General relativity I and II

Spring 2025

Syksy Räsänen

Contents

Preface

These are the lecture notes of the courses General relativity I and II at the University of Helsinki. They owe a heavy debt to the lecture notes of Hannu Kurki-Suonio, which were in turn influenced by Sean Carroll's lecture notes and his book *Spacetime and Geometry*. Most of the figures are from Hannu Kurki-Suonio. I thank Fernando Bracho Blok for LaTeXing parts of these lecture notes. Any errors are my responsibility; let me know when you catch them.[1]

General relativity (GR) is one of the two fundamental theories we currently have. The other is quantum field theory, in particular its application in the Standard Model of particle physics. Both of these pillars grew from addressing the shortcomings of Newtonian mechanics in the 20th century. Quantum physics was developed in bits and pieces in close interaction between theory and observations as the classical description of matter was found to be inadequate. In contrast, observations played second fiddle to theoretical and mathematical arguments in the development of the theory of relativity. Special relativity (SR) was discovered by Albert Einstein in 1905, and GR was unveiled by him and David Hilbert in 1915, after a long process that involved also other collaborators. In contrast to the piecewise development of quantum theory, the structure of SR and GR was laid out almost in its entirety when they were first uncovered.

Quantum field theory describes matter and its non-gravitational interactions, and GR covers spacetime and its interaction with matter, i.e. gravity. GR tells how spacetime behaves when it contains given kind of matter, but it is agnostic on what sort of matter exists. In the 1940s, quantum mechanics was unified with SR in Quantum Electrodynamics, the first quantum field theory. The unification

[1] Apparently, some early 20th century Arabic books included a disclaimer along the lines of "this book contains errors, but they can be easily picked up by an intelligent reader". That guideline is not unsuitable for these lecture notes.

of quantum physics with GR remains one of the greatest open questions in physics. It has been achieved to a limited extent in the theory of cosmic inflation, where linear perturbations of spacetime are quantised, and their quantum properties have been calculated and compared to observations of the cosmic microwave background and large scale structure with great success. We will consider the intersection of quantum physics and GR only in passing when we discuss the Hawking radiation of black holes; otherwise we stick to classical (i.e. non-quantum) physics.

The essence of GR is that gravity is an aspect of spacetime curvature, and that locally spacetime is flat. The first chapter is therefore a refresher on flat spacetime, i.e. SR formulated in terms of spacetime. We will also go through, which was historically important in the development of SR, and will provide a useful comparison to GR. In the first chapter we will be rather informal. In chapter 2 we introduce manifolds to describe a general spacetime, and will be more careful with notation and definitions. In chapter 3 we bring in spacetime curvature. In chapter 4 we discuss how curvature is generated by matter and consider the Newtonian limit. We conclude part I in chapter 5, where we consider the Schwarzschild solution and calculate the precession of the perihelion of Mercury and the bending of light by the Sun. We open part II in chapter 6 with GR in the action formulation, and discuss reducing the number of assumptions in the theory. The rest of part II is dedicated to particularly important solutions of the equations of motion. In chapter 7 we discuss black holes, in chapter 8 we consider perturbation theory around Minkowski space and gravitational waves, and in chapter 9 we conclude with symmetries and cosmology.

1 Special relativity

1.1 Spacetime notation

We assume familiarity with Newtonian physics and SR. The structure of SR is rather simple, the most difficult part is unlearning Newtonian ideas about space and time. Relativistic spacetime consists of **events**, corresponding to spacetime points that can be labelled with an integer number d of real numbers, called **coordinates.**[2] There are infinitely many different ways to assign numbers to spacetime points, i.e. different **coordinate systems**.

We label the coordinates x^α, where $\alpha = 0, 1, \ldots, d-1$. We will mostly consider four-dimensional spacetime, $d = 4$, where there is one time direction and three spatial directions. (We will soon say what space and time directions mean.) The time coordinate is assigned the number $\alpha = 0$, and it is also denoted $x^0 = ct$, where c is the speed of light (we discuss the physical meaning of c below). We use Greek letters for spacetime indices and Latin indices for spatial indices, which run from 1 to $d-1$. Just as the components of a vector in Newtonian physics are labelled x^i, the coordinates x^α are the components of a **vector** in spacetime. Vectors in four-dimensional spacetime are sometimes called four-vectors for emphasis. We use the following notations for x^α (and similarly for other vectors):

$$x^\alpha = (x^0, x^1, \ldots, x^{d-1}) = (x^0, x^i) = (x^0, \vec{x}) \ . \tag{1.1}$$

[2] When we come to GR in the next chapter, we will be more precise about the structure of spacetime and the meaning of coordinates.

In the arguments of functions we often drop the index and write $f(x)$ instead of $f(x^\alpha)$. The notation x^α refers both to the set of components and to the specific component numbered α, as is also the case for x^i. Once we have understood the relation between the components and the vector, we will even lazily refer to x^α as the vector. We denote three-vectors with an arrow on top, $\vec{x}$, and four-vectors with an underline, $\underline{x}$. In SR, we mostly stick to **Cartesian coordinates** that are familiar from elementary treatments of Newtonian physics. We could also consider SR in general coordinates, but this would require introducing half of the machinery of GR. We begin with a brief overview of Newtonian mechanics in terms of symmetry, and then approach SR the same way.

1.2 Newtonian spacetime

1.2.1 Illustrating the structure of Newtonian spacetime

In Newtonian mechanics, a physical system consists of point particles moving in space. (Extended objects can be understood as composites of point particles.) The state of a system is fully determined by the position x^i of each point particle as a function of time t: the central object is the particle trajectory $x^i(t)$. The line drawn in spacetime (direct product of space and time) by a particle is called its **worldline**. The structure of Newtonian spacetime is illustrated in terms of worldlines in figure 1.

In Newtonian spacetime, spatial slices of constant time are **absolute**, i.e. they are the same for all observers. In other words, Newtonian physics has **absolute space**. This also implies that simultaneity is absolute: all observers agree on whether two points lie on the same time slice. There are three kinds of directions in spacetime, illustrated in figure 2. **Spacelike** directions are parallel to the slices of constant time, **future timelike** directions point up from the slices of constant time and **past timelike** directions point down from the slices of constant time. Particle trajectories always move up (i.e. are future timelike), but are otherwise arbitrary[3]. In particular, the slope of the line, corresponding to the velocity, can have any value.

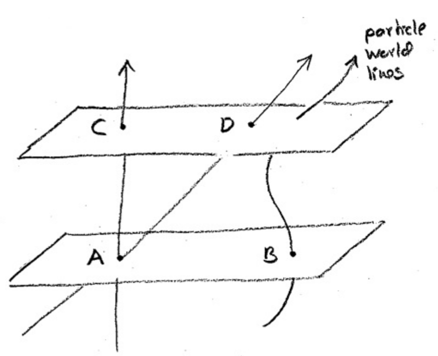


Figure 1: Newtonian time slices. Events A and B are simultaneous, as are C and D. All depicted worldlines are future timelike.

