Relativity

An Introduction to Special and General Relativity Third Edition

Hans Stephani

CAMBRIDGE

CAMBRIDGE www.cambridge.org/9780521811859

This page intentionally left blank

Relativity

An Introduction to Special and General Relativity

Thoroughly revised and updated, and now also including special relativity, this book provides a pedagogical introduction to relativity. It is based on lectures given by the author in Jena over the last decades, and covers the material usually presented in a three-term course on the subject. It is self-contained, but the reader is expected to have a basic knowledge of theoretical mechanics and electrodynamics. The necessary mathematical tools (tensor calculus, Riemannian geometry) are provided. The author discusses the most important features of both special and general relativity, as well as touching on more difficult topics, such as the field of charged pole–dipole particles, the Petrov classification, groups of motions, exact solutions and the structure of infinity.

The book is written as a textbook for undergraduate and introductory graduate courses, but will also be useful as a reference for practising physicists, astrophysicists and mathematicians. Most of the mathematical derivations are given in full and exercises are included where appropriate. The bibliography gives many original papers and directs the reader to useful monographs and review papers.

HANS STEPHANI (1935–2003) gained his Diploma, Ph.D. and Habilitation at the Friedrich-Schiller Universität, Jena. He became Professor of Theoretical Physics in 1992, and retired in 2000. He began lecturing in theoretical physics in 1964 and published numerous papers and articles on relativity over the years. He is also the author of four books.

RELATIVITY

An Introduction to Special and General Relativity

Third Edition

HANS STEPHANI

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo

Cambridge University Press
The Edinburgh Building, Cambridge CB2 2RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521811859

First published in print format 2004

ISBN-13 978-0-511-18904-3 eBook (Adobe Reader)
ISBN-10 0-511-18904-4 eBook (Adobe Reader)

ISBN-13 978-0-521-81185-9 hardback
ISBN-10 0-521-81185-6 hardback

ISBN-13 978-0-521-01069-6 paperback
ISBN-10 0-521-01069-1 paperback

Contents

Preface

Special Relativity originally dealt with the symmetries of the electromagnetic field and their consequences for experiments and for the interpretation of space and time measurements. It arose at the end of the nineteenth century from the difficulties in understanding the properties of light when this light was tested by observers at rest or in relative motion. Its name originated from the surprise that many of the concepts of classical non-relativistic physics refer to a frame of reference ('observer') and are true only relative to that frame.

The symmetries mentioned above show up as transformation properties with respect to Lorentz transformations. It was soon realized that these transformation properties have to be the same for all interacting fields, they have to be the same for electromagnetic, mechanic, thermodynamic, etc. systems. To achieve that, some of the 'older' parts of the respective theories had to be changed to incorporate the proper transformation properties. Because of this we can also say that Special Relativity shows how to incorporate the proper behaviour under Lorentz transformation into all branches of physics. The theory is 'special' in that only observers moving with constant velocities with respect to each other are on equal footing (and were considered in its derivation).

Although the words 'General Relativity' indicate a similar interpretation, this is not quite correct. It is true that historically the word 'general' refers to the idea that observers in a general state of motion (arbitrary acceleration) should be admitted, and therefore arbitrary transformation of coordinates should be discussed. Stated more generally, for a description of nature and its laws one should be able to use *arbitrary coordinate systems*, and in accordance with the *principle of covariance* the form of the laws of nature should not depend essentially upon the choice of the coordinate system. This requirement, in the first place purely mathematical, acquires a physical meaning through the replacement of 'arbitrary coordinate system' by 'arbitrarily moving observer'. The laws of nature should be independent of the state of motion of the observer. Here also belongs the question, raised in particular by Ernst Mach, of whether an absolute acceleration (including an absolute rotation) can really be defined meaningfully, or whether every measurable rotation means a rotation relative to the fixed stars (*Mach's principle*).

But more important for the evolution of General Relativity was the recognition that the Newtonian theory of gravitation was inconsistent with Special Relativity; in it gravitational effects propagate with an infinitely large velocity. So a really new theory of gravitation had to be developed, which correctly reflects the dynamical behaviour of the whole universe and which at the same time is valid for stellar evolution and planetary motion.

General Relativity is the theory of the gravitational field. It is based on Special Relativity in that all laws of physics (except those of the gravitational field) have to be written in the proper special-relativistic way before being translated into General Relativity. It came into being with the formulation of the fundamental equations by Albert Einstein in 1915. In spite of the success of the theory (precession of the perihelion of Mercury, deflection of light by the Sun, explanation of the cosmological redshift), it had retained for a long time the reputation of an esoteric science for specialists and outsiders, perhaps because of the mathematical difficulties, the new concepts and the paucity of applications (for example, in comparison with quantum theory, which came into existence at almost the same time). Through the development of new methods of obtaining solutions and the physical interpretation of the theory, and even more through the surprising astrophysical discoveries (pulsars, cosmic background radiation, centres of galaxies as candidates for black holes), and the improved possibilities of demonstrating general relativistic effects, in the course of the last thirty years the general theory of relativity has become a true physical science, with many associated experimental questions and observable consequences.

The early neglect of relativity by the scientific community is also reflected by the fact that many Nobel prizes have been awarded for the development of quantum theory, but none for Special or General Relativity. Only in 1993, in the laudation of the prize given to J. H. Taylor, Jr. and R. A. Hulse for their detection of the binary pulsar PSR 1913+16, was the importance for relativity (and the existence of gravitational waves) explicitly mentioned.

Modern theoretical physics uses and needs ever more complicated mathematical tools – this statement, with its often unwelcome consequences for the physicist, is true also for the theory of gravitation. The language of the general theory of relativity is differential geometry, and we must learn it, if we wish to ask and answer precisely physical questions. The part on General Relativity therefore begins with some chapters in which the essential concepts and formulae of Riemannian geometry are described. Here suffix notation will be used in order to make the book easier to read for non-mathematicians. An introduction to the modern coordinate-free notation can be found in Stephani *et al.* (2003).

This book is based on the lectures the author gave in Jena through many years (one term Special and two terms General Relativity), and thus gives a rather concise introduction to both theories. The reader should have a good knowledge of classical mechanics and of Maxwell's theory.

My thanks go to all colleagues (in particular in Jena), with whom and from whom I have learnt the theory of relativity. I am especially indebted to J. Stewart and M. Pollock for the translation of most of the parts on General Relativity for the foregoing edition, M. MacCallum for his critical remarks and suggestions, and Th. Lotze for his help in preparing the manuscript.

