# A RELATIVITY AND GRAVITY

Surprisingly, the concepts behind general relativity as a theory of gravity are entirely geometrical and require differential geometry as a description of the geometry and dynamics of space time. In this script for the lecture on general relativity as a master-level course at Heidelberg University we will encounter gravity as a geometric effect of spacetime and the geometrisation of physical laws, understand the structure of the laws of Nature by drawing analogies between classical and relativistic mechanics of point particles, the theory of scalar and vectorial fields, touch on concepts like invariance, covariance and symmetries, and will develop an intuition about gravity. There are three great applications of general relativity: black holes, FLRW-cosmologies and gravitational waves, and in all these areas there have been major experimental and observational advances in the last couple of years.

In this script we will mostly use a coordinate-based description of tensors with explicit indices. For those, we adopt the summation convention, with Greek indices running over all spacetime coordinates and Latin indices over the spatial ones, should the coordinate choice allow this.

*only coordinate choices aligned with the spatial hypersurfaces have this!*

## A.1 *Why is classical Newtonian gravity insufficient?*

It is important to realise that at the time of A. Einstein's thinking about relativity, there was no actual need to abandon I. Newton's theory of gravity. The perihelion shift of Mercury could have easily have had systematic origins, and many of the arguments against Newton gravity to be the ultimate theory of gravity are purely conceptual.

First of all, there is no dynamics of the gravitational potential $\Phi$ in Newton's theory. According to the Poisson-equation as the field equation of Newton-gravity,

$$\Delta\Phi(x^i, t) = 4\pi G\,\rho(x^i, t), \tag{A.1}$$

the potential is source by the density field $\rho$ in an instantaneous way as the Laplace-operator $\Delta$ can, unlike the d'Alembert-operator $\Box$, not generate any retardation.

The missing retardation could be easily fixed, though. Motivated by the ideas of special relativity that there is no clear distinction between the $t$- and $x^i$-coordinates, one could make the replacement

$$\Delta = \delta^{ij}\,\partial_i\partial_j \;\rightarrow\; \Box = \eta^{\mu\nu}\,\partial_\mu\partial_\nu = \partial_{ct}^2 - \Delta \tag{A.2}$$

which gives a Lorentz-covariant field equation,

$$\Delta\Phi(x^i) = 4\pi G\,\rho(x^i) \;\rightarrow\; \Box\,\Phi(x^\mu) = 4\pi G\,\rho(x^\mu) \tag{A.3}$$

with proper retarded (and advanced) potentials. We can quickly check that there is propagation of excitations of $\Phi$ with $c$ along a light cone: Plane waves $\Phi \sim \exp(\pm i\eta_{\mu\nu}\,k^\mu x^\nu)$ yield $\Box\Phi = -\eta_{\mu\nu}\,k^\mu k^\nu\,\Phi = 0$ with a null-vector $k^\mu$, $\eta_{\mu\nu}\,k^\mu k^\nu = k_\mu k^\mu = 0$. That this "covariantised" field equation already allows wave-like excitations of the gravitational field is foreshadowing the emergence of gravitational waves in proper relativistic theory.

Then, the (energy) density $\rho$ should be the $tt$-component of the energy momentum tensor $T^{\mu\nu}$ as suggested by special relativity and follow a Lorentz-transformation rule when boosting from one Lorentz-frame to another: But the gravitational potential in Newton's theory is scalar and would necessarily be equal in all frames. In fact, there should be additional components of the gravitational field beyond $\Phi$ if it was to be sourced by the energy momentum-tensor $T^{\mu\nu}$.

### A.2 *What would be the most general classical theory of gravity?*

It turns out that Newtonian gravity is not even the most general classical (i.e. non-relativisic, and of course non-quantum) theory of gravity! For seeing this, we would approach the construction of a field equation from a variational principle, by writing down an action integral $S = \int d^3x\, \mathcal{L}(\Phi, \partial_i\Phi)$ with a Lagrange-density $\mathcal{L}(\Phi, \partial_i\Phi)$ that is dependent on the potential $\Phi$ and the first derivative $\partial_i\Phi$. Hamilton's principle $\delta S = 0$ would then suggest that

$$\delta S = \delta \int d^3x\, \mathcal{L}(\Phi, \partial_i\Phi) = \int d^3x\, \frac{\partial\mathcal{L}}{\partial\Phi}\delta\Phi + \frac{\partial\mathcal{L}}{\partial\partial_i\Phi}\delta\partial_i\Phi = \int d^3x \left(\frac{\partial\mathcal{L}}{\partial\Phi} - \partial_i\frac{\partial\mathcal{L}}{\partial\partial_i\Phi}\right)\delta\Phi = 0 \tag{A.4}$$

by using $\delta\,\partial_i\Phi = \partial_i\delta\,\Phi$ and performing an integration by parts, so that we can extract the Euler-Lagrange-equation

$$\frac{\partial\mathcal{L}}{\partial\Phi} - \partial_i\frac{\partial\mathcal{L}}{\partial\partial_i\Phi} = 0 \tag{A.5}$$

which determines the field equation once the Lagrange density $\mathcal{L}$ is chosen. Certainly, Newton would have wanted to have a linear field equation such that the superposition principle is valid, and the Ostrogradsky-theorem disallows terms of higher derivative order beyond second derivatives, and we would like an isotropic gravitational field around a spherically symmetric matter distribution. These arguments imply that there can be at most squares of the potential in the Lagrange-density as the Euler-Lagrange equation decreases the order by one through differentiation, and that there should be the invariant $\delta^{ij}\partial_i\Phi\partial_j\Phi$ (as a scalar product of two vectors it is invariant under rotations) as a kinetic term in the Lagrange-density: Therefore, the most general Lagrange-density would be

