this important difference between active (alibi) and passive (alias) transformations, we have found a useful application of the Lorentz transformations as active transformations, transformations which create new, different solutions of the same wave equation.

2.5 Field theories where Lorentz transformations allow to find new solutions

Once the Lorentz transformations allow to create new solutions of the wave equation, another question appears: Are there other interesting equations with the same property, so that we can use Lorentz transformations to find new, different solutions? The answer is positive. There are a lot of other, more complex but very interesting equations where the Lorentz transformations allow to find new solutions.

2.5.1 Mass terms

First of all, one can add a mass term. The Lagrangian would be, with a mass term, the following:

$$S = \frac{1}{2} \int \eta^{\mu\nu} \partial_\mu \phi \partial_\nu \phi + m^2 \phi^2 d^4x \qquad (2.10)$$

This gives a Klein-Gordon equation, which describes a massive scalar particle:

$$\Box u(x^\mu) + m^2 u(x^\mu) = 0. \tag{2.11}$$

This equation is also linear. The solutions of this equation are waves, but their velocity is already lower than c.

2.5.2 Interaction terms between different fields

There may be different fields, let's denote them ϕ^k with $1 \le k \le K$, which interact with each other. The interaction will be defined by some potential term $V(\phi^1(x), \ldots, \phi^K(x))$, which is using only the values of the various functions at the same point x. In this case, a Lorentz transformation will not change the interaction term.

$$S = \frac{1}{2} \int \sum_k \eta^{\mu\nu} \partial_\mu \phi^k \partial_\nu \phi^k + V(\phi^1, \ldots, \phi^K) d^4x \tag{2.12}$$

This gives K different Euler-Lagrange equations, one for each ϕ^k, which have the following form:

$$\Box \phi^k(x) + \frac{\partial}{\partial \phi^k} V(\phi^1(x), \ldots, \phi^K(x)) = 0.$$

In principle, the property that one can apply Lorentz transformation to obtain new solutions holds even to more general interaction terms, with different, unrelated funtions $V^k(\phi^1, \ldots, \phi^K)$ instead of derivatives of the same function $V(\phi^1, \ldots, \phi^K)$. All what we need for this property is that the interaction terms are pointwise functions, thus, depend only on the function values of all the $\phi^k(x)$ at the same point x.

In general, interaction terms make the system of equations itself nonlinear. This makes it usually difficult, usually close to impossible, to find exact solutions. But, despite this, the basic property of the Lorentz transformation remains – if we have an exact solution, the Lorentz transformation creates a new, different solution of the same equation. And it can be applied to approximate solutions too, even if one has to care about the problem that the data which describe the accuracy of the transformation have to be transformed too.

2.5.3 The electromagnetic field

Up to now, the fields we have considered were scalar fields. For scalar fields, the situation is easier, because their transformation rule itself is trivial. Let's now consider some other fields, fields which transform themselves in a nontrivial way if we change the coordinates.

The most important example in our life is the electromagnetic field. The simplest way to describe it is to use its potential, which is a covector field $A_\mu(x)$. The EM field acts as a force on other fields. Different from the case considered above, an arbitrary local reaction term will not be left completely unchanged by a coordinate transformation, because it contains the fields $A_\mu(x)$, which transform in a non-trivial way, as a covector.

Fortunately, the interaction terms with the EM field have a very special form in field theory. Namely, one has to replace all appearences of the partial derivatives ∂_μ of a charged field by ($D_\mu = \partial_\mu - iqA_\mu(x)$), an operation named "gauge covariant derivative".

Now, these gauge covariant derivatives have the same transformation behavior as the original partial derivatives. Thus, if the original terms remain invariant for some coordinate transformations, the same terms with the gauge covariant derivatives will have the same invariance property. Thus, we should not be afraid of the interaction terms.

It remains to study the equations of the EM field itself. Unfortunately, the Maxwell equations for the EM field are formulated in terms of the electric and magnetic fields, which are derivatives of the potentials $A_\mu(x)$, and these potentials are not uniquely defined by the fields themselves. This leaves some freedom in the definition of the potentials. This problem can be handled by fixing the potential using a so-called gauge condition. Only with such a gauge condition, one can obtain an equation for the potential $A_\mu(x)$ from the Maxwell equations for the field strengths. Fortunately, there exists a very beautiful gauge condition, which has a form where it is easy to show that the equations have the invariance property we are interested in. It is the so-called Lorenz gauge[6]. The Lorenz gauge looks more natural in terms of the potentials with upper index $A^\mu(x)$, which can be obtained by raising the index with the Minkowski metric: $A^\mu(x) = \eta^{\mu\nu}A_\nu(x)$. Then, the Lorenz gauge is

$$\partial_\mu A^\mu(x) = 0. \tag{2.13}$$

For this gauge condition, the other equations of the EM field are equivalent to an equation we already know – the wave equation for each of the components.

$$\Box A^\mu(x) = 0. \tag{2.14}$$

For both equations it is now easy to see that they remain unchanged if we change the coordinates using a Lorentz transformation $x^{\mu'} = \Lambda^{\mu'}_\mu x^\mu + c^{\mu'}$. Indeed,

$$\begin{aligned}\partial_{\mu'} A^{\mu'}(x') &= \frac{\partial x^\mu}{\partial x^{\mu'}}\partial_\mu\left(\frac{\partial x^{\mu'}}{\partial x^\nu}A^\nu(x'(x))\right) = (\Lambda^{-1})^\mu_{\mu'}\partial_\mu(\Lambda^{\mu'}_\nu A^\nu(x))\\ &= (\Lambda^{-1})^\mu_{\mu'}\Lambda^{\mu'}_\nu\partial_\mu A^\nu = \delta^\mu_\nu\partial_\mu A^\nu(x) = \partial_\mu A^\mu(x).\end{aligned}$$

Given that we already know that for the wave operator we have $\Box' = \Box$, we also have

$$\Box' A^{\mu'}(x') = \Box'\left(\frac{\partial x^{\mu'}}{\partial x^\mu}A^\mu(x'(x))\right) = \Box(\Lambda^{\mu'}_\mu A^\mu(x)) = \Lambda^{\mu'}_\mu \Box A^\mu(x).$$

And, given that $\Lambda^{\mu'}_\mu$ is a non-degenerated matrix, it follows that $\Box' A^{\mu'}(x') = 0$ if and only if $\Box A^\mu(x) = 0$.