Notation

Minkowski space: $\mathrm{d}s^2 = \eta_{ab}\,\mathrm{d}x^a\,\mathrm{d}x^b = \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 - c^2\mathrm{d}t^2$

$$= \mathrm{d}\mathbf{r}^2 - c^2\mathrm{d}t^2 = -c^2\mathrm{d}\tau^2.$$

Lorentz transformations: $x^{n'} = L^{n'}{}_m\, x^m, \quad L^{n'}{}_a L_{n'}{}^b = \delta^b_a.$

Special Lorentz transformation:

$$x' = \frac{x - vt}{\sqrt{1 - v^2/c^2}}, \quad ct' = \frac{ct - vx/c}{\sqrt{1 - v^2/c^2}}.$$

Addition of velocities: $v = \dfrac{v_1 + v_2}{1 + v_1 v_2/c^2}.$

Four-velocity: $u^n = \mathrm{d}x^n/\mathrm{d}\tau.$

Riemannian space: $\mathrm{d}s^2 = g_{ab}\,\mathrm{d}x^a\,\mathrm{d}x^b = -c^2\mathrm{d}\tau^2,$

$$g^{ab}g_{bm} = \delta^a_m = g^a_m, \quad g = |g_{ab}|.$$

ε-pseudo-tensor: $\varepsilon^{abmn}; \quad \varepsilon^{1234} = 1/\sqrt{-g},$

$$\varepsilon_{abcd}\varepsilon^{abmn} = -2(g^n_c g^m_d - g^m_c g^n_d).$$

Dualization of an antisymmetric tensor: $\tilde{F}^{ab} = \frac{1}{2}\varepsilon^{abmn}F_{mn}.$

Christoffel symbols: $\Gamma^a_{mn} = \frac{1}{2}g^{ab}(g_{bm,n} + g_{bn,m} - g_{mn,b}).$

Partial derivative: $T_{a,m} = \partial T_a/\partial x^m.$

Covariant derivative: $T^a{}_{;m} = \mathrm{D}T^a/\mathrm{D}x^m = T^a{}_{,m} - \Gamma^a_{mn}T_n,$

$$T_{a;m} = \mathrm{D}T_a/\mathrm{D}x^m = T_{a,m} - \Gamma^n_{am}T_n.$$

Geodesic equation: $\dfrac{\mathrm{D}^2x^i}{\mathrm{D}\lambda^2} = \dfrac{\mathrm{d}^2x^i}{\mathrm{d}\lambda^2} + \Gamma^i_{nm}\dfrac{\mathrm{d}x^n}{\mathrm{d}\lambda}\dfrac{\mathrm{d}x^m}{\mathrm{d}\lambda} = 0.$

Parallel transport along a curve $x^i(\lambda)$: $\mathrm{D}T^a/\mathrm{D}\lambda = T^a{}_{;b}\,\mathrm{d}x^b/\mathrm{d}\lambda = 0.$

Fermi–Walker transport: $\dfrac{\mathrm{D}T^n}{\mathrm{D}\tau} - \dfrac{1}{c^2}T_a\left(\dfrac{\mathrm{d}x^n}{\mathrm{d}\tau}\dfrac{\mathrm{D}^2x^a}{\mathrm{D}\tau^2} - \dfrac{\mathrm{d}x^a}{\mathrm{d}\tau}\dfrac{\mathrm{D}^2x^n}{\mathrm{D}\tau^2}\right) = 0.$

Lie derivative in the direction of the vector field $a^k(x^i)$:

$$\mathcal{L}_{\mathbf{a}}T^n = T^n{}_{,k}a^k - T^k a^n{}_{,k} = T^n{}_{;k}a^k - T^k a^n{}_{;k},$$
$$\mathcal{L}_{\mathbf{a}}T_n = T_{n,k}a^k + T_k a^k{}_{,n} = T_{n;k}a^k + T_k a^k{}_{;n}.$$

Killing equation: $\xi_{i;n} + \xi_{n;i} = \mathcal{L}_\xi g_{in} = 0$.

Divergence of a vector field: $a^i{}_{;i} = \dfrac{1}{\sqrt{-g}}(\sqrt{-g}a^i)_{,i}$.

Maxwell's equations: $F^{mn}{}_{;n} = (\sqrt{-g}F^{mn})_{,n}/\sqrt{-g} = j^m/c$,

$$\tilde{F}^{mn}{}_{;n} = 0.$$

Curvature tensor:

$$a_{m;s;q} - a_{m;q;s} = a_b R^b{}_{msq},$$
$$R^b{}_{msq} = \Gamma^b_{mq,s} - \Gamma^b_{ms,q} + \Gamma^b_{ns}\Gamma^n_{mq} - \Gamma^b_{nq}\Gamma^n_{ms},$$
$$R_{amsq} = \tfrac{1}{2}(g_{aq,ms} + g_{ms,aq} - g_{as,mq} - g_{mq,as}) + \text{non-linear terms}.$$

Ricci tensor: $R_{mq} = R^s{}_{msq} = -R^s{}_{mqs}$; $\quad R^m{}_m = R$.

Field equations: $G_{ab} = R_{ab} - \frac{1}{2}Rg_{ab} = \kappa T_{ab}$.

Perfect fluid: $T_{ab} = (\mu + p/c^2)u_a u_b + p g_{ab}$.

Schwarzschild metric:

$$\mathrm{d}s^2 = \frac{\mathrm{d}r^2}{1-2M/r} + r^2(\mathrm{d}\vartheta^2 + \sin^2\vartheta\,\mathrm{d}\varphi^2) - (1-2M/r)c^2\mathrm{d}t^2.$$

Robertson–Walker metric:

$$\mathrm{d}s^2 = K^2(ct)\left[\frac{\mathrm{d}r^2}{1-\varepsilon r^2} + r^2(\mathrm{d}\vartheta^2 + \sin^2\vartheta\,\mathrm{d}\varphi^2)\right] - c^2\mathrm{d}t^2.$$

Hubble parameter: $H(ct) = \dot{K}/K$.

Acceleration parameter: $q(ct) = -K\ddot{K}/\dot{K}^2$.

$\kappa = 2.07 \times 10^{-48}\,\mathrm{g}^{-1}\mathrm{cm}^{-1}\mathrm{s}^2$, $\quad cH = 55\,\mathrm{km/s\,Mpc}$.

$2M_{\mathrm{Earth}} = 0.8876\,\mathrm{cm}$, $\quad 2M_{\mathrm{Sun}} = 2.9533 \times 10^5\,\mathrm{cm}$.

I. Special Relativity

1

Introduction: Inertial systems and the Galilei invariance of Classical Mechanics

1.1 Inertial systems

Special Relativity became famous because of the bewildering properties of length and time it claimed to be true: moving objects become shorter, moving clocks run slower, travelling people remain younger. All these results came out from a theoretical and experimental study of light propagation as seen by moving observers. More technically, they all are consequences of the invariance properties of Maxwell's equations.

To get an easier access to invariance properties, it is appropriate to study them first in the context of Classical Mechanics. Here they appear quite naturally when introducing the so-called 'inertial systems'. By definition, an *inertial system* is a coordinate system in which the equations of motion take the usual form

$$m\ddot{x}^\alpha = F^\alpha, \quad \alpha = 1, 2, 3 \tag{1.1}$$

(Cartesian coordinates $x^1 = x$, $x^2 = y$, $x^3 = z$, $\ddot{x}^\alpha = \mathrm{d}^2x^\alpha/\mathrm{d}t^2$). Experimentally, an inertial system can be realized in good approximation by a system in which the stars are at rest. Inertial systems are not uniquely defined; if Σ is such a system, then all systems Σ' which originate from Σ by performing a spatial translation, a rotation about a constant (time-independent) angle, a shift of the origin of time, or a motion with constant velocity, are again inertial systems. Accelerated systems such as steadily rotating systems are not inertial systems, cp. also (15.2).

We shall now study the abovementioned transformations in more detail.