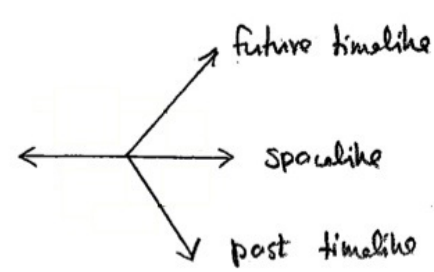


Figure 2: The three kinds of directions in Newtonian spacetime.

[3]Apart from being continuous and sufficiently differentiable.

1.2.2 Symmetries of Newtonian spacetime

Newtonian spacetime is symmetric under the coordinate transformations

$$t \rightarrow t'(t) = t + A \tag{1.2}$$
$$x^i \rightarrow x'^i(t,\vec{x}) = R^i{}_j x^j + A^i + v^i t \ , \tag{1.3}$$

where A is a constant, $R^i{}_j$ is a constant rotation matrix (we discuss this in more detail in section 1.3.3), and A^i and v^i are constant vectors. We have adopted the **Einstein summation convention**, according to which any index that appears once upstairs and once downstairs is summed over, unless otherwise noted. This applies equally to Latin and Greek indices. Such an index is called a summation index, internal index, or dummy index. So $R^i{}_j x^j \equiv \sum_{j=1...3} R^i{}_j x^j$. The spacetime is also symmetric under time reversal, $t \rightarrow -t$, and spatial reflections (also called **parity transformations**), $x^i \rightarrow -x^i$ (for any or all of the spatial directions).

Symmetry under coordinate transformations means that the spacetime before and after the transformation describes the same physical situation. Quantities that do not remain invariant are **relative**, while those that are invariant are called **absolute**. The transformations (1.2) and (1.3) of involve four separate symmetries.

Invariance under constant time shifts means that the value of the time coordinate is not absolute. However, time intervals are absolute (and hence simultaneity is absolute), so we say that time is absolute. Invariance under constant translations means that there is no absolute position in space. However, spatial intervals are absolute, i.e. lengths are absolute. Invariance under spatial rotations means there is no absolute direction in space. However, relative angular positions are absolute. Overall, we say that space is absolute as the symmetry transformations map points on a slice of constant time to points on the same slice. The time-dependent shift by $v^i t$ is called the **Galilei transformation**. Invariance under it, called **Galilei symmetry**, means that there is no absolute velocity. However, velocity differences are absolute as far as their amplitude is concerned. This includes change in velocity over time, so the amplitude of acceleration is absolute. The direction of velocity differences and of acceleration are relative, as they change with rotation.

The laws of physics have to respect the above symmetries. Consider Newton's second law:

$$a^i \equiv \frac{\mathrm{d}^2 x^i}{\mathrm{d}t^2} = \frac{1}{m} F^i \ , \tag{1.4}$$

where a^i is acceleration, F^i is force, and m is the mass of the particle. Under the symmetry transformations, the direction of the acceleration and of the force changes. However, the equation retains its shape under spatial rotations, because both sides are vectors and transform in the same way under rotations. As the transformed quantities satisfy the same equation, the equation is called **covariant** under rotations, rather than invariant. (Although physicists are somewhat sloppy with language, and sometimes the term invariant is used here as well.) Also, equation (1.4) is invariant under the constant translations of time and space, the Galilei transformation, and time reversal, as long as the force F^i is invariant. It is covariant under parity transformations if the components F^i switch their sign under $x^i \rightarrow -x^i$. For observers moving at constant velocity (and only for observers moving at constant velocity) the equations of motion take the form (1.4). Such observers are called

inertial. If an observer is not inertial, there are new terms contributing to (1.4), such as the Coriolis effect. We will discuss this in more detail when we come to GR, whose formalism will make it easy to treat such effects.

An example of a force that is covariant under the above symmetries is Newton's law of gravity, which states that the force exerted by particle 1 on particle 2 is

$$\vec{F}_{12} = -G_{\rm N} m_1 m_2 \frac{\vec{x}_2 - \vec{x}_1}{|\vec{x}_2 - \vec{x}_1|^3} \ , \tag{1.5}$$

where $G_{\rm N}$ is Newton's constant and m_n are the masses of the particles. In index notation, (1.5) reads

$$F^i_{12} = -G_{\rm N} m_1 m_2 \frac{\Delta x^i}{|\Delta\vec{x}|^3} \ , \tag{1.6}$$

where we have introduced the coordinate difference $\Delta x^i \equiv x^i_2 - x^i_1$. As the force depends only on the separation of the particles and not their absolute positions, it is invariant under spatial translation and the Galilei transformation. As the force is independent of time, it is invariant under time translation and time reversal. Its components change sign under a parity transformation.

Note that while symmetries (1.2) and (1.3) constrain the equation of motion (1.4) and the force (1.5), they do not determine them. For example, the symmetry would allow the gravitational force to have $|\Delta\vec{x}|^5$ in the denominator instead of $|\Delta\vec{x}|^3$ – in fact any functional dependence on $|\Delta\vec{x}|$ is allowed. In the case of GR, the laws of physics are covariant under a larger group of transformations (although these are not associated with symmetries of the spacetime; the construction is more subtle), and we will see that this determines the relativistic law of gravity almost uniquely.

In the Newtonian case, let us restrict ourself to space, leaving time aside. The fact that the symmetry under rotations and translations leaves lengths invariant can be expressed by saying that the quantity[4]

$$(\Delta s)^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2 = \delta_{ij} \Delta x^i \Delta x^j \tag{1.7}$$

is invariant; we have labelled the Cartesian coordinates as $x^i = (x, y, z)$. We could reverse this reasoning, starting from the assumption that the interval (1.7) is invariant and looking for the transformations $x^i \to x'^i(\vec{x})$ that leave it invariant. Keeping to linear transformations, we would then arrive at the six-parameter group of spatial rotations and translations (plus reflections).

When we rotate space, the intervals Δx, Δy and Δz change, but the combination (1.7) stays the same, as illustrated in figure 3:

$$(\Delta s)^2 = (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 = (\Delta x')^2 + (\Delta y')^2 + (\Delta z')^2 \ . \tag{1.8}$$

This is why we think of three-dimensional space instead of a set of three one-dimensional spaces. In SR, this idea is extended to the four-dimensional spacetime.

The statement (1.7), Pythagoras' law, expresses the fact that Newtonian space is **Euclidean**. Space is Euclidean precisely when the distance Δs between points with coordinates x^i_1 and x^i_2 is given by (1.7). We say that δ_{ij} is the **metric** (more

[4]The symbol δ_{ij} is the **Kronecker delta**, defined to be 1 if $i = j$ and 0 otherwise. The position of its indices carries no meaning, $\delta^i{}_j = \delta_i{}^j = \delta_{ij} = \delta^{ij}$.