2.5.4 Gauge fields

The electromagnetic field is the simplest example of a much larger class of fields, namely gauge fields. The weak and the strong force of the standard model of particle physics are examples of such gauge fields. The difference to the electromagnetic field is that the fields $A^\mu(x)$ are not simply real fields, but elements of some Lie algebra.

[to be completed]

[6] Note that the "Lorenz" is not a spelling error, it is not named after Hendrik Lorentz as many other things, but after another physicist, Ludvig Lorenz.

2.5.5 The Dirac equation

The Dirac equation is another important equation in modern physics. It is the equation used to describe fermionic fields. It is an equation for a system of four complex (or, equivalently, eight real) fields. The equation was obtained by Dirac in an attempt to take a square root out of the Klein-Gordon equation. But the Klein-Gordon equation is already among the equations for which the Lorentz transformation allows to construct new solutions. So, it is not really a surprise that this works for the Dirac equation too.

In relativistic notations, the equation looks like

$$i\gamma^{\mu}\partial_{\mu}\psi(x^{\mu}) = m\psi(x^{\mu}).$$

The square of the operator $\gamma^{\mu}\partial_{\mu}$ is, by construction of the matrices γ^{μ}, the Laplace operator $\Box$. So, it follows from this equation that $-\Box\psi = m^2\psi$, thus, the equation is a square root of the Klein-Gorden equation.

It also follows immediately that the Lorentz-transformed equation fulfills the same basic property, namely that its square gives the Klein-Gordon equation. Unfortunately, this is nonetheless not exactly the same equation, but a different representation. Fortunately, all these different representations are equivalent, thus, one can find a transformation U so that $\psi'(x^{\mu}) = U\psi(x^{\mu})$ is already again a solution of the Dirac equation in its original representation.

The subtle point is that this operator U is not uniquely defined. The operator -U would do it too.

[to be completed]

2.5.6 And, again, various interaction terms

What has been said about pointwise interaction terms for the scalar wave equation holds also for all the other equations considered - it is possible to add various pointwise interaction terms. The freedom of choice is somewhat restricted, not completely without any restrictions (except for containing no derivatives) as in the scalar case - as the EM field, as the Dirac operator follow some transformation laws, and these transformation laws have to fit each other.

The interaction terms which are important in the standard model of particle physics are:

1.) Non-abelian gauge fields: This is a generalization of the EM field, and formally looks like several such EM fields which additionally interact with each other. 2.) The interaction of these gauge fields with Dirac fermions. The interaction term has a quite special form, namely a replacement of the partial derivative by an additional term:

$$i\partial_{\mu} \to i\partial_{\mu} + gA_{\mu},$$

with the charge of this fermion field being g.

[to be completed]

2.5.7 The equation of the Standard Model of particle physics

So, we have found a lot of wave equations which all share the same nice property: If we have a solution of these equations, we can, using a Lorentz transformation, create new, different solutions of these equations.

How important is this class of wave-like equations? The surprising news is that all the fundamental fields, all the fields used in the Standard Model (SM) of modern particle physics, are described by such equations. And they all share the same constant c - the speed of light. In principle, this SM can be considered as a single big equation, let's name it the SM equation, containing many different parts, which interact with each other. But all these parts, and all the interaction terms, fit into the list of equations above, as well as the reaction terms between these parts. So that the SM equation is also of this type. If we have one solution of the SM equation, we can apply a Lorentz transformation, and will obtain another, different solution of the SM equation, a solution which will be, in comparison with the original solution, Doppler-shifted.

2.6 Relativistic effects

What is the consequence of the fact that all fields, the whole SM, follow an equation where we can apply a Lorentz transformation to obtain a new solution of the same equation? Some surprising results about the behavior of clocks follow.

2.6.1 Time dilation

Let's construct, out of what we have, namely of of things described by the SM, a clock. This clock will be described by some trajectory, with some numbers on it which denote the time measured at this event. Let's simply assume, as an example, that this trajectory is, for the initial solution with the clock at rest, a line with the spatial coordinates (0,0,0) and with a result 0 at time 0 and result 1 at time 1.

What happens now with this solution, if we apply some Lorentz transformation? We obtain another solution. This other solution describes a clock of the same construction, but moving. The Lorentz transformation is linear, thus the point $(x^\mu) = (0,0,0,0)$ remains the same. But the point where the clock shows 1, which was originally at $(x^\mu) = (1,0,0,0)$, will now be in $(x^{\mu'}) = (\gamma, -\frac{v}{c}\gamma, 0, 0)$. The line remains a line, so that the clock is moving with the speed v (which is $\frac{v}{c}$ in terms of the time coordinate $x^0 = ct$. But the clock shows the result 1 at real time γ instead of real time 1. But we have

$$\gamma = \sqrt{1 - \frac{v^2}{c^2}}^{-1} > 1$$

for every $0 < |v| < c$. That means, the moving clock shows clock time 1 only at real time $\gamma > 1$, thus, is dilated.

2.6.2 Length contraction

What happens with a ruler? The ruler at rest is some solution of the SM, but we cannot idealize it as a single point, we need two points of it. The begin we put, again, at $(x^\mu) = (0,0,0,0)$, the mark of the ruler with length 1 will, if we measure the x-direction, be at $(x^\mu) = (0,1,0,0)$. At rest, the begin moves toward $(x^\mu) = (1,0,0,0)$ and the mark with the 1 toward $(x^\mu) = (1,1,0,0)$.