$$\mathcal{L}(\Phi, \partial_i\Phi) = \frac{1}{2}\,\delta^{ij}\,\partial_i\Phi\partial_j\Phi + 4\pi G\rho\Phi + \lambda\Phi + \frac{m^2}{2}\Phi^2, \tag{A.6}$$

with the Newtonian gravitational constant $G$ and two new constants, $m$ and $\lambda$. Of course, as Lagrange-densities only ever appear inside integrals, it is only determined up to an integration by parts, so the kinetic term can equally written as

$$\mathcal{L}(\Phi, \partial_i\Phi) = -\frac{1}{2}\,\Phi\,\delta^{ij}\,\partial_i\partial_j\Phi + \cdots, \tag{A.7}$$

with $\Phi\delta^{ij}\partial_i\partial_j\Phi = \Phi\Delta\Phi$. For carrying out the variation, one needs to substitute the Lagrange-density into the Euler-Lagrange equation.

☠ *please always use new names for the indices in the variation!*

We obtain for the kinetic term

$$\mathcal{L} = \frac{1}{2}\,\delta^{ab}\partial_a\Phi\partial_b\Phi \tag{A.8}$$

the derivative

$$\frac{\partial\mathcal{L}}{\partial\partial_i\Phi} = \frac{1}{2}\,\delta^{ab}\left(\frac{\partial\partial_a\Phi}{\partial\partial_i\Phi}\cdot\partial_b\Phi + \partial_a\Phi\frac{\partial\partial_b\Phi}{\partial\partial_i\Phi}\right) = \frac{1}{2}\left(\delta^{ab}\,\delta^i_a\,\partial_b\Phi + \delta^{ab}\,\partial_a\Phi\cdot\delta^i_b\right) = \partial^i\Phi \tag{A.9}$$

and finally

$$\partial_i\frac{\partial\mathcal{L}}{\partial\partial_i\Phi} = \partial_i\partial^i\Phi = \Delta\Phi, \tag{A.10}$$

while the derivative with respect to the field $\Phi$ itself is very easy,

$$\frac{\partial\mathcal{L}}{\partial\Phi} = 4\pi G\rho + \lambda + m^2\,\Phi, \tag{A.11}$$

such that Newton's field equation should be of Yukawa-form, and being inhomogeneous even in vacuum,

$$\left(\Delta - m^2\right)\Phi = 4\pi G\,\rho + \lambda. \tag{A.12}$$

While the value of the gravitational constant $G \simeq 10^{-11}\mathrm{m}^3/\mathrm{kg/s}^2$ is well known, the cosmological constant $\lambda$ is in fact non-zero and plays a role on scales above $10^{25}$m, but is completely irrelevant inside the Solar System. It would, however, have the funny consequence that there would be gravitational effects in empty space! Specifically for $\rho = 0$ the field equation becomes

*Many people claim that the cosmological constant is a part of a relativistic theory of gravity, but this is just untrue.*

$$\Delta\Phi = \lambda \quad \text{such that} \quad \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\Phi}{\partial r}\right) = \lambda \tag{A.13}$$

and the solution for the gravitational acceleration $g_r$ reads

$$\frac{\partial\Phi}{\partial r} = -g_r = \frac{\lambda}{r^2}\int \mathrm{d}r\,r^2 = \frac{\lambda}{3}r, \tag{A.14}$$

increasing linearly with distance: This is in fact observed in cosmology on very large scales in the distance-redshift-relation of supernovae!

The additional constant $m$ is very difficult to interpret classically, but we should see what effects it might have, by solving the resulting field equation. In three dimensions or more, and on scales below $10^{25}$m, the case $m = 0$ reduces the field equation in vacuum to $\Delta\Phi = 0$, i.e. to

$$\Delta\Phi = \frac{1}{r^{n-1}}\frac{\partial}{\partial r}\left(r^{n-1}\frac{\partial}{\partial r}\Phi\right) = 0, \tag{A.15}$$

suggesting that $\left(r^{n-1}\frac{\partial}{\partial r}\Phi\right)$ should be constant for equation to be valid. Specifically in 3 dimensions on obtains a scale-invariant Coulomb-potential, $\Phi \propto 1/r$ whereas a nonzero value for $m$ introduces a scale in the form of a Yukawa-potential

$\Phi \propto \exp(-mr)/r$. For Newton, it must have been an empirical fact that $m$ would be vanishing as there are perfectly Keplerian orbits in the Solar System.

*Debye-screened electrostatic fields in polarisable media is a good example of Yukawa-potentials.*

Let us come to the fact that orbits of planets in the Solar System are almost perfectly Keplerian: A very instructive derivation of Kepler's third law is to use mechanical similarity transforms of the classical action S, in particular for power-law potentials just as the Coulomb-potential.

$$S = \int \mathrm{d}t\, L(x_i, \dot{x}_i) \quad \text{with} \quad L(x_i, \dot{x}_i) = \frac{m}{2}\,\delta^{ij}\,\dot{x}_i\dot{x}_j - m\,\Phi \tag{A.16}$$

If we introduce a scaling of distance and time with the transformations $x \to \alpha x$ and $t \to \beta t$, the kinetic term transforms according to $T \to \frac{\alpha^2}{\beta^2} T$ and the potential term with $\Phi \to \alpha^n\,\Phi$ if the potential is in fact a power law of with exponent $n$, $\Phi \sim x^n$.