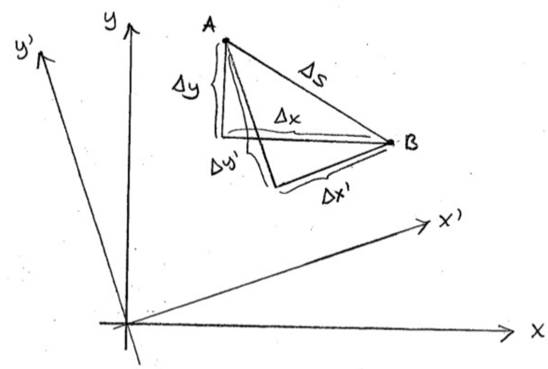


Figure 3: Rotation on the xy-plane.

precisely, components of the metric) of Euclidean space, in Cartesian coordinates. The metric depends on the chosen coordinate system. For example, if we used spherical rather than Cartesian coordinates, both the components of the vectors and the metric would be different, but $(\Delta s)^2$ would be the same. If the distance is not given by (1.7) or a coordinate transformation thereof (but the notion of distance is still defined), the space is **non-Euclidean**. In SR, we encounter the simplest (four-dimensional) example of a non-Euclidean space(time).

1.3 Minkowski space

1.3.1 The Minkowski metric

The spacetime of SR is called Minkowski space. (A misleading name, as it is a spacetime, not a space!) Just as Euclidean space is defined by the invariance of the spatial interval (1.7), Minkowski space is defined by the invariance of the following spacetime interval between two events (using the Cartesian coordinates $x^\alpha = (ct, x, y, z)$):

$$(\Delta s)^2 \quad = \quad -c^2(\Delta t)^2 + (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 \equiv \eta_{\alpha\beta}\Delta x^\alpha \Delta x^\beta \ , \qquad (1.9)$$

where we have introduced the **Minkowski metric** in Cartesian coordinates, $\eta_{\alpha\beta} \equiv \mathrm{diag}(-1, 1, 1, 1)$.[5] The difference between Euclidean geometry and the geometry of Minkowski space is entirely due to the minus sign in front of $c^2(\Delta t)^2$. Depending on the relative size of the time separation $c\Delta t$ and the spatial separation, the interval $(\Delta s)^2$ can be positive, negative or zero. If $(\Delta s)^2$ is positive or zero, the distance is $\sqrt{(\Delta s)^2}$. If $(\Delta s)^2$ is negative or zero, the distance is $\sqrt{-(\Delta s)^2}$. In this case, $(\Delta s)^2$ should not be thought of as the square of anything, despite the notation.

If $(\Delta s)^2 > 0$, the separation of the events is **spacelike**. If $(\Delta s)^2 < 0$, the separation is **timelike**.[6] If $(\Delta s)^2 = 0$, the separation is **lightlike**, also called **null**. The set of points whose separation from a given point P is null (i.e. whose distance to P is zero) form the **lightcone** at that point.

[5] Let the reader beware: some authors, especially in particle physics, prefer to define $\eta_{\alpha\beta}$ with the opposite overall sign. Also note that the metric used in SR and GR is not a metric in the sense that the term is usually used in mathematics, as it is not positive-definite.

[6] Expressed in a way that is independent of the sign convention of $\eta_{\alpha\beta}$, the spacetime interval is timelike if the time interval is longer than the spatial interval, and spacelike if the reverse is true.

There are correspondingly five kinds of directions in spacetime. Straight lines from P to a point inside lightcone are future timelike if oriented up and past timelike if oriented down. A point at timelike separation from P is in the future of P if its time coordinate is larger than that of P, and in the past of P if its time coordinate is smaller than that of P. Massive particles and objects move on future timelike lines, i.e. worldlines. Directions along the lightcone are called **future lightlike** (also called **future null**) or **past lightlike** (**past null**), depending on whether they are in the positive or negative time direction, respectively. Massless particles and electromagnetic waves travel on such lines, called lightlike lines (or equivalently null lines). A line from P to a point outside the cone is spacelike, and there is no absolute ordering into the past or future of P for such points. Particles and objects cannot move on spacelike lines. Observers whose worldlines have a different tilt have different time directions: time is relative, as is space. There is no absolute simultaneity of different events. Also, the time ordering of events that are not connected by timelike or lightlike lines is not absolute. Unlike in Newtonian spacetime, there are no absolute time slices. These features are illustrated in figures 4 and 5.

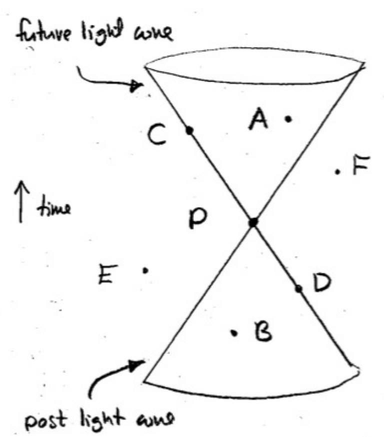


Figure 4: Intervals in spacetime.

- $(\Delta s)^2 < 0$: timelike separated from P (points A and B).
- $(\Delta s)^2 = 0$: null separated from P (points C and D).
- $(\Delta s)^2 > 0$: spacelike separated from P (points E and F).
- Events A and C are in the future of P.
- Events B and D are in the past of P.
- The time ordering between P and events E and F depends on the observer.

The speed of light c has a similar role in SR as infinite speed in Newtonian physics: it cannot be exceeded. Just as in Newtonian physics going faster than at infinite speed would mean that the time taken to travel between two points would be less than zero, in SR moving faster than the speed of light would correspond to moving backward in time. Heuristically, for $c \to \infty$ the SR lightcone flattens

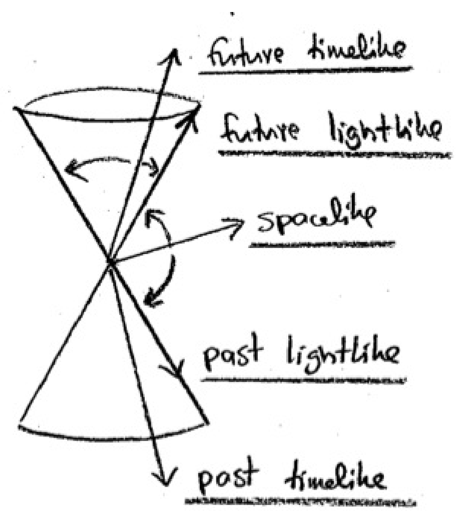


Figure 5: The five kinds of directions in Minkowski space.

into a plane, and spacetime structure reduces to Newtonian physics with absolute time slices, as illustrated in figure 6. However, we have to be careful with limits like this, because c is a dimensional quantity, so it can only be large or small with respect to another quantity of the same dimension. This argument is best regarded as a geometric heuristic for why SR looks like Newtonian physics for velocities much smaller than c.

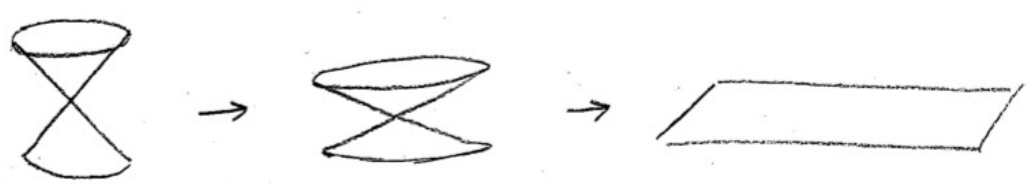


Figure 6: Flattening of the lightcone.