We apply the same procedure and obtain a solution of the same type, thus, also a ruler, but moving. The four resulting points will be the following: The begin will be, like for the clock, moving along the line from $(x^{\mu'}) = (0,0,0,0)$ to $(x^{\mu'}) = (\gamma, -\frac{v}{c}\gamma, 0, 0)$. The point with mark 1 will, instead, move from $(x^{\mu'}) = (-\gamma\frac{v}{c}, \gamma, 0, 0)$ to $(x^{\mu'}) = (\gamma(1-\frac{v}{c}), \gamma(1-\frac{v}{c}), 0, 0)$. Where is the mark 1 of this moving ruler at $t = 0$? We have to compute where the line between the two points intersects the line $t' = 0$. The result is $(x^{\mu'}) = (0, \gamma(1-\frac{v^2}{c^2}), 0, 0)$. Thus, the mark 1 of the moving ruler will be, at time $t = 0$, at a distance $\gamma^{-1} < 1$ from the origin. The moving ruler is shorter.

2.6.3 Impossibility to identify global contemporaneity

A similar fate waits for devices which allow to measure absolute contemporaneity. Such a device would tell us that the two events $(x^\mu) = (0,0,0,0)$ and $(x^\mu) = (0,1,0,0)$ have happened at the same moment of time. But now we apply a Lorentz transformation to this device, and what is the result? The same device, only in a moving state, claims that $(x^\mu) = (0,0,0,0)$ and $(x^\mu) = (-\gamma\frac{v}{c}, \gamma, 0, 0)$ have happened at the same time.

2.6.4 Impossibility to measure absolute rest

And now it is already clear what happens if we try to measure, with some physical device, what is absolute rest. Suppose we have such a device. This device measures that the line from $(x^\mu) = (0,0,0,0)$ to $(x^\mu) = (1,0,0,0)$ is at absolute rest. We do the same trick again, and apply the Lorentz transformation to this measurement of absolute rest. What do we obtain? A solution which describes the same measurement device, which claims that the line from $(x^\mu) = (0,0,0,0)$ to $(x^\mu) = (\gamma, -\frac{v}{c}\gamma, 0, 0)$ is at absolute rest. Which is, of course, wrong.

The consequence is the **relativity principle**: Once all our physical equations allow the application of Lorentz transformations to obtain a new solution of our equations, only moving relative to the original solution, there cannot exist a physical device which identifies the own state of movement.

2.7 The Lorentz ether vs. Minkowski spacetime

What follows from the relativity principle?

2.7.1 The Lorentz ether

In the Lorentz ether interpretation, nothing follows. Wave equations appear, most natural, as a sort of sound wave equations of some material. Once we observe something nicely described by wave equations, the reasonable hypothesis is that they are all similar to sound waves of some ether.

The relativity principle is, then, only an unfortunate consequence of the fact that all what we can actually observe are sound waves of the same ether. For usual condensed matter, there exist other things, like light rays, which are not sound waves of this condensed matter. With these additional possibilities accessible to use, the relativity principle would become invalid, and we would be able to measure as absolute rest, as absolute contemporaneity, and measure absolute lengths and time. Without them, we are simply unable to distinguish things which, in reality, are different.

In the Lorentz ether, time dilation and length contraction are distortions of our measurement devices caused by their motion relative to the ether. True distances and true time could be measured - if our measurement devices would be, by accident, really at rest. Once we cannot know if they are really at rest, one cannot be sure that our measurements are undistorted by the ether.

2.7.2 The Minkowski spacetime

There is another interpretation of special relativity - the spacetime interpretation proposed by Minkowski. It is the one accepted by the mainstream of physics. At this place I have to admit that, given that I'm not a proponent of the spacetime interpretation, the argumentation in its favor may be suboptimal.

From point of view of the spacetime interpretation, once there is no possibility, by any measurement, to distinguish if we are at rest, then no such animal like "absolute rest" exists in reality. The same for absolute contemporaneity: Once there is no possibility to establish it, uniquely, by measurement, there is no absolute contemporaneity in our real world.

This is, of course, in sharp contradiction with classical common sense. Common sense has no problem with the mathematical possibility of a spacetime. But this would be nothing but a simply a collection of the states of space, in various moments of absolute time. It would not be a description of what exists, but of all what has existed in the past, exists now, and will exist in the future. All this would have to be taken together, and given a common status of existence, without any reference of when it existed. This would be, in fact, a new concept of timeless existence. There is no more any difference between events which have happened and those which will happen, they all simply exist.

In this four-dimensional world there would be no present at all. What is the present which we experience? It can be only something particular, derivative, restricted to the particular trajectory in this spacetime which describes our own lifestream, our worldline, which contains all the events of our own life, from our birth to our death.

Chapter 3

Relativistic Gravity

Special relativity in itself does not describe gravity. Moreover, initial attempts to develop a special-relativistic variant of Newtonian gravity in a straightforward way, in particular the attempt made by Poincare in his 1905 paper, have failed. So, the incorporation of gravity requires some more serious considerations.

It is interesting to consider what the two interpretations of relativity suggest about the ways how the theory could be modified to incorporate the gravitational field.

Special Relativity, in the spacetime interpretation, does not suggest any modifications of the basic principles. Once the Lorentz symmetry is, according to the spacetime interpretation, a fundamental symmetry of nature, it follows that the gravitational field also has to be described in agreement with this symmetry. This is, essentially, what has been proposed by Poincare 1905 and what has failed. One may reject the spacetime interpretation as some artificial, mystical object, acceptable only to fatalists who think that our future is predefined, but this does not change the fact that it is, in itself, a consistent theory, which does not have any internal problems, which would be a starting point for developing general relativity.