If the two scaling factors are related through $\frac{\alpha^2}{\beta^2} \sim \alpha^n$, L changes only by an overall factor, which can not matter because the Hamiltonian principle is invariant under affine transformations of the action (or equivalently, of the Lagrange function):

$$L \to aL + b \quad \text{implies} \quad S \to aS + b \quad \text{with} \quad S = \int \mathrm{d}t\, L \tag{A.17}$$

such that

$$\delta S = 0 \;\to\; \delta\,(aS + b) = a\,\delta S = 0, \tag{A.18}$$

and $a$ cancels. Therefore, the two scalings in length and time can not be independent and their relation must depend on the exponent of the power law of the potential: This is summarised by the similarity condition $t^2 \sim x^{2-n}$, which is sometimes referred to as classical similarity: Motion inside a potential with a given exponent is described by an equivalence class of Lagrange-functions (and their solutions), which get mapped onto each other by a similarity transform. The most basic choices of $n$ correspond to well known problems in classical mechanics:

$$n = 2 \qquad t^2 \sim x^0 \quad \text{isochronism of a pendulum} \tag{A.19}$$

$$n = 1 \qquad t^2 \sim x \quad \text{inclined plane, constant acceleration} \tag{A.20}$$

$$n = 0 \qquad t^2 \sim x^2 \quad \text{inertial motion with constant velocity} \tag{A.21}$$

$$n = -1 \qquad t^2 \sim x^3 \quad \text{Kepler's third law} \tag{A.22}$$

Supposedly, the first case was discovered by G. Galilei himself, who noticed that the oscillation period of a pendulum (measured with the pulse on his wrist) did not depend on the amplitude. The last case, Kepler's third law of planetary motion, is necessarily a consequence of the $1/r$-form of the potential and any Yukawa-type contribution would break the scaling relation.

Besides, Kepler's third law provides a neat trick to remember the units of the gravitational constant, $G \sim 10^{-11}\mathrm{m}^3/\mathrm{kg}/\mathrm{s}^2$, where one can immediately recognise the three powers of length divided by the two powers of time! For our Sun, $GM_\odot \sim 10^{19}\,\mathrm{m}^3/\mathrm{s}^2 \sim (1\,\mathrm{AU})^3/(1\,\mathrm{yr})^2$. But mechanical similarity applied to the Solar System is a really powerful concept: All planetary orbits are scaled versions of each other, and for measuring distances one really only needs a calendar!

We have obtained $\Phi \propto 1/r$ from direct solution of Poisson's equation in the case $\rho = 0$, but there needs to be a fundamental argument why this is necessary, and this argument comes in the shape of Gauß's law. The gravitational acceleration $g_i$ is given as the gradient $g_i = -\partial_i\Phi$, and has in a spherically symmetric case only a radial component, $g_r = -\partial_r\Phi$. It is linked to the Poisson equation by

$$\Delta\,\Phi = \delta^{ij}\,\partial_i\partial_j\Phi = -\delta^{ij}\,\partial_i g_j = -\mathrm{div}\boldsymbol{g} = 4\pi G\,\rho \tag{A.23}$$

suggesting that the divergence of the acceleration is $\Delta\Phi$ and proportional to $\rho$. Recasting the Poisson-equation into integral form yields

$$\int_V \mathrm{d}^3 r\,\mathrm{div}\boldsymbol{g} = \int_{\partial V} \mathrm{d}\mathbf{A}\cdot\boldsymbol{g} = -4\pi G\cdot\int_V \mathrm{d}^3 r\rho = -4\pi G M. \tag{A.24}$$

Here, $\int \mathrm{d}\mathbf{A}\cdot\boldsymbol{g}$ is the flux of the field through the surface $\partial V = 4\pi r^2$ for a spherical integration volume V appropriate for the isotropic case. As a consequence, the acceleration decreases $\propto 1/r^2$ as the flux needs to be the same at every distance and surfaces increase $\propto r^2$! Now we can set up an entire chain of arguments: The flux of $\vec{g}$ through surfaces $\partial V$ is constant, so $g$ needs to be $\propto 1/r^2$ and $\Phi \propto 1/r$. Then, mechanical similarity requires that $t^2 \propto r^3$. And in addition, Bertrand's theorem makes sure that the orbits are closed ellipses.

At this point it is a very large surprise that Mercury, the planet closest to the Sun where perhaps the gravitational field behaves unusual, shows a tiny violation of Kepler's third law and in fact of Bertrand's theorem, too: Neither is the orbit a closed ellipse nor is Kepler's law fulfilled. There is a small precession of the point of closest approach to the Sun, called perihelion precession, amounting to 43 arcseconds in about 1000 orbits (The number is usually stated as 43 arcseconds in 100 years, but this refers to Earth years!). By now, we know many systems that show pericentre precession, even much more pronounced than Mercury in the Solar System. For instance, PSR 1913+10 with 4 arcseconds per orbit, PSR J0737-3039 with 20 arcseconds per orbit, and the system OJ287 with 40° per orbit! For a precession to appear, the gravitational field needs to be stronger in the vicinity of a massive object compared to the Newtonian prediction, and neither $m$ nor $\lambda$ could achieve this: They both correspond to long-distance modifications of gravity: That would be a very strong argument for the necessity of a new theory of gravity. And we can see a tiny glimpse onto geometry. Combining the constant of gravity G with the speed of light $c$,

$$\frac{G}{c^2} \sim 10^{-28}\mathrm{m/kg} \tag{A.25}$$

which assigns a length scale to the field generating mass. With the specific value $M_\odot \simeq 10^{30}$kg for the mass of the Sun on obtains

$$\frac{GM_\odot}{c^2} \sim 10^2\mathrm{m} \tag{A.26}$$

which we will encounter later as the Schwarzschild radius of the Sun. Perhaps we can change how surfaces scale with $r$?

## A.3 *Lorentz-geometry*

The foundational idea of general relativity is differential geometry, i.e. a varying geometry of spacetime, with locally Minkowskian properties, i.e. we will see that the laws of special relativity will be valid locally in freely falling reference frames. Lorentz-transforms and rotations apply locally to the transitions between frames with different orientation relative to each other or moving at constant velocities $\upsilon$ relative to each other. The homogeneity of spacetime should be respected by the coordinate choice, meaning that it should not single out certain spacetime points.