1.3.2 Natural units

The value of c has no physical meaning, it is a piece of historical baggage that just relates the units of length and time. Due to the fact that velocities in the everyday environment relevant for our survival are much smaller than the speed of light, our brains have evolved to view time and space as fundamentally different kinds of entities, so we are used to measuring them with different units. In SR, time and space directions are not absolute but mix, so it is unnecessary and unwieldy to use different units for them. Doing so would be analogous to measuring spatial distance along the surface of the Earth with nautical miles and in the radial direction in fathoms: c has the role of the conversion factor that tells how many fathoms there are in a nautical mile.

We use **natural units**, where $c = 1$. In SI units $c = 299\ 792\ 458$ m/s (exactly), so this means that 1 s = 299 792 458 m, and 1 second = 1 light second, 1 year =

1 light year, and so on. Velocity is a dimensionless quantity, which is smaller than one for massive objects.

In natural units, the reduced Planck constant is also unity, $\hbar = h/2\pi = 1$. In SI units $h = 6.626\ 070\ 15\ \times 10^{-34}$ Js (exactly), so in natural units the dimensions of mass, energy, momentum and wavenumber are the same, and equal to the dimension of 1/time or 1/distance. In natural units Boltzmann's constant is also unity, $k_B = 1$, so temperature and energy are measured in the same units.

1.3.3 Poincaré transformations

Taking the coordinate difference in (1.9) to be infinitesimal, we can write the interval in the form commonly used in SR,

$$\mathrm{d}s^2 \quad = \quad -\mathrm{d}t^2 + \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 = \eta_{\alpha\beta}\mathrm{d}x^\alpha \mathrm{d}x^\beta \ . \tag{1.10}$$

The infinitesimal quantity $\mathrm{d}s^2$ is called the **line element**.

Taking the line element (1.10) as fundamental, we now look for transformations that leave it invariant. We demand that the coordinate differences transform in the same way everywhere in spacetime (i.e. there are no preferred points in spacetime: spacetime is **homogeneous**), so the transformations have the form

$$x^\alpha \to x'^\alpha(x) = \Lambda^\alpha{}_\beta x^\beta + A^\alpha \ , \tag{1.11}$$

where $\Lambda^\alpha{}_\beta$ and A^α are constants. (We use x in the argument to collectively denote all spacetime coordinates, as mentioned earlier.) An infinitesimal coordinate difference then transforms as

$$\mathrm{d}x^\alpha \to \mathrm{d}x'^\alpha = \Lambda^\alpha{}_\beta \mathrm{d}x^\beta \ , \tag{1.12}$$

and the interval transforms as

$$\mathrm{d}s^2 = \eta_{\alpha\beta}\mathrm{d}x^\alpha \mathrm{d}x^\beta \to \eta_{\alpha\beta}\mathrm{d}x'^\alpha \mathrm{d}x'^\beta = \eta_{\alpha\beta}\Lambda^\alpha{}_\gamma \Lambda^\beta{}_\delta \mathrm{d}x^\gamma \mathrm{d}x^\delta \ . \tag{1.13}$$

The invariance of $\mathrm{d}s^2$ therefore implies

$$\eta_{\gamma\delta} = \eta_{\alpha\beta}\Lambda^\alpha{}_\gamma \Lambda^\beta{}_\delta \ . \tag{1.14}$$

Written in matrix form, this reads $\Lambda^T \eta \Lambda = \eta$, where T denotes transpose. (Recall that in matrix multiplication, the rightmost index of the matrix on the left is summed over with the leftmost index on the right, $(AB)^i{}_j = \sum_k A^i{}_k B^k{}_j$.) The solutions to this equation are the **Lorentz matrices** that are a representation of the **Lorentz transformations**, which form the **Lorentz group**. The Lorentz group is denoted $O(1,3)$. Excluding reflections and time reversal by demanding $\Lambda^0{}_0 > 0, \det\Lambda > 0$, we get the **proper orthochronous** Lorentz matrices and transformations. They form the **proper orthochronous Lorentz group**, denoted $SO(1,3)^\uparrow$.

As the Lorentz transformations are described by the 4×4 matrices $\Lambda^\alpha{}_\beta$ (note how loosely we identify the matrix and its components), they have 16 components. As (1.14) is a symmetric 4×4 matrix equation, it has 10 independent components, leaving the Lorentz matrices with 6 independent components. So the Lorentz group is a six-parameter group. Including spacetime translations, we get the **Poincaré**

	Euclidean space	Minkowski space
line element	$\mathrm{d}s^2 = \delta_{ij}\mathrm{d}x^i\mathrm{d}x^i$	$\mathrm{d}s^2 = \eta_{\alpha\beta}\mathrm{d}x^\alpha\mathrm{d}x^\beta$
symmetry transformation	$x^i \to R^i{}_j x^j + A^i$	$x^\alpha \to \Lambda^\alpha{}_\beta x^\beta + A^\alpha$
condition for transformation matrix	$R^T R = \mathbb{1}$	$\Lambda^T \eta \Lambda = \eta$
full symmetry group	$E(3)$	$P(1,3)$
symmetry group without translations	$O(3)$	$O(1,3)$
symmetry group w/o translations, reflections or time reversal	$SO(3)$	$SO(1,3)^\uparrow$

Table 1: Comparison of 3d Euclidean space and 4d Minkowski spacetime in terms of symmetry.

transformations written in (1.11), which form the 10-parameter **Poincaré group** $P(1,3)$.

We could have followed the same route in the case of three-dimensional Euclidean space. To do so, we would replace the Poincaré transformations (1.11) with the spatial rotations and translations $x^i \to x'^i = R^i{}_j x^j + A^i$, and the Minkowski metric $\eta_{\alpha\beta}$ with the Euclidean metric δ_{ij}. This would give the condition $R^T R = \mathbb{1}$. The matrices that satisfy this condition are a representation of the **orthogonal group** $O(3)$. Together with the translations, they form the **Euclidean group** $E(3)$. Excluding reflections by demanding $\det R > 0$ gives the **special Euclidean group** $SE(3)$. Excluding translations then gives the **rotation group** $SO(3)$. Euclidean and Minkowski space are summarised in terms of symmetries in table 1.

In Cartesian coordinates, the rotation matrix on the xy-plane (one solution of the equation $R^T R = \mathbb{1}$) reads

$$R^i{}_j = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} . \tag{1.15}$$

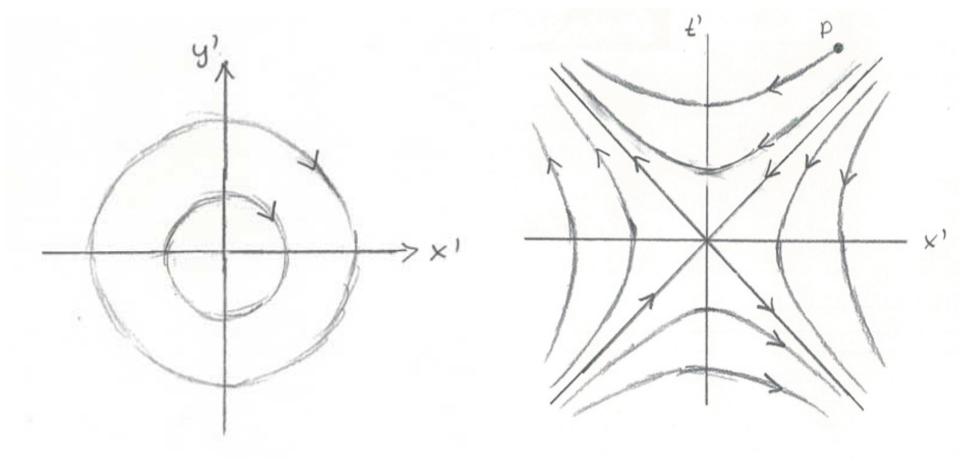
We see that the transformation parameter is restricted to the range $0 \leq \theta < 2\pi$. Equivalently, we can say that rotations are periodic, and take $-\infty < \theta < \infty$. Either way, we see that the rotation group is compact.