One may object that Einstein, developing general relativity, was following philosophical ideas which were quite similar to those he used to develop special relativity. So, in both cases, some equivalence principle played a central role, and, in general, relativism and observer-dependence of many things. But to follow Einstein may not be a good idea. The ideas Einstein has used to find general relativity were in part simply wrong in the final theory (Mach's principle is wrong in GR), in part only approximately true (a non-trivial gravitational field has non-zero curvature, while pure acceleration gives zero curvature, so that there is no exact equvialence between gravity and acceleration in GR), in part misguided (Kretschmann's objection: every physical theory can be presented in a general covariant form, thus, this cannot be a physical property of GR).

And, in fact, the resulting theory is qualitatively very different from special relativity. In particular, the central, distinguishing property of GR is background independence, its completely local character. It does not contain any global object. Special relativity is, in this relation, much closer to Newtonian theory. All what has been done was to combine two absolute global objects of Newtonian theory – absolute space and absolute time – into one bigger absolute global object, namely the Minkowski spacetime. As Newtonian space and time, as the Minkowski spacetime define a predefined, fixed stage, which is not influenced in any way by matter.

So, if we look at special relativity and general relativity, we see two conceptually very different theories. There is no natural connection between them. And there is, first of all, no open problem in SR itself so that its solution would lead us to GR.

So, essentially one of the simplest ways to introduce GR is simply to define it as it is, as a covariant, background-independent metric theory of gravity.

3.1 Toward a compressible and instationary Lorentz ether

Surprisingly, this is different for the Lorentz ether.

One could think about an interpretation slightly different from the Lorentz ether, let's name it Lorentz spacetime, which would be equally sterile as the Minkowski spacetime. It would be simply the Minkowski spacetime with an unobservable preferred frame. This preferred frame would define the global notion of "now", but would not have any meaning in terms of some ether. There would be no ether, so that the wave equation would describe sound waves of the ether. There would be simply some universal abstract absolute wave operator $\square = c^{-2}\partial_t^2 - \Delta$ without further physical interpretation. This Lorentzian spacetime would have absolute space and absolute time, and no Galilean invariance, thus, conceptually even closer to Newton's spacetime. Nothing would lead from this interpretation toward another theory.

The Lorentz ether is different. The Lorentz ether assumes that this Newtonian background space is filled with some ether, and that the wave equations describe various sound waves of the ether. Now, this attempt to explain the universal wave equation as some physical equation, analogical to those we already know from condensed matter theory, leads immediately to some problems, inconsistencies of the explanation, which tell us that the equations for the ether as given in special relativity can be only an approximation. Indeed, the ether is assumed to be ideally rigid, incompressible, homogeneous and static. But, of course, a wave of the ether defines some compression, some deformation, some inhomogeneity, some change in time. So, there has to be some more fundamental theory, which defines how the Lorentz ether deformes and changes in time.

So, the Lorentz ether contains, in itself, acceptance that the theory as it is is only an approximation of a different, more complex theory, which contains an inhomogeneous, instationary, compressible ether.

At a first look, this seems to be not much information, not enough to hope that this allows to extract sufficient information to derive some fundamental ether theory. Surprisingly, it appears quite sufficient. There is a quite straightforward way from what we have found here toward that more fundamental ether theory, and the resulting ether theory of gravity appears to be very close to GR, so close that we obtain the Einstein equations of GR in a natural limit of this ether theory.

3.1.1 Continuity and Euler equations

The natural way to start to look for a dynamic ether theory is to use standard equations of standard condensed matter theory. The most general equations are the continuity and Euler equations. Let's denote the Cartesian coordinates of the Newtonian background space and Newtonian absolute time with $\mathfrak{x}^i, \mathfrak{t}$. The ether would be described by its density $\rho(\mathfrak{x}^i, \mathfrak{t})$, its velocity $v^i(\mathfrak{x}^i, \mathfrak{t})$, and in the most general case by a stress tensor $\sigma^{ij}(\mathfrak{x}^i, \mathfrak{t})$ instead of a scalar pressure. The continuity equation would be:

$$\partial_{\mathfrak{t}}\rho + \partial_{\mathfrak{x}^i}(\rho v^i) = 0, \tag{3.1}$$

and the Euler equations:

$$\partial_{\mathfrak{t}}(\rho v^j) + \partial_{\mathfrak{x}^i}(\rho v^i v^j - \sigma^{ij}) = 0. \tag{3.2}$$

Now one can rewrite these condensed matter theory equations in four-dimensional denotations, introducing the four-dimensional fields $\mathfrak{g}^{00} = \rho, \mathfrak{g}^{0i} = \rho v^i, \mathfrak{g}^{ij} = \rho v^i v^j - \sigma^{ij}$. Then, the continuity and Euler equations obtain the form

$$\partial_{\mathfrak{x}^\mu}\mathfrak{g}^{\mu\nu}(\mathfrak{x}) = 0.$$

The next interesting observation is that, if we introduce the four-dimensional wave operator

$$\Box = \partial_{\mathfrak{x}^\mu}\mathfrak{g}^{\mu\lambda}(\mathfrak{x})\partial_{\mathfrak{x}^\lambda},$$

then the equations can be rewritten as a simple wave equation for the four preferred coordinates $\mathfrak{x}^\nu$:

$$\partial_{\mathfrak{x}^\mu}\mathfrak{g}^{\mu\nu}(\mathfrak{x}) = \Box\mathfrak{x}^\nu = 0.$$

In particular, the continuity equation is equivalent to $\Box\mathfrak{t} = 0$, and the Euler equations to $\Box\mathfrak{x}^i = 0$.

Instead of the densities $\mathfrak{g}^{\mu\nu}(x)$ one can also use a pseudometric $g^{\mu\nu}(x)$ defined by $\mathfrak{g}^{\mu\nu} = g^{\mu\nu}\sqrt{-g}$. Then the operator $\Box = \partial_{\mathfrak{x}^\mu}g^{\mu\nu}(\mathfrak{x})\sqrt{-g}\partial_\nu$ is an invariant operator defined by the metric $g^{\mu\nu}(x)$. So, we obtain a set of four equations known as the harmonic condition:

$$\Box\mathfrak{x}^\alpha(x) = \partial_\mu\mathfrak{g}^{\mu\alpha}(x) = \partial_\mu(g^{\mu\alpha}(x)\sqrt{-g}) = 0. \tag{3.3}$$

What have we found up to now? A classical condensed matter theory, which follows continuity and Euler equations, can be described by a pseudometric $g^{\mu\nu}(x)$ so that the preferred Newtonian coordinates are harmonic coordinates of this metric.