1.2 Invariance under translations

Experimental results should not depend on the choice of the origin of the Cartesian coordinate system one is using ('homogeneity of space'). So if there is a system of masses m_N, then their equations of motion

$$m_N \ddot{\mathbf{r}}_N = \mathbf{F}_N \tag{1.2}$$

should be invariant under a translation by a constant vector $\mathbf{b}$, i.e. under the substitution

$$\mathbf{r}'_N = \mathbf{r}_N + \mathbf{b}, \ \ \dot{\mathbf{r}}'_N = \mathbf{r}_N, \ \ \mathbf{F}'_N = \mathbf{F}_N. \tag{1.3}$$

Substituting (1.3) into (1.2), the invariance seems to hold trivially. But a closer inspection of (1.2) shows that if we write it out as

$$m_N \ddot{\mathbf{r}}_N = \mathbf{F}_N(\mathbf{r}_M, \dot{\mathbf{r}}_M, t) \tag{1.4}$$

(the forces may depend on the positions and velocities of all masses), then the substitution $\mathbf{r}'_N = \mathbf{r}_N + \mathbf{b}$ leads to

$$m_N \ddot{\mathbf{r}}'_N = \mathbf{F}_N(\mathbf{r}'_M - \mathbf{b}, \dot{\mathbf{r}}'_M, t). \tag{1.5}$$

This has the form (1.4) only if the force on a mass does not depend on the positions $\mathbf{r}_M$ of the (other) masses, but only on the distances $\mathbf{r}_N - \mathbf{r}_M$, because then we have $\mathbf{F}_N = \mathbf{F}_N(\mathbf{r}_N - \mathbf{r}_M, \dot{\mathbf{r}}_M, t) \rightarrow \mathbf{F}'_N = \mathbf{F}_N(\mathbf{r}'_N - \mathbf{r}'_M, \dot{\mathbf{r}}'_M, t)$; the $\mathbf{b}$ drops out. Closed systems, for which the sources of all forces are part of the system, usually have that property.

Examples of equations of motion which are invariant against translation are $m\ddot{\mathbf{r}} = \mathbf{g}$ (motion in a homogeneous gravitational field) and the motion of a planet (at position $\mathbf{r}$) in the field of the Sun (at position $\mathbf{r}_S$)

$$m\ddot{\mathbf{r}} = f \frac{\mathbf{r} - \mathbf{r}_S}{|\mathbf{r} - \mathbf{r}_S|^3}. \tag{1.6}$$

In a similar way, experimental results should not depend on the choice of the origin of time ('homogeneity of time'), the equations of motion should be invariant under a time translation

$$t' = t + b. \tag{1.7}$$

An inspection of equations (1.4) shows that the invariance is only guaranteed if the forces do not *explicitly* depend on time (they are then time-dependent only via the motion of the sources of the forces); this again will hold if there are no *external* sources of the forces.

We thus can state that for closed systems the laws of nature do not permit an experimental verification, or a sensible definition, of an absolute location in space and time.

1.3 Invariance under rotations

Rotations such as the simple rotation about the z-axis

$$x' = x\cos\varphi + y\sin\varphi, \quad y' = -x\sin\varphi + y\cos\varphi, \quad z' = z, \tag{1.8}$$

are best described using matrices. To do this, we first denote the Cartesian coordinates by

$$x_1 = x^1 = x, \;\; x_2 = x^2 = y, \;\; x_3 = x^3 = z. \tag{1.9}$$

The convention of using x^α as well as x_α for the same set of variables looks rather strange and even clumsy; the reason for this will become clear when dealing with vectors and tensors in both Special and General Relativity. As usual in relativity, we will use the Einstein summation convention: summation over two repeated indices, of which always one is lowered and one is raised.

The general rotation (orthogonal transformation) is a linear transformation and can be written in the two equivalent forms

$$x^{\alpha'} = D^{\alpha'}{}_\beta \, x^\beta, \quad x_{\alpha'} = D_{\alpha'}{}^\beta x_\beta \tag{1.10}$$

(note the position of the indices on the Ds!). Here, and on later occasions in Special and General Relativity, we prefer a notation which distinguishes the new coordinates from the old not by a new symbol (say y^α instead of x^α), but by a prime on the index. This convention is advantageous for many calculations of a general kind, although we shall occasionally deviate from it. The transformation matrices $D^{\alpha'}{}_\beta$ mediating between the two systems thus have two kinds of indices.

Rotations leave angles and lengths fixed; so if there are two arbitrary vectors x^α and ξ^α, then their scalar product has to remain unchanged. With

$$x_{\alpha'} = D_{\alpha'}{}^\beta x_\beta, \quad \xi^{\alpha'} = D^{\alpha'}{}_\gamma \, \xi^\gamma \tag{1.11}$$

that gives the condition

$$x_{\alpha'}\xi^{\alpha'} = D_{\alpha'}{}^\beta D^{\alpha'}{}_\gamma \, x_\beta \xi^\gamma = x_\beta \xi^\beta. \tag{1.12}$$

For arbitrary vectors $\mathbf{x}$ and $\boldsymbol{\xi}$ this can be true only for

$$D_{\alpha'}{}^\beta D^{\alpha'}{}_\gamma = \delta^\beta_\gamma, \quad \alpha, \beta, \gamma = 1, 2, 3. \tag{1.13}$$

Equation (1.13) characterizes the general orthogonal transformation. By taking the determinants on both sides of it (note that $D_{\alpha'}{}^\beta$ and $D^{\alpha'}{}_\gamma$ are numerically identical) we get

$$\left\| D_{\alpha'}{}^\beta \right\|^2 = 1. \tag{1.14}$$

The transformations with $\|D_{\alpha'}{}^{\beta}\| = +1$ are rotations; an example is the rotation (1.8) with

$$D_{\alpha'}{}^{\beta} = \begin{pmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} . \tag{1.15}$$

Transformations with $\|D_{\alpha'}{}^{\beta}\| = -1$ contain reflections such as, for example, the inversion

$$\begin{aligned} x' &= -x, \\ y' &= -y, \\ z' &= -z, \end{aligned} \qquad \Rightarrow \qquad D_{\alpha'}{}^{\beta} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} . \tag{1.16}$$

To apply a rotation to the equations of motion, we first observe that for time-independent rotations we have

$$x^{\alpha'} = D^{\alpha'}{}_{\beta}\, x^{\beta} \quad \Rightarrow \quad \ddot{x}^{\alpha'} = D^{\alpha'}{}_{\beta}\, \ddot{x}^{\beta} . \tag{1.17}$$

We then note that the force $\mathbf{F}$ is a vector, i.e. its components F^{α} transform in the same way as the components of the position vector x^{α}. If we now multiply both sides of equation (1.1) by $D^{\alpha'}{}_{\beta}$, we get

$$D^{\alpha'}{}_{\beta}\, \ddot{x}^{\beta} = m\ddot{x}^{\alpha'} = D^{\alpha'}{}_{\beta}\, F^{\beta} = F^{\alpha'} ; \tag{1.18}$$

the form of the equation remains unchanged. But if we also take into account the arguments in the components of the force,

$$m\ddot{x}^{\alpha'} = D^{\alpha'}{}_{\beta}\, F^{\alpha}(x^{\beta}, \dot{x}^{\beta}, t) = F^{\alpha'}(x^{\beta}, \dot{x}^{\beta}, t), \tag{1.19}$$

we see that the $F^{\alpha'}$ may depend on the wrong kind of variables. This will not happen if the F^{α} depend only on invariants, which in practice happens in most cases.

An example of an invariant equation is given by (1.6): the $\mathbf{r} - \mathbf{r}_S$ is a vector, and the distance $|\mathbf{r} - \mathbf{r}_S|$ is rotationally invariant.

We thus can state: since the force is a vector, and for closed systems the force-components depends only on invariants, the equations of motions are rotationally invariant and do not permit the definition of an absolute direction in space.