An observer looking at two coordinate choices could measure the rate at which the coordinates $x^\mu$ and $x'^\mu$ are drifting by as a function of her or his proper time $\tau$, defining the velocity

$$\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} = \text{const.}\ \text{, with } x^\mu = \begin{pmatrix} t \\ x^i \end{pmatrix} \sim \text{ 4-vector} \tag{A.27}$$

which is constant for inertial motion and suitably chosen coordinates, and the corresponding acceleration

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2} = 0, \text{ and identically in S}' : \frac{\mathrm{d}^2 x'^\mu}{\mathrm{d}\tau^2} = 0 \tag{A.28}$$

which then vanishes. Then, the relation between the two velocities and accelerations is given by

$$\frac{\mathrm{d}x'^\mu}{\mathrm{d}\tau} = \frac{\partial x'^\mu}{\partial x^\nu}\frac{\mathrm{d}x^\nu}{\tau}, \text{ with Jacobian } \frac{\partial x'^\mu}{\partial x^\nu} \tag{A.29}$$

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\tau^2} = \frac{\partial x'^\mu}{\partial x^\nu}\frac{\mathrm{d}^2 x^\nu}{\mathrm{d}\tau^2} + \frac{\partial^2 x'^\mu}{\partial x^\nu \partial x^\rho}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau}\frac{\mathrm{d}x^\rho}{\mathrm{d}\tau}, \tag{A.30}$$

where

$$\frac{\partial^2 x'^\mu}{\partial x^\nu \partial x^\rho} = 0 \tag{A.31}$$

for transformations between frames that are linear and therefore conserve homogeneity. The solution for $x'^\mu(x^\nu)$ follows then as

$$x'^\mu = \mathrm{A}^\mu{}_\nu\, x^\nu + a^\mu, \tag{A.32}$$

implying that the transformation between frames should be affine.

Let's construct this transform from the most general transition between two frames, where we align for simplicity the coordinate axes with the direction of relative motion, taken to be the $x$-axis. There is an event with coordinates $\begin{pmatrix} t \\ x^i \end{pmatrix}$ in S and $\begin{pmatrix} t' \\ x'^i \end{pmatrix}$ in S', and the two frames move with a relative (constant) velocity $\upsilon$.

A linear transform would then the only one to respect the homogeneity of spacetime (nonlinear transforms would always single out certain spacetime points), so we make the ansatz:

$$x' = ax + bt, \;\; a, b \text{ arbitrary, but } x = vt \text{ must imply } x' = 0 \tag{A.33}$$

$$x' = 0 = avt + bt = (av + b)\, t \;\Rightarrow\; b = -av, \text{ and:} \tag{A.34}$$

$$x' = a(x - vt)\ (*) \tag{A.35}$$

Reversing the roles of S and S′ implies that

$$x = ax' + bt' \text{ but } x' = -vt \text{ must imply } x = 0 \tag{A.36}$$

$$x = 0 = -avt' + bt' = (-av + b)\, t' \;\Rightarrow\; b = +av, \text{ and:} \tag{A.37}$$

$$x = a(x' + vt')\ (**) \tag{A.38}$$

But this relation between $x$ and $x'$ is not yet fixed without an additional assumption that determines the value of $a$. Here, Nature would have in fact a choice! Either, Nature could work with a universal time coordinate (or rather, a parameter, as it does not participate in transforms unlike the other coordinates). A universal time parameter would require that $t = t'$, which is the defining property of Galilei-transforms. Then,

$$x' = a(x - vt) \tag{A.39}$$

$$x = a(x' + vt) = a(a(x - vt) + vt) = a^2 x + (1 - a)\, vt = x \tag{A.40}$$

which can only be realised if $a = 1$. Nature chose instead, for very good reasons, the speed of light to be equal in all frames, $c = c'$, which requires Lorentz- instead of Galilei-transforms between frames. In this choice,

$$x' = ct' = a(ct - vt) \tag{A.41}$$

$$x = ct = a(ct' - vt') \tag{A.42}$$

$$\Rightarrow\; c^2 tt' = a^2 (c - v)(c + v) \cdot tt', \tag{A.43}$$

where the third equation was obtained by multiplying the first two. Dividing by $tt'$ and solving for $a$ yields the Lorentz-factor $\gamma$,

$$a = \gamma = \frac{1}{\sqrt{1 - \beta^2}}, \quad \text{with} \quad \beta = \frac{v}{c} \tag{A.44}$$

We should note that Lorentz-transformations, due to their linearity, do not 'mix' the spatial coordinates. The Lorentz-factor $\gamma$ diverges at $\beta = 1$ and would indeed become imaginary for values $\beta > 1$. Taylor-expanding $\gamma$ for small velocities $\beta$ gives the result that

$$\gamma \sim 1 + \left.\frac{\partial^2 \gamma}{\partial \beta^2}\right|_{\beta=0} \cdot \frac{\beta^2}{2} = 1 + \frac{\beta^2}{2}, \text{ with } \left.\frac{\partial \gamma}{\partial \beta}\right|_{\beta=0} = 1 \tag{A.45}$$

which is perfectly consistent with the fact that for low velocities $\beta \ll 1$ and $\gamma \simeq 1$, Lorentz- and Galilei-transforms are indistinguishable.