3.1.2 A Lagrange formalism for the Lorentz ether

Let's now look for a Lagrange formalism for the Lorentz ether. In general, condensed matter theories may not have a Lagrange formalism. The problem is that in many condensed matter equations thermodynamic effects, diffusion, dissipation, friction and other irreversible things play a role, which the Langrange formalism usually gives equations which are reversible in time. So, one can reasonably hope for a nice Lagrange formalism only for some ideally elastic materials, where irreversible thermodynamic processes play no role. But let's assume the Lorentz ether has such ideal properties. To describe the Lorentz ether, we will use, instead of the ρ, v^i, σ^{ij}, the pseudometric $g^{\mu\nu}(x)$ together with the preferred coordinates of the Newtonian background, with $\mathfrak{t} = \mathfrak{x}^0$ as Newtonian absolute time and $\mathfrak{x}^i$ as the Cartesian coordinates for the Newtonian absolute space, which we will use as four scalar functions $x^\alpha(x)$ on the background. But there will be also other material properties of the ether, let's denote them $\varphi^m(x)$. The m can be interpreted also as denoting "matter fields" instead of "material properties of the ether". Then the Lagrangian for our generalized Lorentz ether would have the form

$$S_{GLE} = \int \mathcal{L}_{GLE}(g_{\mu\nu}, \varphi^m, \mathfrak{x}^\mu) d^4x.$$

What we certainly need, by construction, are the four harmonic conditions $\square \mathfrak{x}^\alpha(x) = 0$. Fortunately, we already know a way how to obtain the harmonic equation for a scalar function $\square \mathfrak{x}^\alpha(x) = 0$ we know, the action (2.7) for the scalar field $\phi(x)$ in general coordinates, which gives $\frac{\delta S}{\delta \phi} = \square \phi$, can be used also for a general pseudometric. Using this type of construction, we can construct a Lagrangian which gives the conservation laws as the Euler-Lagrange equations of the preferred coordinates $\mathfrak{x}^\alpha(x)$. Say, the Lagrangian

$$S_{harm} = \frac{1}{2} \int \Xi_{\alpha\beta} g^{\mu\nu} \partial_\mu x^\alpha \partial_\nu \mathfrak{x}^\beta \sqrt{-g} d^4x$$

would give as the Euler-Lagrange equations for the preferred coordinates $\mathfrak{x}^\alpha(x)$:

$$\frac{\delta}{\delta \mathfrak{x}^\alpha} S_{harm} = \Xi_{\alpha\beta} \square \mathfrak{x}^\beta(x) = 0.$$

3.1.3 The Noether theorem

This is the time to remember an old theorem of Emmy Noether. It proves that if there is a translational symmetry in time in the Lagrange formalims, this gives a conservation law of energy. Moreover, if there is a translational symmetry in space, this gives a corresponding conservation law for momentum. And, more general, a symmetry of the Lagrangian leads to a conservation law.

There is no need to look at the proof as given by Emmy Noether, because in our particular case the proof becomes exceptionally simple. Note that our proposed Lagrangian has a quite special form – it looks formally covariant, given that it has the same form in arbitrary coordinates. We have, indeed, written down this Langrangian in an explicit covariant form, in general coordinates x. On the other hand, it depends on the preferred Newtonian coordinates. But not in the usual way, using an expression in these particular coordinates, which would be different in other coordinates. No, the dependence on the preferred coordinates has been made explicit – via the special functions $\mathfrak{x}^\alpha(x)$. This makes it easy to find out if there is a translational symmetry – in this case, the Lagrangian should have the same value if we add constants $\mathfrak{x}^\alpha(x) \to \mathfrak{x}^\alpha(x) + c^\alpha$. But that means, the Lagrangian cannot depend on the $\mathfrak{x}^\alpha(x)$ themselves, but can depend only on its various partial derivatives.

But now look at (2.4) and you will see that the only part of $\frac{\delta S}{\delta \phi}$ which is not under some partial derivative is $\frac{\partial L}{\partial \phi}$ itself. Thus, once this term is zero for the preferred Newtonian coordiantes $\mathfrak{x}^\alpha$, the Euler-Langrange equation for the $\mathfrak{x}^\alpha$ automatically has the form of a conservation law $\partial_\mu(\ldots)^\mu = 0$. So, if the Lagrangian is presented in a form where all the dependence on the preferred coordinates is an explicit dependence on the functions $\mathfrak{x}^\alpha(x)$ and their derivatives, then the Noether conservation laws are simply the Euler-Lagrange equations of the preferred coordinates. Thus, from translational symmetry $\mathfrak{x}^\alpha(x) \to \mathfrak{x}^\alpha(x) + c^\alpha$ it follows that

$$\frac{\delta}{\delta \mathfrak{x}^\alpha} S_{GLE} = \partial_\mu T^\mu_\alpha(x). \tag{3.4}$$

for the energy-momentum tensor $T^\mu_\alpha(x)$ of the theory. All what we have to do now is the identification of this general energy-momentum tensor $T^\mu_\alpha(x)$ obtained via our variant of the Noether theorem with the components of the gravitational field $\mathfrak{g}^{\mu\nu}(x) = g^{\mu\nu}\sqrt{-g}$, which give the continuity and Euler equations of , in the form $T^\mu_\alpha(x) = \Xi_{\alpha\beta} g^{\mu\beta}(x)\sqrt{-g}$ for some constants $\Xi_{\alpha\beta}$.