The other solutions $R^i{}_j$ of the equation $R^T R = \mathbb{1}$ that satisfy the condition $\det R > 0$ can be written similarly as rotations on the xy- or the zx-plane, or as a product of these three rotations. The three parameters of $SO(3)$ can be taken to be the rotation angles on the three two-dimensional planes on which the transformations operate. The fact that the group is compact reflects the fact that in the combination $(\Delta s)^2 = (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2$ any one coordinate difference can be reduced or increased only by a finite amount when $(\Delta s)^2$ is held constant.

Returning to Minkowski space, rotations are a subgroup of the Lorentz group: they are those transformations that leave the time interval untouched,

$$\Lambda^\alpha{}_\beta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{1.16}$$

The remaining three Lorentz transformations act on the tx-, ty- and tz-planes, changing the time interval Δt and one of the length intervals in $(\Delta s)^2 = -(\Delta t)^2 +$



(a) Lorentz transformation on the xy-plane (a rotation).

(b) Lorentz transformation on the tx-plane (a boost).

Figure 7: Illustrating the orbits of Lorentz transformations.

$(\Delta x)^2+(\Delta y)^2+(\Delta z)^2$. Because of the minus sign, the orbit of the transformations is not a circle but a hyperbola, so we get hyperbolic instead of trigonometric functions, as illustrated in figure 7. For example, the transformations on the tx-plane read

$$\Lambda^\alpha{}_\beta = \begin{pmatrix} \cosh\psi & -\sinh\psi & 0 & 0 \\ -\sinh\psi & \cosh\psi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} , \tag{1.17}$$

where the transformation parameter ψ is called the **rapidity**. It has the range $-\infty < \psi < \infty$, and the transformations are non-periodic.

Exercise: Show that the matrices $\Lambda^\alpha{}_\beta$ in (1.16) and (1.17) satisfy (1.14).

The Lorentz transformations that involve time are called **boosts**, for reasons that will become clear in section 1.3.6. Like rotations, they form a subgroup of the Lorentz group. Unlike the rotation subgroup, the boost subgroup is non-compact. This corresponds to the fact that $(\Delta t)^2$ and, say, $(\Delta x)^2$ can both be increased without limit while keeping their difference constant. The 10 parameters of the Poincaré group correspond to the transformation parameters on the six planes on which rotations and boosts operate plus the translation parameters along the four coordinate axes on which the translations happen. (In d dimensions, we would have $d(d-1)/2+d = d(d+1)/2$ parameters.) Rotations and boosts on different planes do not commute among themselves nor with each other, so the Lorentz group (like the rotation group) is non-Abelian.

Lorentz boosts mix time and space, so

$$\Delta t \neq \Delta t'$$
$$(\Delta x)^2+(\Delta y)^2+(\Delta z)^2 \neq (\Delta x')^2+(\Delta y')^2+(\Delta z')^2 .$$

Only the spacetime interval (1.9) remains invariant. This is why it is useful to think in terms of four-dimensional spacetime instead of a set of three-dimensional spaces stacked together.

1.3.4 Dot product and raising and lowering indices

In addition to defining distances in spacetime, the metric is used to define the dot product of two vectors, and to lower and raise indices.[7] For a four-vector with components U^α, we define

$$U_\alpha \equiv \eta_{\alpha\beta} U^\beta \ . \tag{1.18}$$

The object with components U_α is called a **covariant vector**[8], and the object with components U^α is called a **contravariant vector**.

The **inverse metric**, with components $\eta^{\alpha\beta}$, is defined as the inverse matrix of the matrix whose components are $\eta_{\alpha\beta}$; it is described by the same matrix, $\eta^{\alpha\beta} = \mathrm{diag}(-1,1,1,1)$. So $\eta_{\alpha\gamma}\eta^{\gamma\beta} = \delta_\alpha{}^\beta$. Indices are raised with the inverse metric as

$$U^\alpha \equiv \eta^{\alpha\beta} U_\beta \ . \tag{1.19}$$

The dot product between vectors $\underline{U}$ and $\underline{V}$ is defined as

$$\underline{U} \cdot \underline{V} \equiv \eta_{\alpha\beta} U^\alpha V^\beta = \eta^{\alpha\beta} U_\alpha V_\beta = U^\alpha V_\alpha = U_\alpha V^\alpha \ . \tag{1.20}$$

Under Poincaré transformations, contravariant vectors transform with the Lorentz matrix and covariant vectors with the inverse matrix, so the dot product (1.20) is invariant:

$$\begin{aligned} U^\alpha &\to U'^\alpha = \Lambda^\alpha{}_\beta U^\beta \\ U_\alpha &\to U'_\alpha = (\Lambda^{-1})^\beta{}_\alpha U_\beta \ . \end{aligned} \tag{1.21}$$

The rule is easy to remember: up indices transform with the Lorentz transformation, down indices transform with the inverse transformation. By definition, $\Lambda^\alpha{}_\gamma (\Lambda^{-1})^\gamma{}_\beta = \delta^\alpha{}_\beta$.

In the case $\underline{U} = \underline{V}$, the dot product (1.20) gives the squared **norm** $\|U\|$ of the vector:

$$\|U\|^2 \equiv \underline{U} \cdot \underline{U} = \eta_{\alpha\beta} U^\alpha U^\beta \ . \tag{1.22}$$

If $\underline{U} \cdot \underline{U} < 0$, the vector $\underline{U}$ is timelike. If $\underline{U} \cdot \underline{U} = 0$, the vector $\underline{U}$ is lightlike, also called null. If $\underline{U} \cdot \underline{U} > 0$, the vector $\underline{U}$ is spacelike.

1.3.5 Spatial and timelike distances

Let us now look at distances in spacetime. Consider an arbitrary differentiable curve with coordinates $x^\alpha(\lambda)$, where λ is a real number that parametrises the position on the curve. We have (see figure 8)

$$\mathrm{d}x^\alpha = \frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda} \mathrm{d}\lambda \ . \tag{1.23}$$

[7] We will savor the topics of this subsection in more detail in the next chapter, this is just a quick snack.

[8] This meaning of the term covariant is unrelated to its use to describe equations that retain their form under transformations.

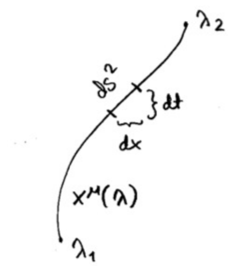


Figure 8: The line element of a parametrized curve $x^\mu(\lambda)$.