Writing $ct$ and arranging the temporal and spatial coordinates into a vector $x^\mu = \begin{pmatrix} ct \\ x \end{pmatrix}$ allows to use the standard matrix-form of the Lorentz-transformation by eliminating $x'$ from (*) and (**):

$$x' = \gamma\,(x - vt) = \gamma\,(x - \beta ct) \tag{A.46}$$

$$ct' = \gamma\,(ct - \beta x), \tag{A.47}$$

so that one arrives at

$$\begin{pmatrix} ct' \\ x' \end{pmatrix} = \begin{pmatrix} \gamma & -\beta\gamma \\ -\beta\gamma & \gamma \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix} \tag{A.48}$$

encapsulating the Lorentz-transform in a matrix $\Lambda^\mu{}_\nu$ with

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu. \tag{A.49}$$

We have seen that coordinates undergo a joint transformation and that any physical statement on coordinates of an event is only sensibly within a specified frame S. From that one might wonder if there is a way to make true statements about physical properties of a system independent from a specification of a frame: That is exactly the idea of a Lorentz-invariant. Similarly to rotations, where $r^2 = \delta_{ij} x^i x^j$ are invariant, which essentially corresponds to the statement $\cos^2\alpha + \sin^2\alpha = 1$ if the rotation is parameterised by an angle $\alpha$, one can define invariants for Lorentz transforms and relate them to the rapidity $\psi$ which is indicative of the relative velocity between the frames.

Setting $\cosh\psi = \gamma$ and $\sinh\psi = \beta\gamma$ (which is sensible if you look at the range of values of $\gamma$ and $\beta\gamma$, and compare with the hyperbolic functions), one obtains the relation $\tanh\psi = \frac{\beta\gamma}{\gamma} = \beta$ between the rapidity and the dimensionless velocity $\beta = \upsilon/c$. Lorentz-transformations can then be written compactly as

$$\begin{pmatrix} ct' \\ x' \end{pmatrix} = \begin{pmatrix} \cosh\psi & \sinh\psi \\ \sinh\psi & \cosh\psi \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix}, \tag{A.50}$$

as a hyberbolic rotation, suggesting an invariant through $\cosh^2(\psi) - \sinh^2(\psi) = \gamma^2 - \beta^2\gamma^2 = \gamma^2(1-\beta^2) = 1$, which we have already derived by direct calculation, $(ct')^2 - x'^2 = (ct)^2 - x^2$.

Analogous to rotations we write the Lorentz-invariant as $s^2 = (ct)^2 - x^2$ by introducing the Minkowski-metric,

$$\eta_{\mu\nu} = \begin{pmatrix} +1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{A.51}$$

such that one can write $s^2 = \eta_{\mu\nu}\, x^\mu x^\nu$. In contrast to the invariant $r^2$ in Euclidean space, Lorentz-invariants can be positive, negative or zero, and as the sign of the Lorentz-invariant is of course conserved under transforms, too, the classification into timelike ($s^2 > 0$), spacelike ($s^2 < 0$) and lightlike (or null, $s^2 = 0$) is very suggestive.

Let's now imagine the motion of a point through spacetime: The Lorentz-invariant reads

$$s^2 = (ct)^2 - x^2 = (ct')^2 - x'^2 = (c\tau)^2 \tag{A.52}$$

in two frames S and S′, and the choice of comoving coordinates $x' = 0$ defines proper time $t' = \tau$, which is read off a clock in the rest frame S′. Rewriting the Lorentz-invariant for infinitesimal coordinate differences,

$$\mathrm{d}s^2 = (c\mathrm{d}t)^2 - \mathrm{d}x^2 = (c\mathrm{d}\tau)^2, \tag{A.53}$$

then shows that the passage of coordinate time $\mathrm{d}t$ and proper time $\mathrm{d}\tau$ differ by an inverse Lorentz-factor,

$$\mathrm{d}\tau = \sqrt{1-\beta^2}\mathrm{d}t \quad \text{with} \quad \beta = \frac{1}{c}\frac{\mathrm{d}x}{\mathrm{d}t} \tag{A.54}$$

That implies that the length $\mathrm{d}s$ of the spacetime curve that a point takes is actually measure by the comoving clock displaying proper time $\mathrm{d}\tau$, at least for timelike motion with velocities $\beta < 1$!

At this point, by merging the temporal coordinate with the spatial coordinates we obtained $\mathbb{R}^4$ with a particular geometric structure, given by the Minkowski scalar product $\langle x, y\rangle = \eta_{\mu\nu}x^\mu y^\nu$, trading the positive definiteness of the Euclidean scalar product for the ability to define general invariants.

### A.3.1 Lie-groups and the generation of the Lorentz-group

Rotations and Lorentz-boosts are the fundamental transforms that leave a Lorentzian spacetime invariant. Both transformations are (non-Abelian) groups and are parameterised by real numbers, the rotation angles in the first and the rapidities in the second case. One might ask now the question whether there is something analogous to a basis of these groups, such that all group elements can be addressed by a suitable choice of the rotation angle or the rapidity: It turns out that this presumption is true, and it brings us to the topic of Lie-groups. Lie-groups are continuously parameterised groups and are generated from a basic building block, called, well, a generator.