We also conclude that the full energy-momentum tensor of the Lorentz ether has to be $\Xi_{\alpha\beta}\mathfrak{g}^{\mu\beta}(x)$. So, in some sense we do not even need the continuity and Euler equations – we can as well start with the Noether law (3.4) and then introduce the pseudometric as defined by the energy-momentum tensor, $T^\mu_\alpha = \Xi_{\alpha\beta}\mathfrak{g}^{\mu\beta}$.

3.1.4 The derivation of the general Lagrangian

So, following the Noether theorem, we have established the following requirement:

$$\frac{\delta}{\delta \mathfrak{x}^\alpha} S_{GLE} = \Xi_{\alpha\beta}\Box \mathfrak{x}^\beta(x) = 0.$$

But that simply means that

$$\frac{\delta(S_{GLE} - S_{harm})}{\delta x^\alpha} = 0.$$

That means, the difference should not depend in any way on the preferred coordinates – it should be covariant.

3.1.5 The Lagrangian of general relativity

Covariance is the most important, distinguishing property of the Lagrangian of General Relativity (GR), as for the Lagrangian of the gravitational field itself, which is

$$L_{GR} = \int R\sqrt{-g} - \Lambda\sqrt{-g} d^4x,$$

as well as of the Lagrangians of the various matter fields of GR. That means, we can identify the difference $S_{gen} - S_{harm}$ with the most general action of general relativity.

There are, of course, more general possibilities to define covariant expressions, in particular any function $f(R)$ of the scalar curvature would be covariant, as well as, say, other possibilities to obtain scalar expressions out of the curvature tensor like $R^{\mu\nu}R_{\mu\nu}$. Here, the Lagrangian of GR is simply a restriction to the lowest order terms, which makes sense, because these are the terms most relevant in a large distance approximation. This restriction to lowest order terms makes sense for the Lorentz ether theory too. Thus, we have obtained now, as the natural Lagrangian for a Lorentz ether, the Lagrangian of GR, together with a simple additional term which enforces a preference for harmonic coordinates.

$$S_{GLE} = S_{harm}(g_{\mu\nu}, \mathfrak{x}^{\mu}) + S_{GR}(g_{\mu\nu}) + S_{matter}(g_{\mu\nu}, \varphi^m). \tag{3.5}$$

3.2 Basic properties of the Lorentz ether

The general Lorentz ether we have derived above is obviously a theory different from GR, given that it has a different Lagrangian, and, as a consequence, different equations. Nonetheless, it shares some key properties with GR.

3.2.1 The Einstein Equivalence Principle

The first remarkable observation is what the additional term does not influence at all – namely the equations for the matter fields. Given that the harmonic part of the Lagrangian does not depend on the matter fields $\varphi^m(x)$ at all, we have

$$\frac{\delta S}{\delta \varphi^m} = \frac{\delta S_{matter}}{\delta \varphi^m}.$$

Thus, the equations for the matter fields are the same as in GR. And that means that the equations for the matter fields do not depend on the preferred coordinates, that means, the Einstein Equivalence Principle (EEP) holds exactly for the Lorentz ether.

The EEP includes the Weak Equivalence Principle, that the local effects of motion in a curved spacetime (gravitation) are indistinguishable from those of an accelerated observer in flat spacetime, without exception. Moreover, the outcome of any local non-gravitational experiment in a freely falling laboratory is independent of the velocity of the laboratory and its location in spacetime.

But the EEP is restricted to non-gravitational experiments. The Strong Equivalence Principle (SEP), which covers also gravitational experiments, does not hold: The equations for the gravitational field contain terms which depend on the preferred coordinates.

These results are consequences of the "action equals reaction" principle, which is a consequence of the Lagrange formalism. Formally, the "action equals reaction" principle is simply the consequence that the functional derivatives do not depend on the order:

$$\frac{\delta}{\delta u}\frac{\delta}{\delta v}S(u,v) = \frac{\delta}{\delta v}\frac{\delta}{\delta u}S(u,v).$$

So, if the Euler-Lagrange equation for a variable u depends on v, then the Euler-Lagrange equation for v also depends on u, and both dependencies are equal.

This can be applied here too. The equation for the preferred coordinates, which is the harmonic equation, depends only on the gravitational field – this is how we have defined the gravitational field in this theory, namely as the field which defines the coefficients $\mathfrak{g}^{\mu\nu}(\mathfrak{x},\mathfrak{t})$ of the harmonic equation. It follows, that only the equations of the gravitational field, but not the equations for matter fields, can depend on the preferred coordinates. So, the EEP appears in the

Lorentz ether as a natural consequence of the "action equals reaction" principle of the Lagrange formalism.

3.2.2 The GR limit

The Lagrangian also shows that there is a natural limit, namely $\Xi_{\alpha\beta} \to 0$, which gives the Einstein equations of general relativity (GR).

Why is it justified to name this limit "natural"? The point is that the additional terms share an important property with Einstein's cosmological term – they do not depend on derivatives of the metric. This has the consequence that they become more important on the cosmological scale. If one, instead, considers non-cosmological situations, the local variations, which depend on the derivatives of the metric, become more important.

Let's note that the limit leads to some non-trivial modifications of the theory itself: The symmetry of the Lagrange formalism changes. For $\Xi_{\alpha\beta} \neq 0$, the Lagrangian depends on the preferred coordinates, for $\Xi_{\alpha\beta} \neq 0$ there is no longer any such dependence.

That means, in the limit we also have no longer an Euler-Lagrange equation for the preferred coordinates, $\frac{\delta S}{\delta \mathfrak{x}^\alpha} = 0$.

On the other hand, if one considers the limit not in the Lagrange formalism, but in the equations, the situation looks different. While the harmonic condition, as obtained as Euler-Lagrange equations for the preferred coordinates, contain the factors $\Xi_{\alpha\beta}$, they are only an irrelevant factor, and one can simply multiply the equations with the inverse matrix $\Xi^{\alpha\beta}$ to obtain the harmonic equations in the form $\square \mathfrak{x}^\alpha = \partial_\mu \mathfrak{g}^{\mu\alpha} = 0$ which does not contain these constants. Once they are not changed at all during the limit process, they remain in the limit of the equations too.