1.4 Invariance under Galilei transformations

We consider two systems which are moving with a constant velocity $\mathbf{v}$ with respect to each other:

$$\mathbf{r}'_N = \mathbf{r}_N - \mathbf{v}t, \quad t' = t \tag{1.20}$$

(*Galilei transformation*). Because of $\dot{\mathbf{r}}'_N = \dot{\mathbf{r}}_N - \mathbf{v}$, $\ddot{\mathbf{r}}'_N = \ddot{\mathbf{r}}_N$, the equations of motion (1.4) transform as

$$m_N \ddot{\mathbf{r}}'_N = m_N \ddot{\mathbf{r}}_N = \mathbf{F}_N(\mathbf{r}'_M + \mathbf{v}t, \dot{\mathbf{r}}'_M + \mathbf{v}, t). \tag{1.21}$$

Although the constant $\mathbf{v}$ drops out when calculating the acceleration, the arguments of the force may still depend on $\mathbf{v}$. The equations are invariant, however, if only relative positions $\mathbf{r}_M - \mathbf{r}_N$ (as discussed above) and relative velocities $\dot{\mathbf{r}}_N - \dot{\mathbf{r}}_M$ enter. This is usually the case if the systems are closed and the equations are properly written. Take for example the well known example of a motion in a constant gravitational field $\mathbf{g}$ under the influence of friction,

$$m\ddot{\mathbf{r}} = -a\dot{\mathbf{r}} - m\mathbf{g}. \tag{1.22}$$

At first glance, because of the explicit $\dot{\mathbf{r}}$ occurring in it, this equation seems to be a counterexample. But what is really meant, and is the cause of the friction, is the relative velocity with respect to the air. The equation (1.22) should correctly be written as

$$m\ddot{\mathbf{r}} = -a(\dot{\mathbf{r}} - \mathbf{v}_{\mathrm{Air}}) - m\mathbf{g}, \tag{1.23}$$

and the invariance is now obvious.

For closed systems, the equations of motions are invariant under Galilei transformations; an absolute velocity cannot be defined. Stated differently: only relative motions can be defined and measured (*Galilei's principle of relativity*).

We close this section with two remarks. In all three cases of invariances we had to refer to closed systems; how far do we have to go to get a really closed system? Is our Galaxy sufficient, or have we to take the whole universe? Second, we saw that only relative velocities matter; what about acceleration – why is this absolute?

1.5 Some remarks on the homogeneity of time

How can one check that space and time are really homogeneous? We want to discuss that problem a little bit for the case of time.

We start with the notion 'constant velocity'. How can one check that a mass is moving with constant velocity? Of course by measuring distances and reading clocks. How does one know that the clocks are going uniformly? After some consideration, and looking at standard procedures, one concludes that good clocks are made by taking a periodic process (rotation of the Earth, harmonic oscillator, vibration of a molecule) and dividing that into smaller parts. But how does one know that this fundamental process is really periodic – no clock to measure it is available! The only way out is *to define* that process as being periodic. But which kind of process should one use for that?

Of course, one has to consult Newton's equations of motion

$$m\ddot{\mathbf{r}} = \frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}t^2} = \mathbf{F} \tag{1.24}$$

and to take a process, such as the rotation of the Earth around the Sun, which is periodic when these equations hold.

To see that really a definition of the time is hidden here in the equations of motion, consider a transformation

$$T = f(t) \quad \Rightarrow \quad \mathrm{d}T = f'\mathrm{d}t,\ \ \mathrm{d}/\mathrm{d}t = f'\mathrm{d}/\mathrm{d}T \tag{1.25}$$

of the time. In the new time variable T the equations of motion (1.24) read

$$f'^2\frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}T^2} + f'f''\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}T} = \mathbf{F}; \tag{1.26}$$

they no longer have the Newtonian form.

We conclude that the correct, appropriate time coordinate is that in which the equations of motion take the simple form (1.24); the laws of mechanics guarantee that such a time really exists. But it here remains an open question whether this time coordinate, which is derived from planetary motion, is also the appropriate time to describe phenomena in other fields of physics such as light propagation. This questions will be answered by Special Relativity – in the negative.

Exercises

1.1 Is the equation $m\ddot{\mathbf{r}} - k\mathbf{r} = 0$ (harmonic oscillator) invariant under translations?

1.2 Show that a rotation $D^{\alpha'}{}_{\beta}$ always has one real eigenvector $\mathbf{w}$ with $D^{\alpha'}{}_{\beta}\, w^{\beta} = \lambda w^{\alpha}$, and that $w^{\alpha} = (1, \mathrm{i}, 0)$ is a complex eigenvector of the rotation (1.15). What are the corresponding eigenvalues?

1.3 Is $m\ddot{\mathbf{r}} = f(x)\mathbf{r}$ rotationally invariant?

1.4 Show that the Laplacian is invariant under rotations, i.e. that $\partial^2/\partial x^{\alpha}\partial x_{\alpha} = \partial^2/\partial x^{\alpha'}\partial x_{\alpha'}$ holds.

2

Light propagation in moving coordinate systems and Lorentz transformations

2.1 The Michelson experiment

At the end of the nineteenth century, it was a common belief that light needs and has a medium in which it propagates: light is a wave in a medium called ether, as sound is a wave in air. This belief was shattered when Michelson (1881) tried to measure the velocity of the Earth on its way around the Sun. He used a sensitive interferometer, with one arm in the direction of the Earth's motion, and the other perpendicular to it. When rotating the instrument through an angle of $90°$, a shift of the fringes of interference should take place: light propagates in the ether, and the velocity of the Earth had to be added that of the light in the direction of the respective arms. The result was zero: there was no velocity of the Earth with respect to the ether.

This negative result can be phrased differently. Since the system of the ether is an inertial system, and that of the Earth is moving with a (approximately) constant velocity, the Earth's system is an inertial system too. So the Michelson experiment (together with other experiments) tells us that the velocity of light is the same for all inertial systems which are moving with constant velocity with respect to each other (*principle of the invariance of the velocity of light*). The speed of light in empty space is the same for all inertial systems, independent of the motion of the light source and of the observer.

This result does not violate Galilei's principle of relativity as stated at the end of Section 1.4: it confirms that also the ether cannot serve to define an absolute velocity. But of course something is wrong with the transformation law for the velocities: light moving with velocity c in the system of the ether should have velocity $c+v$ in the system of the Earth.

This contradiction can be given a geometric illustration (see Fig. 2.1). Consider two observers Σ (coordinates x, y, z, t) and Σ' (coordinates $x' = x - vt$, $y' = y$, $z' = z$, $t' = t$), moving with constant velocity $\mathbf{v}$ with respect to each other. At $t = 0$, when their coordinate systems coincide, a light signal is emitted at the origin. Since for both of them the light velocity is c, after a time T the light signal has reached the sphere

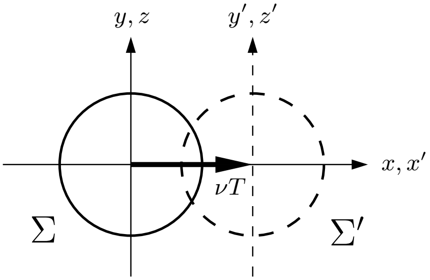

Fig. 2.1. Light propagation as seen by two observers in relative motion; $t = t' = T$.

$x^2 + y^2 + z^2 = c^2T^2$ for Σ, and $(x - vT)^2 + y^2 + z^2 = c^2T^2$ for Σ'. But this a contradiction, the light front cannot be simultaneously at the two spheres!