We define a section of a curve to be timelike, null, or spacelike if the tangent vector $\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda}$ everywhere in the section is timelike, null, or spacelike, respectively. Most physically relevant curves are everywhere timelike, or everywhere null, or everywhere spacelike. In the case of a timelike curve that corresponds to particle motion, $x^\alpha(\lambda)$ generalises the Newtonian particle trajectory $x^i(t)$.

The infinitesimal **proper time** interval along a timelike curve (i.e. a worldline) is

$$\mathrm{d}\tau = \sqrt{-\mathrm{d}s^2} = \sqrt{-\eta_{\alpha\beta}\mathrm{d}x^\alpha\mathrm{d}x^\beta} = \sqrt{-\eta_{\alpha\beta}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda}}\mathrm{d}\lambda\ . \tag{1.24}$$

Note that λ does not have to be the proper time along the curve, although that is often a convenient choice. The proper time from point A to B along the curve is

$$\Delta\tau = \int_{\lambda_A}^{\lambda_B} \sqrt{-\eta_{\alpha\beta}\mathrm{d}x^\alpha\mathrm{d}x^\beta} = \int_{\lambda_A}^{\lambda_B} \mathrm{d}\lambda\sqrt{-\eta_{\alpha\beta}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda}}\ . \tag{1.25}$$

A straight line maximises the proper time between two events whose separation is timelike. One example of this is the "twin paradox", where two observers start with the same age, but one undergoes acceleration (i.e. travels in a line that is not straight) and is thus younger when the two meet again.

Exercise. Consider the twins Alice and Betty. Alice stays on Earth. Betty leaves Alice and travels to Alpha Centauri (distance 4 light years) at the speed $v = 0.8$, turns around, and returns at the same speed. How much have Alice and Betty aged when they meet again? Draw a spacetime diagram of the worldlines of Alice and Betty
a) in the frame of Alice (K),
b) in the frame of Betty traveling towards αCen (K'),
c) in the frame of Betty returning (K'').

Similarly, the infinitesimal **proper length** along a spacelike curve is

$$\mathrm{d}s = \sqrt{\mathrm{d}s^2} = \sqrt{\eta_{\alpha\beta}\mathrm{d}x^\alpha\mathrm{d}x^\beta} = \sqrt{\eta_{\alpha\beta}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda}}\mathrm{d}\lambda\ , \tag{1.26}$$

and the proper length along a spacelike curve from A to B is

$$s = \int_{\lambda_A}^{\lambda_B} \sqrt{\eta_{\alpha\beta}\mathrm{d}x^\alpha\mathrm{d}x^\beta} = \int_{\lambda_A}^{\lambda_B} \mathrm{d}\lambda\sqrt{\eta_{\alpha\beta}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda}}\ . \tag{1.27}$$

A straight line minimises the proper length between two events whose separation is spacelike.

For a null curve $\mathrm{d}s^2 = 0$ everywhere, so its proper length is zero.

1.3.6 Lorentz boosts and velocity

So far, we have not said anything about the relation of Lorentz boosts and velocity. From the spacetime point of view, the core assumption of SR is homogeneity of spacetime. The relation of the boosts to velocities is then a result rather than an assumption, in contrast to elementary treatments of SR. Let us see how the velocity comes into play from the spacetime point of view.

Consider a straight worldline with coordinates $x^\alpha(\lambda)$. The components of the vector $\underline{u}$ tangent to the curve are

$$u^\alpha = \frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda} \ . \tag{1.28}$$

Let us use the proper time of the observer as the parameter along the curve, $\lambda = \tau$. We first find the relation between four-velocity and coordinate three-velocity. The norm of $\underline{u}$ is

$$\underline{u} \cdot \underline{u} = \eta_{\alpha\beta} u^\alpha u^\beta \ . \tag{1.29}$$

As the norm is independent of the coordinates, we can evaluate it in **comoving coordinates**. They are coordinates attached to the observer, so that $t(\lambda) = \tau, x^i(\lambda) = 0$. Then $u^\alpha = \delta^{\alpha 0}$, so we see that $\underline{u} \cdot \underline{u} = -1$. We define the coordinate three-velocity as

$$v^i \equiv \frac{\mathrm{d}x^i}{\mathrm{d}t} \ . \tag{1.30}$$

Note that v^i are not the spatial components of a four-vector. Instead,

$$u^\alpha = \left(\frac{\mathrm{d}t}{\mathrm{d}\tau}, \frac{\mathrm{d}t}{\mathrm{d}\tau} v^i \right) \equiv \gamma(1, v^i) \ , \tag{1.31}$$

where we have introduced the **Lorentz factor** $\gamma \equiv \frac{\mathrm{d}t}{\mathrm{d}\tau}$. From the condition $\underline{u} \cdot \underline{u} = -1$ we get $\gamma = (1 - v^2)^{-1/2}$, where $v^2 \equiv \delta_{ij} v^i v^j$.

Now that we have the coordinate three-velocity, let's consider the Lorentz boosts (1.11). Under a Lorentz transformation, the four-velocity (like other four-vectors) transforms as

$$u^\alpha \to u'^\alpha = \Lambda^\alpha{}_\beta u^\beta \ . \tag{1.32}$$

Let us take the original coordinates to be comoving with the observer, $u^\alpha = \delta^{\alpha 0}$. A Lorentz boost on the tx-plane then gives, applying (1.17),

$$u^\alpha = (1, 0, 0, 0) \to (\cosh\psi, -\sinh\psi, 0, 0) \ . \tag{1.33}$$

Comparing to (1.31), we have $\gamma = \cosh\psi, \gamma v = -\sinh\psi$. The minus sign comes from the fact that as the new coordinates are moving with velocity v with respect to the old coordinates, the old coordinates are moving with velocity $-v$ with respect to the new coordinates. Switching the sign (this is a matter of convention), we get $\psi = \text{artanh} v$, and the Lorentz boost matrix (1.17) can be written as

$$\Lambda^\alpha{}_\beta = \begin{pmatrix} \gamma & -\gamma v & 0 & 0 \\ -\gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{1.34}$$

The Lorentz boost hence reads

$$\begin{aligned} t &\to t' = \gamma(t - vx) \\ x &\to x' = \gamma(x - vt) \\ y &\to y' = y \\ z &\to z' = z . \end{aligned} \tag{1.35}$$

Just as rotations correspond to observers with different orientation in spatial directions, boosts correspond to observers with different orientation in directions involving both space and time, which is to say with different velocities. The fact that the group of Lorentz boosts is non-compact corresponds to the fact that we can never reach the speed of light: transformations with ever larger values of the rapidity ψ bring v ever closer to 1, but the value 1 is never reached.

A boosted coordinate system moves with constant velocity with respect to the original, i.e. in a straight line. Considering motion in the x direction, velocity is just a tilt on the tx-plane, as illustrated in figure 9. Note that the relativity of constant velocity, which is one of the fundamental assumptions in Newtonian mechanics, is a derived property in SR. From the assumption that the metric is the Minkowski metric (and remains invariant under coordinate transformations), it follows that constant velocity is relative, just as it follows that constant angles are relative.