3.3 The spacetime interpretation

A situation where different physical states – like the Lorentz ether being at rest or moving – cannot be distinguished by observation seems inacceptable to many physicists. So, in particular, Einstein argued:

> For the theoretician such an asymmetry in the theoretical structure, with no corresponding asymmetry in the system of experience, is intolerable. If we assume the ether to be at rest relatively to K, but in motion relatively to K', the physical equivalence of K and K' seems to me from the logical standpoint, not indeed downright incorrect, but nevertheless inacceptable.

The situation in GR is even more serious. Here, we have not only a single global fixed velocity of the ether which appears unobservable. We have, instead, states of the ether with or without sound waves which appear indistinguishable. Thus, not only the ether at rest and a moving ether would be indistinguishable, but even a homogeneous static ether and an inhomogeneous oscillating ether have to be indistinguishable. So, the argument becomes even stronger.

Moreover, GR appears to be much closer to a philosophy named "relationalism", which goes back to Descartes [4]:

> According to Descartes, there is no "space" at all, but only physical objects which can be in touch with each other. The "position" or location, respectively, of an object is only defined by the naming of other physical objects close to it, i.e. the position of a body is the set of those objects to which the body is contiguous. Equally important is the concept of "motion", which is defined as the change of position. Thus motion is determined by the change of contiguity, i.e. only in relation to other objects. This point of view is denoted as relationalism. [3]

Let's compare this with the concept of absolute space and time proposed by Newton:

> According to Newton, "space" exists by itself, independently of the objects in it. Motion of a body can be defined with respect to space alone, irrespectively whether other objects are present. . . . according to Newton, space exists independently of objects, weather they are present or not. The location of objects is the part of space that they occupy. This implies that motion can be understood without regard to surrounding objects. Similarly, Newton uses absolute time, leading to a space–time picture which provides an always present fixed background over which physics takes place. Objects can always be localized in space and time with respect to this fixed non-dynamical background. [3]

This absolute space and time is realized in the Lorentz ether. In the Minkowski spacetime interpretation of SR, we have yet an absolute spacetime. And this absolute spacetime defines some global properties. In particular, Newton's rotating bucket argument remains valid, and shows that at least acceleration is absolute. Similarly, light rays follow straight lines. But what is a straight line is defined by some absolute, global geometry, which is not influenced by physics. And this property of light rays could be used, via the synchronization procedure known as "Einstein synchronization", to define also a notion of global contemporaneity. It was not uniquely defined, but it depended only on a single parameter – the velocity of the observer who decided to use this "Einstein synchronization".

But in GR we do not have any such global aspects of absolute spacetime. Say, assume that there are no sound waves of the ether near you. So, in your environment everything looks like in flat Minkowski space. But this does not tell us anything about the presence of ether sound waves outside your environment. Thus, even if you know the true preferred coordinates in your environment, you don't know how to extend them to infinity. In SR, knowing the ether velocity locally would have been sufficient to know it globally, and to know how spacetime splits into space and time. In GR, this is no longer possible.

So, special relativity has removed from space only one single information – the velocity of the ether in a single point. This has caused some confusion with common sense ideas about space and time, but at least their combination, the Minkowski spacetime, remained untouched. SR was far away from the philosophy of Descartes that there is no space at all, only physical objects which can be in touch with each other. Instead, GR is quite close to this.

The philosophy of the spacetime interpretation is, essentially, this philosophy of Descartes that space does not exist. In the harmonic GR limit of the Lorentz ether, absolute space and time yet exist, we are only, by unfortunate circumstances that we are unable to see anything beyong this approximation, unable to observe the differences between really different states of the ether. The spacetime interpretation goes beyond this. Not only absolute space and time, but even absolute Minkowski spacetime does not exist. What exists is the gravitational field. And this gravitational field defines all what locally looks like space and time. But the absolute Minkowski spacetime of SR does not exist.

3.3.1 The mathematical formalism of getting rid of the background

So, all the properties of the background of space and time which are yet present in the GR limit of the Lorentz ether (like, in particular, the translational symmetry in space and time) but unobservable, have to be removed from the theory. On the other hand, some of the properties remain. For example, that we can, at least approximately, describe the solutions as functions of four real continuous parameters named "spacetime".

From a mathematical point of view, this is not difficult, because the mathematical apparatus to describe similar things is well-known, and it was necessary to develop an adequate mathematical apparatus anyway, simply because (as usual for mathematics) a similar problem existed in some completely different situation. One such situation was the use of curved coordinates. Curved coordinates were useful already in classical theory. If one describes a sperically symmetric solution, it makes sense to describe it in spherical coordinates, for the simple reason that the solution will look much simpler in these coordinates. But this does not mean that anything of fundamental importance will look simple in these coordinates. One can define distances, volumes and so on in spherical coordinates too, but this becomes a quite complex mathematical object, which has to be defined separately, and nothing as simple as the same objects of Euclidean geometry in Cartesian coordinates. Another situation was the geometry of curved surfaces. A curved surface will be, locally, described by two local coordinates, which are two real numbers. But the distances on this surface will be, clearly, something which depends on the particular properties of the surface, on the embedding of this surface into three-dimensional space. Here, the formula for the distance will not only look more ugly than that of Euclidean distance in Cartesian coordinates, but there is no such Euclidean distance defined on the surface.