It will turn out that it is exactly this 'simultaneously' which has to be amended.

2.2 The Lorentz transformations

Coordinates The wave front of light emitted at $t = 0$ at the origin has reached the three-dimensional light sphere

$$x^2 + y^2 + z^2 - c^2t^2 = 0 \tag{2.1}$$

at the time t. Space and time coordinates enter here in a very symmetric way. Therefore we adapt our coordinates to this light sphere and take the time as a fourth coordinate $x^4 = ct$. More exactly, we use

$$x^a = (x,\, y,\, z,\, ct), \quad x_a = (x,\, y,\, z, -ct), \quad a = 1, \ldots, 4. \tag{2.2}$$

The two types of coordinates are obviously related by means of a matrix $\boldsymbol{\eta}$, which can be used to raise and lower indices:

$$\begin{aligned} x_a &= \eta_{ab}x^b, \\ x^a &= \eta^{ab}x_b, \end{aligned} \qquad \eta_{ab} = \eta^{ab} = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & -1 \end{pmatrix}, \qquad \eta^a_b = \delta^a_b. \tag{2.3}$$

Using these coordinates, (2.1) can be written as

$$x^a x_a = \eta_{ab}\, x^a x^b = x^2 + y^2 + z^2 - c^2t^2 = 0. \tag{2.4}$$

Invariance of light propagation and Lorentz transformations We now determine the coordinate transformations which leave the light sphere (2.4) invariant, thus ensuring that the light velocity is the same in both systems. Unlike the Galilei transformations (1.20), where the time coordinate was kept constant, it too is transformed here: the definition of the time scale will be adjusted to the light propagation, as it is adjusted to the equations of motion in Newtonian mechanics, cp. Section 1.4.

The transformations we are looking for should be one-to-one, and no finite point should go into infinity; they have to be linear. Neglecting translations, they have the form

$$x^{n'} = L^{n'}{}_a\, x^a, \quad x_{m'} = L_{m'}{}^b\, x_b, \quad L_{m'}{}^b = \eta_{m'n'}\eta^{ab}L^{n'}{}_a \tag{2.5}$$

(for the notation, see the remarks after equation (1.10); note that $\eta_{m'n'}$ and η_{ab} have the same numerical components).

To give the light sphere the same form $x_n x^n = 0 = x^{n'} x_{n'}$ in both coordinates, the transformations (2.5) have to satisfy

$$x^{n'} x_{n'} = L^{n'}{}_a L_{n'}{}^b x^a x_b = x^b x_b, \tag{2.6}$$

which for all x^a is possible only if

$$L^{n'}{}_a L_{n'}{}^b = \delta^b_a, \quad a, b, n' = 1, \ldots, 4. \tag{2.7}$$

These equations define the *Lorentz transformations*, first given by Waldemar Voigt (1887). The discussion of these transformations will fill the next chapters of this book.

If we also admit translations,

$$x^{n'} = L^{n'}{}_a\, x^a + c^{n'}, \quad c^{n'} = \text{ const.}, \tag{2.8}$$

we obtain the *Poincaré transformations.*

Lorentz transformations, rotations and pseudorotations Equation (2.7) looks very similar to the defining equation (1.13) for rotations, $D^{v'}{}_\alpha D_{\nu'}{}^\beta = \delta^\beta_\alpha$, to which it reduces when the time (the fourth coordinate) is kept fixed:

$$L^{n'}{}_a = \begin{pmatrix} D^{\nu'}{}_\alpha & 0 \\ 0 & 1 \end{pmatrix}. \tag{2.9}$$

Rotations leave $x^\alpha x_\alpha = x^2 + y^2 + z^2$ invariant, Lorentz transformations $x^a x_a = x^2 + y^2 + z^2 - c^2t^2$.

We now determine the special Lorentz transformation which corresponds to a motion (with constant velocity) in the x-direction. We start from

$$\begin{aligned} x' &= Ax + Bct, \; y' = y \\ ct' &= Cx + Dct, \; z' = z \end{aligned} \quad \Longleftrightarrow \quad L^{n'}{}_a = \begin{pmatrix} A & 0 & 0 & B \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ C & 0 & 0 & D \end{pmatrix}. \tag{2.10}$$

When we insert this expression for ${L^{n'}}_a$ into (2.7), we get the three conditions $A^2 - C^2 = 1$, $D^2 - B^2 = 1$, $AB = CD$, which can be parametrically solved by $A = D = \cosh\varphi$, $B = C = -\sinh\varphi$, so that the Lorentz transformation is given by

$$\begin{aligned} x' &= \quad x\cosh\varphi - ct\sinh\varphi, \qquad & y' &= y, \\ ct' &= -x\sinh\varphi + ct\cosh\varphi, \qquad & z' &= z. \end{aligned} \tag{2.11}$$

The analogy with the rotations

$$\begin{aligned} x' &= \quad x\cos\varphi - y\sin\varphi, \qquad & z' &= z, \\ y' &= -x\sin\varphi + y\cos\varphi, \qquad & t' &= t \end{aligned} \tag{2.12}$$

is obvious – but what is the physical meaning of φ in the case of the pseudorotations (2.11)?

To see this, we consider the motion of the origin $x' = 0$ of the moving coordinate system Σ' as seen from Σ. From $x' = 0$ and (2.11) we have

$$v = \frac{\mathrm{d}x}{\mathrm{d}t} = \frac{c\sinh\varphi}{\cosh\varphi} \quad \Rightarrow \quad \tanh\varphi = \frac{v}{c}, \tag{2.13}$$

φ is in a simple way related to the velocity v. If we substitute v for φ in the pseudorotations (2.11), we get the well-known form

$$x' = \frac{x - vt}{\sqrt{1 - v^2/c^2}}, \quad ct' = \frac{ct - vx/c}{\sqrt{1 - v^2/c^2}}, \quad y' = y, \quad z' = z \tag{2.14}$$

of the special Lorentz transformation. This transformation describes the transformation between a system Σ and a system Σ' which moves in the x-direction with constant velocity v with respect to Σ.

For small velocities, $v/c \ll 1$, we regain $x' = x - vt$, i.e. the Galilei transformation; we see that Newtonian mechanics is valid for small velocities, discrepancies will appear only if the particles are moving very fast. We shall come back to this question in Chapter 4.

If we solve (2.14) for the x^a, we will get the same equations, with the primed and unprimed coordinates exchanged and v replaced by $-v$.

2.3 Some properties of Lorentz transformations

In this section we shall discuss some of the more mathematical properties of the Lorentz transformations. Many of the physical implications will be dealt with in the following chapters, in particular in Chapter 3.

Group property The Lorentz transformations form a group. To prove this, we remark that matrix multiplication is associative, and see by inspection that the identity $L^{n'}{}_a = \delta^n_a$ is contained. Two successive transformations yield

$$x^{m''} = L^{m''}{}_{n'}\, x^{n'} = L^{m''}{}_{n'} L^{n'}{}_a\, x^a = L^{m''}{}_a\, x^a. \tag{2.15}$$

This will be a Lorentz transformation if $L^m{}_a$ satisfies (2.7), which is indeed the case:

$$L^{m''}{}_a L_{m''}{}^d = L^{m''}{}_{n'} L^{n'}{}_a L_{m''}{}^{b'} L_{b'}{}^d = \delta^b_n L^{n'}{}_a L_{b'}{}^d = L^{b'}{}_a L_{b'}{}^d = \delta^d_a. \tag{2.16}$$

In a similar way one can show that the inverse of a Lorentz transformation is again such a transformation.