In the limit $v \ll 1$, the t and x transformations (1.35) reduce to $t \to t - vx$, $x \to x - vt$. If we consider the separation between two points along the worldline of a particle moving with a velocity that is $\ll 1$, the transformation reduces to $\Delta t \to \Delta t$, $\Delta x \to \Delta x - v\Delta t$, as $\Delta x \ll \Delta t$. In other words, in the limit of small velocity, Lorentz boosts reduce to the Galilei transformation. Thus the Lorentz transformations unify spatial rotations and Galilei transformations, which in Newtonian mechanics seemed independent. This is an example of the feature that a unified theory is generally not merely the sum of the theories it unifies, but rather a larger whole that has those theories as its limits. In the same vein, we will see that GR is not SR plus Newtonian gravity, but reduces to them in the appropriate limits.

1.4 Dynamics in SR

1.4.1 Generalising Newton's second law

We have introduced the velocity of an observer (note that observers are treated as pointlike) as the proper time derivative of their coordinate $x^\alpha(\tau)$,

$$u^\alpha = \frac{\mathrm{d}x^\alpha}{\mathrm{d}\tau} . \tag{1.36}$$

As the proper time is invariant under coordinate changes, taking a derivative of the components of a vector that depend on the proper time gives the components of

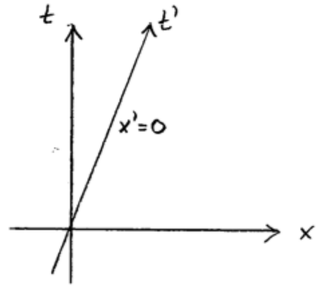


Figure 9: The coordinate system of an observer boosted in the x direction.

a vector. The acceleration is similarly defined as the proper time derivative of the velocity,

$$a^\alpha \equiv \frac{\mathrm{d}u^\alpha}{\mathrm{d}\tau} = \frac{\mathrm{d}^2 x^\alpha}{\mathrm{d}\tau^2} \ . \tag{1.37}$$

Note that by use of the chain rule we have $\frac{\mathrm{d}u^\alpha}{\mathrm{d}\tau} = u^\beta \partial_\beta u^\alpha$, where $\partial_\alpha \equiv \frac{\partial}{\partial x^\alpha}$. From the condition $\underline{u} \cdot \underline{u} = -1$ it follows that $\underline{a} \cdot \underline{u} = 0$: velocity and acceleration are always orthogonal. (**Exercise:** Show this.) This is analogous to a Newtonian particle on a circular orbit: because the norm of the position vector is constant, velocity and position are orthogonal.

Particle momentum is defined as

$$p^\alpha \equiv m u^\alpha = m\gamma(1, \vec{v}) \equiv (E, \vec{p}) \tag{1.38}$$

where m is the mass of the particle, which is a constant, and we have expressed the velocity using (1.31). So energy is the zero component of the four-momentum. In Newtonian mechanics, we have

$$E = \frac{1}{2} m v^2 = \frac{p^2}{2m} \qquad \text{and} \qquad \vec{p} = m\vec{v} \ , \tag{1.39}$$

where $p \equiv \sqrt{\delta_{ij} p^i p^j}$. In SR we have

$$E = m\gamma \qquad \text{and} \qquad \vec{p} = m\gamma\vec{v} \ . \tag{1.40}$$

The Newtonian relations are limiting cases of the SR ones for $v \ll 1$ (with the addition of the rest mass to the energy). To obtain the relation between momentum and energy in SR, let us consider the norm of $\underline{p}$. On the one hand, $\underline{p} \cdot \underline{p} = m^2 \underline{u} \cdot \underline{u} = -m^2$. On the other hand,

$$\underline{p} \cdot \underline{p} = \eta_{\alpha\beta} p^\alpha p^\beta = -E^2 + p^2 \ . \tag{1.41}$$

Combining these expressions, we get

$$E = \sqrt{m^2 + p^2} \simeq m + \frac{p^2}{2m} \simeq m + \frac{1}{2} m v^2 \ , \tag{1.42}$$

where we have expanded in the non-relativistic limit $p/m \ll 1$.

We can now generalise Newton's second law (1.4) to SR. We simply write

$$f^\alpha = ma^\alpha \ , \tag{1.43}$$

where f^α is the four-force. This equation carries little meaning unless we specify f^α. Of course, this is also true for Newton's second law. We can characterise the four-force in terms of the three-force as

$$f^\alpha = \gamma(\vec{F}\cdot\vec{v}, \vec{F}) = \gamma\frac{\mathrm{d}p^\alpha}{\mathrm{d}t} = \gamma\left(\frac{\mathrm{d}E}{\mathrm{d}t}, \frac{\mathrm{d}\vec{p}}{\mathrm{d}t}\right) \ . \tag{1.44}$$

(**Exercise:** Show that this is consistent.) This shows that the zeroth component of the four-momentum has a similar relation to force as in Newtonian physics.

It follows from (1.43) and the generalisation of Newton's third law that the four-momentum of an isolated system is conserved:

$$\sum_{\substack{\text{all}\\ \text{particles}}} p^\alpha = \text{constant} \ , \tag{1.45}$$

which is equivalent to the conservation of energy and three-momentum.

We can also consider SR in arbitrary coordinate systems. However, in SR, as in Newtonian mechanics, the laws of physics only retain the forms given above in coordinate systems that move at constant velocity, i.e. for inertial observers. As an example of a different coordinate system, consider the merry-go-round coordinates, defined in terms of Cartesian coordinates as

$$\begin{aligned} t' &= t \\ x' &= \sqrt{x^2+y^2}\cos(\varphi - \omega t) \\ y' &= \sqrt{x^2+y^2}\sin(\varphi - \omega t) \\ z' &= z \ , \end{aligned} \tag{1.46}$$

where $\varphi = \arctan(y/x)$ and ω is a constant. We will return to the merry-go-round coordinates when we develop the formalism for describing physics in arbitrary coordinate systems in GR.

Exercise: Find the metric tensor in the merry-go-round coordinates, starting from the line element of Minkowski space in Cartesian coordinates (1.10).

1.5 Electrodynamics

1.5.1 Maxwell equations

In section 1.2 we gave the Newtonian gravitational force as an example of a force that is covariant under the Newtonian symmetry transformations (1.2) and (1.3). Let us now give an example of a force that is covariant under the Poincaré transformations, i.e. consistent with SR. The first such force known predates SR: it is the **Lorentz force**, to be considered together with the **Maxwell equations**. In three-vector notation, the Maxwell equations read

$$\begin{aligned} \nabla\cdot\vec{B} &= 0 & \qquad \nabla\cdot\vec{E} &= \rho_q \\ \nabla\times\vec{E} &= -\frac{\partial\vec{B}}{\partial t} & \qquad \nabla\times\vec{B} &= \vec{J} + \frac{\partial\vec{E}}{\partial t} \ , \end{aligned} \tag{1.47}$$

where $\vec{E}$ is the electric field, $\vec{B}$ is the magnetic field, ρ_q is the electric charge density, and $\vec{J}$ is the electric current. The Lorentz force law is

$$\vec{F} = q(\vec{E} + \vec{v} \times \vec{B}) \ , \tag{1.48}$$

where q is the electric charge of the particle and $\vec{v}$ is its coordinate three-velocity.