If we choose the set of Pauli-matrices,

*There are many different definitions of Pauli-matrices, we're using the real-valued ones here, which in a real-valued linear combination, are a basis of the space of 2 × 2-matrices.*

$$\sigma^{(0)} = \begin{pmatrix} +1 & 0 \\ 0 & +1 \end{pmatrix}, \quad \sigma^{(1)} = \begin{pmatrix} +1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \sigma^{(2)} = \begin{pmatrix} 0 & +1 \\ -1 & 0 \end{pmatrix}, \quad \sigma^{(3)} = \begin{pmatrix} 0 & +1 \\ +1 & 0 \end{pmatrix}, \tag{A.55}$$

to begin with, we can investigate which type of transformation could be generated by substituting them into and exponential series, for instance

$$\Lambda = \exp\left(\Psi \cdot \sigma^{(3)}\right) = \sum_n \frac{1}{n!}\left(\psi \cdot \sigma^{(3)}\right)^n. \tag{A.56}$$

For evaluating the matrix-valued exponential series, one needs to know all powers of the matrix in question. In the case of the Pauli-matrices, it is easy to show that only ever other Pauli-matrices appear. Specifically for $\sigma^{(3)}$ one gets:

$$\left(\sigma^{(3)}\right)^0 = \sigma^{(0)}, \left(\sigma^{(3)}\right)^1 = \sigma^{(3)}, \left(\sigma^{(3)}\right)^2 = \sigma^{(0)}, \left(\sigma^{(3)}\right)^3 = \sigma^{(3)}. \tag{A.57}$$

Then, the exponential series can be summed,

$$\Lambda = \sigma^{(0)} \cdot \sum_n \frac{\psi^{2n}}{(2n)!} + \sigma^{(3)} \sum_n \frac{\psi^{2n+1}}{(2n+1)!} = \sigma^{(0)} \cdot \cosh\psi + \sigma^{(3)} \cdot \sinh\psi = \begin{pmatrix} \cosh\psi & \sinh\psi \\ \sinh\psi & \cosh\psi \end{pmatrix}, \tag{A.58}$$

and one recovers the expression for the Lorentz-transform as a hyperbolic rotation. T he invariant $\det\Lambda = \cosh^2\psi - \sinh^2\psi = 1$, which otherwise appears as a property of the hyperbolic function, comes out naturally like this: Using $\ln\det\Lambda = \mathrm{tr}\ln\Lambda$ with $\Lambda = \exp(\psi\cdot\sigma^{(3)})$ implies that $\ln\det\Lambda = \psi\cdot\mathrm{tr}\sigma^{(3)} = 0$, because Pauli-matrices (with the exception of $\sigma^{(0)}$) are traceless. Then, the determinant needs be equal to one.

Surely, boosts and rotations are groups, but how does one need to combine their continuous parameters? This question is readily answered by the tools that Lie-groups provide: For instance, two successive boosts

$$\Lambda(\phi)\cdot\Lambda(\psi) = \exp\left(\phi\cdot\sigma^{(3)}\right)\cdot\exp\left(\psi\cdot\sigma^{(3)}\right) = \exp\left(\left(\phi+\psi\right)\sigma^{(3)}\right) = \Lambda\left(\phi+\psi\right), \tag{A.59}$$

implying that rapidities (and not the velocities!) are in fact additive parameters for boosts. If one wants to revert to velocities, one can use the addition theorem for the hyperbolic tangent:

$$\tanh(\phi) + \tanh\psi = \tanh(\phi+\psi)\cdot\left[1 + \tanh(\phi)\cdot\tanh\psi\right] \tag{A.60}$$

We have just shown that rapidities add for boosts, and from the commutativity of the addition of real numbers one should then obtain the commutativity of the boosts into the same direction:

$$\Lambda(\phi)\cdot\Lambda(\psi) = \Lambda\left(\phi+\psi\right) = \Lambda\left(\psi+\phi\right) = \Lambda(\psi)\cdot\Lambda(\phi), \tag{A.61}$$

which implies for the inverse boost that

$$\Lambda(\psi)\cdot\lambda(-\psi) = \Lambda(\psi-\psi) = \Lambda(0) = \mathrm{id} \;\Rightarrow\; \Lambda(\psi)^{-1} = \Lambda(-\psi) \tag{A.62}$$

as a perfectly intuitive result: The inverse boost is that with the inverse velocity or rapidity. In complete analogy we would have obtained rotations by starting the constructing with $\sigma^{(2)}$ instead of $\sigma^{(3)}$.

*Lorentz-transforms are orthogonal, but with respect to $\eta$, not $\delta$.*

Finally, one could ask the question what happens if rotations around different axes and boosts into different directions are combined. If both transformations are generated by basis elements A and B in an exponential series, their successive application $\exp(A)\exp(B)$ is only equal to $\exp(A+B)$ if the generators commute, $[A,B] = AB - BA = 0$, which is not the case in every of our examples. The Rubik's cube demonstrates nicely that rotations in 3 dimensions do not commute, and neither do boosts: In fact, if one moves from one inertial frame into another by a combination of two boosts and moving back by interchanging the two boosts leaves you with a rotation! In the case of non-commuting generators, the two transformation need to be combined using the Baker-Hausdorff-Campbell-formula,

*my most favourite formula of all of physics!*

$$\exp(A)\exp(B) = \exp(A+B)\cdot\exp\left(-\frac{1}{2}[A,B]\right), \tag{A.63}$$

to lowest order, or exactly if $[A, [A, B]] = 0$ and $[B, [B, A]] = 0$ is valid. In fact, one can define a set of generators in 4d for the group comprising rotations around all three axes and boosts into all 3 directions with a very rich algebra of generators, called the Lorentz-algebra.

## A.4 *Relativistic motion through spacetime*

It might come as a surprise that variational principles, being so typical of classical mechanics, only make sense in the context of relativity: Here, there is a well define geometric interpretation, the Lagrange-function and the action are measurable quantities, many properties such as their convexity and their affine invariance are made sure by geometry, they are naturally invariant under Lorentz-transforms and there is a natural pathway to include gravity.