Classification of Lorentz transformations The 4×4 matrices $L^{n'}{}_a$ which describe Lorentz transformations have 16 parameters which are subject to the ten conditions (2.5); there are six independent Lorentz transformations, corresponding to three motions (e.g. in the direction of the axes) and three rotations. As we shall show now, there are four distinct types of Lorentz transformations.

From the defining equations (2.5) and (2.7) we immediately get

$$\left\| L^{n'}{}_a L_{n'}{}^b \right\| = \left\| \delta^b_a \right\| = 1, \quad \left\| L_{n'}{}^b \right\| = \left\| \eta_{n'm'} \right\| \cdot \left\| \eta^{ab} \right\| \cdot \left\| L^{m'}{}_a \right\| = \left\| L^{m'}{}_a \right\|, \tag{2.17}$$

so that

$$\left\| L^{n'}{}_a \right\| = \begin{cases} +1 \\ -1 \end{cases} \tag{2.18}$$

holds. Evaluating the (4,4)-component of (2.5), we obtain (remember that indices are raised and lowered by means of $\boldsymbol{\eta}$!)

$$\begin{aligned} 1 &= -\eta_{44} L^{n'}{}_4 L_{n'}{}^4 \\ &= -\eta^{n'm'} L_{n'4} L_{m'4} = (L_{4'4})^2 - (L_{1'4})^2 - (L_{2'4})^2 - (L_{3'4})^2 \end{aligned} \tag{2.19}$$

and conclude that

$$L^{4'}{}_4 = \begin{cases} \geq +1 \\ \leq -1 \end{cases}. \tag{2.20}$$

Equations (2.18) and (2.20) show that there are four distinct classes of Lorentz transformations. Those which do not contain reflections have $\|L^n{}_a\| = +1$ and are called proper. Transformations with $L^4{}_4 \geq +1$ are called orthochronous; because of $ct' = L^4{}_4\, ct + \cdots$ they preserve the direction of time.

Normal form of a proper orthochronous Lorentz transformation By using an adapted coordinate system, any proper orthochroneous Lorentz transformation can be written in the form

$$L^{n\prime}{}_a = \begin{pmatrix} \heartsuit & 0 & 0 & \heartsuit \\ 0 & \times & \times & 0 \\ 0 & \times & \times & 0 \\ \heartsuit & 0 & 0 & \heartsuit \end{pmatrix} \tag{2.21}$$

of direct product of a special Lorentz transformation (motion) ($\heartsuit$) and a rotation ($\times$) in the plane perpendicular to that motion. We leave the proof to the reader, see Exercise 2.2.

Lorentz transformation for an arbitrarily directed velocity We start with a question: how does a Lorentz transformation between two systems whose spatial axes are parallel, as in Fig. 2.2, look? By 'parallel' we mean that, for a fixed time, x' (for example) does not change if only y and z vary: in

$$x' = L^{1'}{}_a\, x^a = L^{1'}{}_1\, x + L^{1'}{}_2\, y + L^{1'}{}_3\, z + L^{1'}{}_4\, ct \tag{2.22}$$

the $L^{1'}{}_2$ and $L^{1'}{}_3$ are assumed to be zero, and from the y'- and z'-equations we see that also $L^{2'}{}_1$, $L^{2'}{}_3$, $L^{3'}{}_1$and $L^{3'}{}_2$ should vanish. There should be at least one component of the velocity, so we assume $L^{1'}{}_4 \neq 0$. Inserting all this into the defining equations (2.7), the result may be a surprise to the reader: the Lorentz transformation necessarily is of the form (2.11) of a motion in the x-direction (which is preferred here because of the assumption $L^{1'}{}_4 \neq 0$). So if the spatial axes of the two systems should be parallel, then the motion must be in the direction of one of the axes! For all other cases, the Lorentz transformations contain also terms which cause a rotation of the spatial system. For rotations the analogous effect is well known: none of the axes of a coordinate system can remain unchanged unless it coincides with the axis of the rotation.

So one should not be surprised that the Lorentz transformation describing the motion of the system Σ' with an arbitrarily directed velocity V^α (with no 'extra' rotation) looks rather complicated:

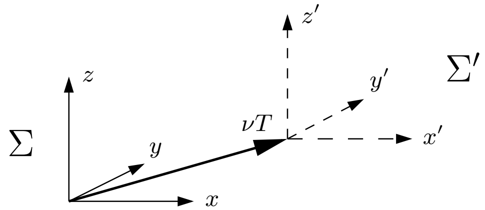

Fig. 2.2. Lorentz transformations between parallel systems.

$$L^{a'}{}_{b} = \begin{pmatrix} (\gamma-1)n^\alpha n_\beta + \delta^\alpha_\beta & -v\gamma n^\alpha/c \\ -v\gamma n_\beta/c & \gamma \end{pmatrix}, \quad V^\alpha = vn^\alpha/c,$$
$$n^\alpha n_\alpha = 1, \quad \gamma \equiv (1-v^2/c^2)^{-1/2}, \quad \alpha,\beta = 1,2,3. \tag{2.23}$$

Note that the rotational part in (2.23), the term $(\gamma-1)n^\alpha n_\beta$, is of second order in v/c.

Velocity addition formula for parallel velocities What is the result if we perform two successive Lorentz transformations, both corresponding to motions in the x-direction? Since the Lorentz transformations form a group, of course again a transformation of that type – but with what velocity?

Lorentz transformations are pseudorotations, i.e. they satisfy

$$\begin{aligned} x' &= x\cosh\varphi_1 - ct\sinh\varphi_1, & ct' &= -x\sinh\varphi_1 + ct\cosh\varphi_1, \\ x'' &= x'\cosh\varphi_2 - ct'\sinh\varphi_2, & ct'' &= -x'\sinh\varphi_2 + ct'\cosh\varphi_2. \end{aligned} \tag{2.24}$$

To get (x'', ct'') in terms of (x, ct), we observe that one adds rotations about the same axis by adding the angles:

$$\begin{aligned} x'' &= x\cosh\varphi - ct\sinh\varphi, \\ ct'' &= -x\sinh\varphi - ct\cosh\varphi, \end{aligned} \qquad \varphi = \varphi_1 + \varphi_2. \tag{2.25}$$

To translate this relation into one for the velocities, we have to use (2.13), i.e. $\tanh\varphi = v/c$, and the well-known theorem for the hyperbolic tangent,

$$\tanh\varphi = \tanh(\varphi_1+\varphi_2) = \frac{\tanh\varphi_1 + \tanh\varphi_2}{1+\tanh\varphi_1\tanh\varphi_2}. \tag{2.26}$$

We obtain

$$v = \frac{v_1+v_2}{1+v_1v_2/c^2}. \tag{2.27}$$

For small velocities, $v_n/c \ll 1$, we get the Galilean addition formula $v = v_1 + v_2$. If we take the velocity of light as one of the velocities (as a limiting case, since the Lorentz transformations (2.14) are singular for $v = c$), we get

$$v = \frac{c+v_2}{1+v_2/c} = c, \tag{2.28}$$

the velocity of light cannot be surpassed.