When Maxwell originally wrote the equations, he did not use vectors. Writing them in terms of three-vectors makes their structure more transparent. Further clarity is achieved by using quantities adapted to Minkowski space. We define the electromagnetic **field strength** in component form in Cartesian coordinates as

$$F_{\alpha\beta} = \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & B_3 & -B_2 \\ E_2 & -B_3 & 0 & B_1 \\ E_3 & B_2 & -B_1 & 0 \end{pmatrix} \Leftrightarrow F^{\alpha\beta} = \begin{pmatrix} 0 & E^1 & E^2 & E^3 \\ -E^1 & 0 & B^3 & -B^2 \\ -E^2 & -B^3 & 0 & B^1 \\ -E^3 & B^2 & -B^1 & 0 \end{pmatrix} \ . \tag{1.49}$$

Note that $F_{\beta\alpha} = -F_{\alpha\beta}$. We have written the spatial indices down in $F_{\alpha\beta}$ and up in $F^{\alpha\beta}$, but there's really no difference, as $E^i = \delta^{ij} E_j$, and similarly for B^i.

The source term of Maxwell equations can be written as the electric current four-vector

$$j^\alpha \equiv (\rho_q, \vec{J}) \ . \tag{1.50}$$

In terms of $F_{\alpha\beta}$ and j^α, Maxwell equations assume simple form:

$$F^{\alpha\beta}{}_{,\beta} = j^\alpha \tag{1.51}$$

$$F_{[\alpha\beta,\gamma]} = 0 \ , \tag{1.52}$$

where $F_{\alpha\beta,\gamma} \equiv \partial_\gamma F_{\alpha\beta}$, and [] refers to antisymmetrisation over all indices; this is defined in (2.33). Using the fact that $F_{\beta\alpha} = -F_{\alpha\beta}$, equation (1.52) can be written as

$$F_{\alpha\beta,\gamma} + F_{\gamma\alpha,\beta} + F_{\beta\gamma,\alpha} = 0 \ . \tag{1.53}$$

The Lorentz force law reads

$$f^\alpha = q F^\alpha{}_\beta v^\beta \ , \tag{1.54}$$

where v^α is the four-velocity of the charged particle (not to be confused with the observer four-velocity u^α).

The field strength can further be written in terms of the electromagnetic **vector potential** A^α:

$$F_{\alpha\beta} = 2\partial_{[\alpha} A_{\beta]} = \partial_\alpha A_\beta - \partial_\beta A_\alpha \ . \tag{1.55}$$

The constraint equation (1.52) is now satisfied identically – (1.55) is the general solution of (1.52). The Maxwell equations can be further simplified by using the freedom to make transformations of A^α. The field strength (and hence the physics) is invariant under the **gauge transformation**

$$A_\alpha \to A'_\alpha = A_\alpha + \partial_\alpha \sigma \ , \tag{1.56}$$

where σ is an arbitrary scalar function. A choice of σ is called **a gauge choice**. With (1.55), the equation of motion (1.51) becomes

$$\partial_\alpha \partial_\beta A^\beta - \eta^{\beta\gamma} \partial_\beta \partial_\gamma A_\alpha \equiv \partial_\alpha \partial_\beta A^\beta - \Box A_\alpha = j_\alpha \ , \tag{1.57}$$

where $\Box \equiv \eta^{\alpha\beta}\partial_\alpha\partial_\beta$. Doing a gauge transformation (1.56) and choosing σ as the solution of the equation $\Box\sigma = -\partial_\alpha A^\alpha$, we get $\partial_\alpha A'^\alpha = 0$.[9] So (dropping the prime) the equation of motion for the components A^α separates, and the Maxwell equations reduce to

$$\Box A^\alpha = -j^\alpha \ . \tag{1.58}$$

We will come back to electromagnetism and this equation when discussing how to generalise the non-gravitational laws of physics (in particular, electromagnetism) when going from SR to GR.

1.5.2 The road in reverse

Let us conclude the section on electrodynamics and the chapter on SR by commenting on the route by which Einstein originally discovered SR via electrodynamics. Einstein started from the assumption that the laws of physics are the same in all coordinate systems that move with constant velocity (inertial observers), as is the case in Newtonian physics. In particular, he demanded that this holds for Maxwell equations. Since the speed of electromagnetic waves follows from (1.47) and is constant (in vacuum), it follows that the speed of light is the same for all observers. Assuming that the coordinate transformations between different inertial coordinates are linear (i.e. that coordinate differentials transform in the same way everywhere in space and time – spacetime is homogeneous), Einstein derived the Lorentz transformation

$$\begin{aligned} t \to t' &= \gamma(t - vx) \\ x \to x' &= \gamma(x - vt) \\ y \to y' &= y \\ z \to z' &= z \ . \end{aligned} \tag{1.59}$$

Taking (1.59) and demanding that Maxwell equations (1.47) retain their form, Einstein found that the components E^i and B^i transform as

$$\begin{aligned} E^x \to E'^x &= E^x & B^x \to B'^x &= B^x \\ E^y \to E'^y &= \gamma(E^y - vB^z) & B^y \to B'^y &= \gamma(B^y + vE^z) \\ E^z \to E'^z &= \gamma(E^z + vB^y) & B^z \to B'^z &= \gamma(B^z - vE^y) \ . \end{aligned} \tag{1.60}$$

Exercise: You can check Einstein's result by showing that, indeed, Maxwell equations retain their form if you apply the transformations (1.59) and (1.60). (Hint: start with the fact that $F_{\alpha\beta}$ defined in (1.49) transforms with an inverse Lorentz transformation matrix for each index.)

Hermann Minkowski, who had taught mathematics and mechanics to Einstein, realised in 1907 that SR was not only a theory of relative space and time, but a theory of absolute spacetime. We have followed the road in the opposite direction. We started from absolute spacetime: from our point of view the Maxwell equations are covariant under Poincaré symmetry because they live in Minkowski space, instead of Minkowski space following from the covariance of Maxwell equations. From this perspective, the association of the speed of light to SR is fortuitous. SR has a

[9] This is called the **Lorenz gauge**. Note that there is no "t". The gauge is named after Ludvig Lorenz, the transformations etc. after Hendrik Lorentz.

maximum velocity for causal influence. Massless fields travel at this maximum velocity. In particular this is true for electromagnetic waves, as shown by the Maxwell equations (1.47). Since electromagnetic radiation is the first massless field known, the maximum signal velocity c was identified with the speed of light. But light as such, or its laws, has no fundamental role in SR. If the laws of electromagnetism were modified by giving the photon a small mass, it would travel at sublight speed, but this would have no impact on SR.

In 1908 Minkowski expressed the radical new vision of spacetime as follows:

> *Henceforth space by itself, and time by itself, are doomed to fade away into mere shadows, and only a kind of union of the two will preserve an independent reality.*

With GR, Einstein, Hilbert, and their coworkers took the next step and discovered that spacetime is not a passive stage where the interactions of matter play out, but a dynamical entity that interacts with matter and itself.