### A.4.1 Variational principles for relativistic mechanics

The fundamental idea of variational principles (and which ironically is not present clearly in classical mechanics) is to link invariant quantities of a system in the form of the Lagrange-function with a covariant equation of motion. Specifically, the Lagrange function $L(x^i, \dot{x}^i)$ of classical mechanics

$$L(x^i, \dot{x}^i) = \frac{m}{2}\delta_{ij}\dot{x}^i\dot{x}^j - m\,\Phi \tag{A.64}$$

is invariant as the norm of $\dot{x}^i$ is unaffected by rotations of the coordinate systems and because the scalar potential $\Phi$ does not have any internal degrees of freedom. With Hamilton's principle $\delta S = 0$ for the variation of the action

$$S = \int \mathrm{d}t\; L(x^i, \dot{x}^i) \tag{A.65}$$

one obtains through the Euler-Lagrange equation a covariant equation of motion

$$\ddot{x}^i = -\partial^i\Phi, \tag{A.66}$$

which sets two vectors in relation to each other, namely the acceleration $\ddot{x}^i$ and the potential gradient $\partial^i\Phi$, which of course have the same transformation properties. While this is a perfectly valid example of covariance generated from an invariant Lagrange-function, one should note that while it is invariant unter rotations, it is not invariant under Galilei-transforms.

Let's take a leap of faith and replace the Lagrange-function by something relativistic, for instance the proper time $\tau = \int \mathrm{d}\tau$, which is measurable with a clock, fully Lorentz-invariant as

$$\mathrm{d}s^2 = c^2\mathrm{d}\tau^2 = c^2\mathrm{d}t^2 - \gamma_{ij}\mathrm{d}x^i\mathrm{d}^j = \eta_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu, \tag{A.67}$$

and has an intuitive geometric interpretation as the arc-length of the trajectory through spacetime in the geometry defined by the Minkowski-metric $\eta_{\mu\nu}$.

Suppose that a particle travels through spacetime along a trajectory $x^\mu(\tau)$. Then, we can define the 4-velocity $u^\mu$ as the rate at which the coordinates pass by the observer,

$$u^\mu = \frac{\mathrm{d}}{\mathrm{d}\tau}x^\mu(\tau) = \frac{\mathrm{d}}{\mathrm{d}\tau}x^\mu(\tau) = \frac{\mathrm{d}t}{\mathrm{d}\tau}\frac{\mathrm{d}}{\mathrm{d}t}x^\mu = \gamma \cdot \begin{pmatrix} c \\ \upsilon^i \end{pmatrix} \tag{A.68}$$

with Lorentz-factor $\frac{\mathrm{d}t}{\mathrm{d}\tau} = \gamma$. The normalisation of the 4-velocity can be computed straightforwardly,

$$\eta_{\mu\nu}u^\mu u^\nu = u_\mu u^\mu = \gamma^2 \cdot \left(c^2 - \upsilon_i \upsilon^i\right) = c^2, \tag{A.69}$$

because $\gamma^2\left(1 - \beta^2\right) = 1$ for $\beta^i = \frac{\upsilon^i}{c}$. The 4-velocity, or the tangent to the trajectory $x^\mu(\tau)$ is therefore timelike $\eta_{\mu\nu}\, u^\mu u^\nu = c^2 > 0$ and the particle moves inside the light cone.

Inertial motion of a free particle should proceed along a straight line as a natural result of the relativistic variational principle. Indeed, starting with the arc-length $s$

$$S = \int_A^B \mathrm{d}s = \int_A^B c\mathrm{d}\tau = c \cdot \int_A^B \frac{\mathrm{d}t}{\gamma} \tag{A.70}$$

of a trajectory linking the spacetime points A to B we obtain the elapsed proper time $\tau$ (which can be measured by a clock carried along by the particle) or the integrated laboratory time $\int \mathrm{d}t/\gamma$, weighted by the Lorentz-factor, which is responsible for relativistic time dilation: $\mathrm{d}\tau = \mathrm{d}t \cdot \sqrt{1 - \delta_{ij}\,\beta^i\beta^j} = \frac{1}{\gamma} \cdot \mathrm{d}t$, and because $\gamma \geq 1$, $\mathrm{d}\tau$ is always smaller than $\mathrm{d}t$ and proper time elapses slower.

This would imply that the Lagrange function of a free particle is $L(\dot{x}^i) = 1/\gamma$, and that the action S is in fact the arc-length of a trajectory. In the slow-motion limit $|\beta| \ll 1$ one should recover the classical Lagrange-function: Taylor-expanding yields

$$S \simeq -mc^2 \int_A^B \mathrm{d}t \cdot \left(1 - \frac{\delta_{ij}}{2}\beta^i\beta^j\right) = +mc^2 \int_A^B \mathrm{d}t \cdot \delta_{ij}\beta^i\beta^j + \text{const.} = m \cdot \int_A^B \mathrm{d}t \cdot \delta_{ij}\upsilon^i\upsilon^j \tag{A.71}$$

with irrelevant prefactors, as affine transformations $S \to aS + b$ with two constants $a$, $b$ drop out in the Euler-Lagrange-equation. Effectively, the non-relativistic limit yields the kinetic energy as the leading-order term of the proper time integral.

Funnily, Lorentz-covariance is lost in the non-relativistic limit and Galilei-invariance is not generated: If one carries out a Galilei-transform by setting $x^i \to x^i + \upsilon^i t$, and $\dot{x}^i \to \dot{x}^i + \upsilon^i$ one obtains:

$$L = \frac{1}{2}\delta_{ij}\dot{x}^i\dot{x}^j \to \frac{1}{2}\delta_{ij}\dot{x}^i\dot{x}^j + \delta_{ij}\dot{x}^i\upsilon^j + \frac{1}{2}\delta_{ij}\upsilon^i\upsilon^j = L + \frac{\mathrm{d}}{\mathrm{d}t}\left(\delta_{ij}x^i\upsilon^j + \delta_{ij}\upsilon^i\upsilon^j \cdot t\right), \tag{A.72}$$

where the additional term is a total time derivative with no influence on the variational principle: We find ourselves in the weird situation that we need a new concept to remedy the error made by classical Galilei-invariance!