On the other hand, if we take two velocities smaller than that of light, we have, with $v_1 = c - \lambda,\ v_2 = c - \mu,\ \lambda, \mu > 0$,

$$\begin{aligned} v &= \frac{2c-\lambda-\mu}{1+(c-\lambda)(c-\mu)/c^2} = c\frac{2c-\lambda-\mu}{2c-\lambda-\mu+\lambda\mu/c} \\ &= \frac{c}{1+\lambda\mu/[c(2c-\lambda-\mu)]} \leq c, \end{aligned} \tag{2.29}$$

it is not possible to reach the velocity of light by adding velocities less than that of light. The velocity addition formula (2.27) seems to indicate that the velocity of light plays the role of a maximum speed; we shall come back to this in the next chapter.

The addition of two non-parallel velocities will be considered in Section 4.4.

Exercises

2.1 Show that the inverse of a Lorentz transformation is again a Lorentz transformation.

2.2 Show by considering the eigenvalue equation $L^{a'}{}_b\, x_b = \lambda x^a$ that the four eigenvalues λ_a of a proper orthochroneous Lorentz transformation obey $\lambda_1\lambda_2 = 1 = \lambda_3\lambda_4$, and that by using the eigenvectors the Lorentz transformation can be written as indicated in (2.21).

2.3 Show that the transformation (2.23) is indeed a Lorentz transformation, and that origin of the system Σ' obtained from Σ by (2.23) moves with the velocity V^α.

2.4 Show by directly applying (2.14) twice that (2.26) is true.

2.5 In a moving system Σ', a rod is at rest, with an angle φ' with respect to the x'-axis. What is the angle φ with respect to the x-axis?

3

Our world as a Minkowski space

In this chapter we will deal with the physical consequences of the Lorentz transformations. Most of them were first found and understood by Einstein (1905), although most of the more technical properties considered in the last chapter were known before him.

3.1 The concept of Minkowski space

We have seen that the velocity of light is the same for all inertial systems, i.e. for all observers which move with constant velocity with respect to each other. The velocity of light is just one aspect of Maxwell equations, so that in fact the Michelson experiment shows that Maxwell equations are the same in all inertial systems. Since the elements of our world interact not only by electromagnetic fields, but also by gravitation, heat exchange, and nuclear forces, for example, the same must be true for all these interactions. *The laws of physics are the same for all inertial systems* (principle of relativity).

The principle of relativity does not exclude the Galilei transformations of mechanics, if one does not specify the transformations between inertial systems. This can be done by demanding that *the velocity of light is the same for all inertial systems* (principle of the invariance of the velocity of light).

Both principles together characterize Special Relativity. They are most easily incorporated into the laws of physics if one uses the concept of Minkowski space.

The four-dimensional Minkowski space, or world, or space-time, comprises space and time in a single entity. This is done by using Minkowski coordinates

$$x^a = (x^\alpha, ct) = (\mathbf{r}, ct), \quad x_a = \eta_{ab}\, x^b = (x_\alpha, -ct). \tag{3.1}$$

A point in this space is characterized by specifying space and time; it may be called an event.

The metrical properties of Minkowski space (in Minkowski coordinates) are given by its line element

$$\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 - c^2\,\mathrm{d}t^2 = \mathrm{d}\mathbf{r}^2 - c^2\,\mathrm{d}t^2 = \eta_{ab}\,\mathrm{d}x^a\,\mathrm{d}x^b. \tag{3.2}$$

This line element is invariant under Lorentz transformations

$$x^{n'} = L^{n'}{}_a\, x^a, \quad L^{n'}{}_a L_{n'}{}^b = \delta^b_a \tag{3.3}$$

since $x^n x_n$ is. Note that $\mathrm{d}s^2$ is not positive definite!

3.2 Four-vectors and light cones

A four-vector $a^n = (a^1, a^2, a^3, a^4) = (\mathbf{a}, a^4)$ is a set of four elements which transforms like the components x^n of the position vector,

$$a^{n'} = L^{n'}{}_m\, a^m. \tag{3.4}$$

An example is the vector connecting two points P_1 and P_2 of Minkowski space, $\overrightarrow{P_1P_2} = (x^2 - x^1, y^2 - y^1, z^2 - z^1, ct^2 - ct^1)$.

Obviously, a Lorentz transformation mixes the spacelike and the timelike parts of a four-vector, but leaves the 'length' fixed:

$$a^{n'} a_{n'} = L^{n'}{}_m L_{n'}{}^b a^m a_b = a^n a_n = \text{ inv.} \tag{3.5}$$

This invariant can have either sign, or can be zero, depending on the relative size of the spacelike and timelike parts of the vector. This leads to the following invariant classification of four-vectors:

$$a^n a_n = \mathbf{a}^2 - (a^4)^2 \begin{cases} > 0 & \text{spacelike vector} \\ = 0 & \text{null vector} \\ < 0 & \text{timelike vector} \end{cases} . \tag{3.6}$$

For a given vector a^n, one can always perform a (spatial) rotation of the coordinate system so that $\mathbf{a}$ points in the x-direction: $a^n = (a^1, 0, 0, a^4)$. A special Lorentz transformation (2.14) then yields

$$a^{1'} = \frac{a^1 - va^4/c}{\sqrt{1 - v^2/c^2}}, \quad a^{4'} = \frac{a^4 - va^1/c}{\sqrt{1 - v^2/c^2}}. \tag{3.7}$$

For $\left|a^1/a^4\right| > 1$, one can make $a^{4'}$ vanish by choice of v (note that v has to be smaller than c !), and similarly in the other cases. So one gets the following normal forms of four-vectors.

$$\text{Normal forms:} \qquad \begin{array}{rl} \text{spacelike vector:} & a^n = (a, 0, 0, 0) \\ \text{null vector:} & a^n = (a, 0, 0, a) \\ \text{timelike vector:} & a^n = (0, 0, 0, a). \end{array} \tag{3.8}$$

If we have two four-vectors a^n and b^n, then we can define the scalar product of the two by

$$|ab| = a^i b_i = \eta_{in} a^i b^n. \tag{3.9}$$

This is of course an invariant under Lorentz transformations. When $|ab|$ is zero, the two vectors are called orthogonal, or perpendicular, to each other. Note that in this sense a null vector is perpendicular to itself.

A light wave emanating at $t = 0$ from the origin of the coordinate system will at time t have reached the points $\mathbf{r}$ with

$$\mathbf{r}^2 - c^2t^2 = 0. \tag{3.10}$$

If we suppress one of the spatial coordinates, equation (3.10) describes a cone in (x, y, ct)-space. Therefore one calls (3.10) the *light cone*. As Fig. 3.1 shows, the light cone separates timelike vectors inside it from the spacelike vectors outside; null vectors are tangent to it.

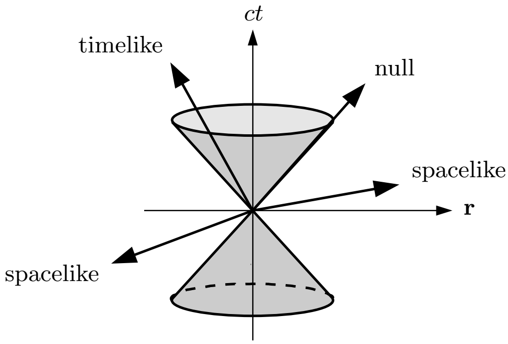

Fig. 3.1. Light cone structure of Minkowski space.

One can attempt to visualize the special Lorentz transformation (2.14) in Minkowski space by drawing the lines $x' = 0$ and $ct' = 0$ as ct'-axis or x'-axis, respectively, for a given value of v/c, see Fig. 3.2. This figure clearly shows that the new ct'-axis always lies inside the light cone (and the new x'-axis outside), that the transformation becomes singular for $v = c$, and that any timelike (spacelike) vector can be given its normal form by a suitable Lorentz transformation. But it does not show that the two coordinate systems are completely equivalent as in fact they are.