So, the mathematical apparatus to handle such things exists. It is the apparatus of differential geometry, which describes the techniques how to use general systems of coordinates. This apparatus allows to describe theories which are defined using a particular special system of coordinates in a covariant way, which no longer depends on this special choice of coordinates. Einstein has initially thought that the possibility to describe a theory in such a covariant way is a physical restriction. But it is not. As was argued by Kretschmann [5], all physical theories can be written in a covariant form. Einstein has accepted this point, and with our presentation of the Lorentz ether we have, essentially, given an illustration. We have transformed the continuity and Euler equations from their classical form (3.1), (3.2) in the preferred coordinates into a covariant equation

(3.3) for the preferred coordinates $\mathfrak{x}^\alpha(x)$. We have also transformed the condition that the ether density has to be positive, $\rho(\mathfrak{x}) > 0$, which is defined only in the preferred coordinates, into a covariant condition for the preferred time coordinate $\mathfrak{t}(x)$ itself, namely that it has to be time-like, $g^{\mu\nu}(x)\partial_\mu\mathfrak{t}(x)\partial_\nu\mathfrak{t}(x) > 0$.

So, we have applied the mathematical apparatus of general covariance to the Lorentz ether, and the result was a formulation of the theory which is covariant, that means, can be written in arbitrary coordinates in the same way. But, on the other hand, in our variant the preferred coordinates of the theory remain in the theory even if we use other coordinates to describe it – as four scalar functions $\mathfrak{x}^\alpha(x)$, which follow a covariant equation $\Box\mathfrak{x}^\alpha(x) = 0$.

On the other hand, we can get now rid of the preferred coordinates, as well as the absolute spacetime which they define. All we have to do is to modify the theory in such a way that they do no longer appear in the equations. The first step was the GR limit. The consequence was that the Einstein equations, as the Euler-Lagrange equations of a Lagrangian which does not depend on the $\mathfrak{x}^\alpha(x)$, also do not depend on the $\mathfrak{x}^\alpha(x)$. The next step, which is done by the spacetime interpretation, is to remove the preferred coordinates completely from the theory. As a consequence, the harmonic equation disappears too.

3.3.2 Mathematical consequences of the spacetime interpretation

The consequence of this is that there is no longer anything which requires that the spacetime has the usual form $\mathbb{R}^3 \times \mathbb{R}$. It has to be some four-dimensional manifold, that's all. For example, the space could be a three-dimensional sphere, so that the spacetime would be $S^3 \times \mathbb{R}$. In this case, there would be no single set of three spatial coordinates which could describe, uniquely, all points. At least at one point the coordinates would degenerate.

How to find such coordinates one can illustrate using the surface of a sphere S^2, like the surface of the Earth. Analogical mathematical formulas would, then, work for S^3 too. One choice of coordinates would be, say, to remove the North pole, and to project all other points to a plane through the equator, using a line through that point and the North pole. Or you could use the same scheme, only with the South pole instead of the North pole. The complete solution of GR for this spacetime would consist of the solution for above choices, together with a proof that the solution is the same on the part covered by above sets of coordinates – that means, all except the two poles. Neither the part of the solution without the North pole nor the part without the South pole would be the complete solution. But only such an incomplete part would allow the introduction of coordinates which could be used as preferred coordinates. So, no Lorentz ether interpretation for this solution is possible – only one for a part of it.

So, together with the possibility of solutions with non-trivial topologies, we see also that the notion of what is a complete solution differs. A complete solution of the Lorentz ether is defined for all values $-\infty < \mathfrak{x}^\alpha < \infty$. The complete solution of GR may contain more, with the other parts described by other sets of coordinates, which cover these other parts of the manifold.

What also changes is the notion of symmetry. A solution of the Einstein equations on $S^3 \times \mathbb{R}$ can have rotational symmetry. But what we would obtain if we would use some coordinates, say, those without the North pole, and use the corresponding ether interpretation? The ether would, quite obviously, have maximal density at the South pole, and decreasing density if one moves away from the South pole. The notion of symmetry of the Lorentz ether is much more restricted, because it requires also that the symmetry applies to the Newtonian background space too.

And, last but not least, there also exists no preferred time coordinate. So, the spacetime is not even obliged to split into something like $M^3 \times \mathbb{R}$ for some arbitrary three-dimensional manifold M^3. And even if it does, the removal of a global time coordinate from the theory has important consequences. Remember that the preferred time coordinate has to be time-like, to give a positive ether density. Once no global time coordinate is obliged to exist, there may be solutions which do not have a global time-like coordinate.

The first solution of GR of this type is quite well-known, it is Gödel's rotating universe. From a topological point of view, it is trivial $\mathbb{R}^3 \times \mathbb{R}$. So, you can introduce there preferred coordinates. But none of these coordinates would be time-like everywhere. So, an attempt to find an ether interpretation ends fatally with negative ether densities at some parts of the solution.

3.3.3 Physical consequences of the spacetime interpretation

The main difference is that the Strong Equivalence Principle obtains a completely different meaning.

In the GR limit of the Lorentz ether, solutions with different choices of the preferred coordinates were interpreted as different solutions (they had, for example, different ether densities and velocities), but they were indistinguishable for observation.

In the spacetime interpretation, these same solutions are no longer considered as different at all. The name "equivalence principle" suggests only some equivalence, which implicitly presupposes some difference. But this is misleading – the different solutions are not considered to describe different physical situations at all. They are considered as different descriptions, using different coordinates, of the same physics.

Bibliography

[1] Schmelzer, I.: A generalization of the Lorentz ether to gravity with general-relativistic limit, Advances in Applied Clifford Algebras 22, 1 (2012), p. 203-242, arXiv:gr-qc/0205035

[2] A. Einstein, Ether and the Theory of Relativity, an address delivered on May 5th, 1920, in the University of Leyden

[3] M. Gaul, C. Rovelli: Loop Quantum Gravity and the Meaning of Diffeomorphism Invariance, Lect.Notes Phys. 541 (2000) 277-324, arxiv:gr-qc/9910079

[4] R. Descartes: Dioptrique, Les Météores. In: Discours de la méthode (1637)

[5] E. Kretschmann, Über den physikalischen Sinn der Relativitätspostulate, A. Einsteins neue und seine ursprüngliche Relativitätstheorie, Ann. Phys. 53, 575-614 (1917)