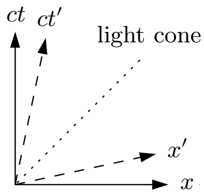

Fig. 3.2. Visualization of a Lorentz transformation.

3.3 Measuring length and time in Minkowski space

The problem One may argue that the results of any measurement should be independent of the observer who made them. If we admit observers in relative motion, then only invariants with respect to Lorentz transformations will satisfy that condition. So for example (spacelike) distances which occur only as a part of a four-vector do not have an invariant meaning.

In practice one is accustomed to measuring spatial distances and time-intervals separately, and one often insists on using these concepts. But then the results of a measurement depend on the state of motion of the observer, as the components of a three-vector depend on the orientation of the Cartesian coordinate system one uses. The typical question which then arises is the following: suppose two observer Σ and Σ' (in relative motion) make some measurements; how are their results related? The answer to this question leads to some of the most spectacular results of Special Relativity theory.

The notion of simultaneity As a prerequisite, we will consider the meaning of 'same place at different times'. If an observer Σ states this for an object, it means the object is at rest at $x = 0$ (for example). For an observer Σ' moving with respect to Σ and to the object, the object changes its position; from (2.14) one gets

$$x' = \frac{x - vt}{\sqrt{1 - v^2/c^2}}, \; x = 0 \quad \Rightarrow \quad x' = \frac{-vt}{\sqrt{1 - v^2/c^2}}. \tag{3.11}$$

There is no *absolute* being at the same place for different times.

This is trivial – but the corresponding result obtained by interchanging the role of space and time is not. If an observer Σ states that two events at different places x_A and x_B are simultaneous (observed at the same time t_0), then the application of a Lorentz transformation gives

$$ct'_A = \frac{ct_0 - vx_A/c}{\sqrt{1 - v^2/c^2}}, \quad ct'_B = \frac{ct_0 - vx_B/c}{\sqrt{1 - v^2/c^2}}, \quad c(t'_A - t'_B) = \frac{(x_B - x_A)v}{c\sqrt{1 - v^2/c^2}}. \tag{3.12}$$

For an observer Σ' the two events are no longer simultaneous: *there is no absolute simultaneity at different places.*

This result has been much debated. In the beginning many people objected to that statement, and most of the attempts to disprove Special Relativity rely on the (hidden) assumption of an absolute simultaneity. There seems to be a psychological barrier which makes us refuse to acknowledge that our personal time which we feel passing may be only relative.

We now shall analyze the notion of simultaneity in more detail, just for a single observer. How can we judge and decide that two events at different places A and B happen at the same time? Just to assume 'we know it' is tantamount to assuming that there are signals with an infinite velocity coming from A and B which tell us that events have taken place; also, though not said in those terms, Newtonian physics uses this concept. To get a more precise notion, our first attempt may be to say: two events are simultaneous if two synchronized clocks situated at A and B show the same time. But how can we be sure that the two clocks are synchronized? We cannot transport one of two identical clocks from A to B, since the transport may badly disturb the clock and we have no way of checking that. Nor can we send a signal from A to B, divide the distance $\overline{AB}$ by the signal's velocity V to get the travelling time, and compare thus the clocks: without a clock at B, we cannot know the velocity V!

Considerations like this tell us that we need to *define* simultaneity. As with the definition of time discussed in Section 1.5, simultaneity has to be defined so that the laws of nature become simple, which means here: so that the Lorentz transformations hold. Einstein showed that a possible definition is like this: two events at A and B are simultaneous if light signals emitted simultaneously with those events arrive simultaneously in the middle of the line $\overline{AB}$. Note that here 'simultaneously' has been used only for events occurring at the same place!

By procedures like this, an observer Σ can synchronize his system of clocks in space-times; for a different observer Σ', this system is of course no longer synchronized.

Time dilatation At $\mathbf{r} = 0$, an event takes place between $t_B = 0$ and $t_E = T$; for an observer Σ at rest with that event the corresponding time-interval is of course $\Delta t = T$. Because of the Lorentz transformation (2.14) we then have $ct'_B = 0$, $c\,t'_E = cT/\sqrt{1 - v^2/c^2}$; for a moving observer Σ' this event lasts

$$\Delta t' = \frac{\Delta t}{\sqrt{1 - v^2/c^2}}. \tag{3.13}$$

A moving clock runs slower than one at rest, any clock runs fastest for an observer who is at rest with respect to it.

Length contraction When we measure the length of a rod at rest, the times t_A and t_B at which we look at the two endpoints $x_A = 0$ and $x_B = L$ are unimportant, its length is always $L = \Delta x$. For a moving observer Σ' this is different: since the rod is moving in his system of reference, he has to take care to determine its two endpoints *simultaneously*! So when using the relations

$$\begin{aligned} x'_A &= \frac{-vt_A}{\sqrt{1 - v^2/c^2}}, \quad & ct'_A &= \frac{ct_A}{\sqrt{1 - v^2/c^2}}, \\ x'_B &= \frac{L - vt_B}{\sqrt{1 - v^2/c^2}}, \quad & ct'_B &= \frac{ct_B - vL/c}{\sqrt{1 - v^2/c^2}}, \end{aligned} \tag{3.14}$$

he has to set $t'_A = t'_B$. Choosing $t'_A = 0$, this amounts to $t_A = 0$, $t_B = vL/c^2$, and thus to $x'_A = 0$, $x'_B = L\sqrt{1 - v^2/c^2}$, or to

$$\Delta x' = \Delta x \sqrt{1 - v^2/c^2}. \tag{3.15}$$

A moving rod is shorter than one at rest, a rod is longest for an observer at rest with respect to it.

3.4 Two thought experiments

The two effects explained above, the time dilatation and the length contraction, are experimentally well confirmed. To get a better understanding of them, we will now discuss in some detail two gedanken (thought) experiments.

3.4.1 A rod moving through a tube

We take a rod of length $2L$, and a tube of length L (both measured at rest), see Fig. 3.3.

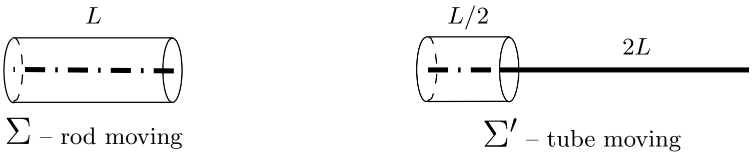

Fig. 3.3. Rod and tube.

System Σ (*Tube at rest, rod moving*) The length of the tube is L. If the rod moves with velocity $v = c\sqrt{3}/2$, application of (3.15) yields $2L\sqrt{1 - v^2/c^2} = L$ as the length of the rod; if it moves through the tube, it just fits in!

System Σ' (*Rod at rest, tube moving*) The rod is four times as long as the tube, it never can fit into the tube!

How can the two results both be true? Observer Σ' will state that Σ did not measure the position of the rod's endpoints simultaneously: Σ determined the position of its tip when it had already reached the end of the tube, and then waited until the end of the rod just entered the tube.

3.4.2 The twin paradox

Imagine a pair of twins; one is travelling around in space with a high velocity, the other just stays on Earth.

System Σ (*Earth at rest*) The travelling twin, assumed to have a constant velocity (except at the turning point), experiences a time dilatation, his biological clock runs slower; when coming back to Earth he is younger.