Is inertial motion really proceeding along a straight line? Hamilton's principle requires that $\delta S = 0$, so

$$\delta S = -mc^2\,\delta\int_A^B \mathrm{d}\tau = -mc^2\int_A^B \frac{\eta_{\mu\nu}}{2\mathrm{d}\tau}\cdot\left[\mathrm{d}x^\mu\cdot\delta\mathrm{d}x^\nu + \delta\mathrm{d}x^\mu\cdot\mathrm{d}x^\nu\right] = -mc^2\int_A^B \eta_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\delta\mathrm{d}x^\nu, \tag{A.73}$$

where we have used the symmetry of the integrand to get rid of the factor 1/2. For continuing, we interchange variation and differentiation, $\delta\mathrm{d}x^\nu = \mathrm{d}\delta x^\nu$ and perform an integration by parts

$$\delta S = +mc^2\int_A^B \mathrm{d}\tau\,\eta_{\mu\nu}\,\frac{\mathrm{d}^2x^\mu}{\mathrm{d}\tau^2}\delta x^\nu, \tag{A.74}$$

where $\mathrm{d}(\mathrm{d}x^\mu/\mathrm{d}\tau) = (\mathrm{d}^2x^\mu/\mathrm{d}\tau^2)\,\mathrm{d}\tau$, and with the assumption of vanishing variation on the boundary. The result then is that the 4-acceleration needs to vanish,

$$\frac{\mathrm{d}^2x^\mu}{\mathrm{d}\tau^2} = 0 \tag{A.75}$$

for fulfilling Hamilton's principle, and the equation of motion is solved to yield $x^\mu(\tau) = a^\mu\tau + b^\mu$ with two integration constants: In fact, the solution is a straight line through spacetime.

### A.4.2 Legendre-transforms and Hamilton-functions

We have seen in the last chapter that the Lagrange-function is much more a statement of causal motion in spacetime and has little to do with energies: Those appear after Legendre-transform, which is always well defined because the Lagrange function is a convex functional in $\dot{x}$ - this is, incidentally, the same reason why the variation yields a unique result and finds a unique extremum. In fact, the relativistic Lagrange-function $L = 1/\gamma$ is perfectly convex as it always lies above its tangent: To visualise this, one can write $1/\gamma = \sqrt{c^2 - \upsilon^2}$, whose graph is a semi-circle!

Not only do convex functions have uniquely defined Legendre-transforms, but the Legendre-transformed function is again convex, making sure that the inverse transform is possible, too. Starting with the relativistic Lagrange-function

$$L(\dot{x}) = \frac{1}{\gamma} = \sqrt{c^2 - \dot{x}_i\dot{x}^i} \tag{A.76}$$

we can define the canonical momentum

$$p^i = \frac{\partial\mathcal{L}}{\partial\dot{x}^i} = \frac{\dot{x}^i}{\sqrt{c^2 - \dot{x}_j\dot{x}^j}} \tag{A.77}$$

which we need to convert into a relation for $\upsilon(p)$: Let's do this in one dimension for simplicity.

$$p^2\left[c^2 - \upsilon^2\right] = \upsilon^2,\ p^2c^2 = \upsilon^2\left(1 + p^2\right) \rightarrow \upsilon = \frac{cp}{\sqrt{1 + p^2}} \tag{A.78}$$

Then, the Legendre-transform, replacing $\dot{x} = \upsilon$ by $p$ can be carried out and the Hamilton-function $\mathcal{H}$ can be obtained:

$$\mathcal{H}(p) = \dot{x}\frac{\partial\mathcal{L}}{\partial\dot{x}} - \mathcal{L}(\dot{x}(p)) = \upsilon \cdot \frac{\upsilon}{\sqrt{c^2 - \upsilon^2}} + \sqrt{c^2 - \upsilon^2} = \upsilon p + \frac{\upsilon}{p} = c \cdot \sqrt{1 + p^2}, \quad (A.79)$$

and if we include the prefactor $mc^2$:

$$\mathcal{H}(p) = \sqrt{(mc^2)^2 + c^2p^2} \simeq mc^2 + \frac{p^2}{2m} + \cdots, \quad (A.80)$$

where $mc^2$ is the rest mass and $\frac{p^2}{2m}$ is the kinetic energy which appear in a Taylor-expansion in the last step.

For massive particles the energy-momentum-relation $\mathcal{H}$ allows statements about dispersion: In fact, phase and group velocities can not be equal for massive particles, $\mathcal{H}/p \neq \mathrm{d}\mathcal{H}/\mathrm{d}p$, but one can show that

$$\upsilon_{\mathrm{ph}} \cdot \upsilon_{\mathrm{gr}} = \frac{\mathcal{H}}{p}\frac{\mathrm{d}\mathcal{H}}{\mathrm{d}p} = \frac{cp}{\sqrt{1+p^2}}\frac{c\sqrt{1+p^2}}{p} = c^2 \quad (A.81)$$

i.e. that the geometric mean of phase and group velocity is the speed of light. That in turn implies that the phase velocity of massive particles needs to be $\upsilon_{\mathrm{ph}} > c$ if their group velocity is subluminal, $\upsilon_{\mathrm{gr}} < c$. And, as a shortcut,

$$c^2 = \frac{\mathcal{H}}{p}\frac{\mathrm{d}\mathcal{H}}{\mathrm{d}p} = \frac{\mathrm{d}(\mathcal{H}^2)}{\mathrm{d}(p^2)} \quad (A.82)$$

which can be integrated to give $\mathcal{H}^2 = (cp)^2 + \mathrm{const}$, with the rest mass as the integration constant.

We have already encountered the classification of Lorentz-vectors in timelike, spacelike and lightlike, and we saw that 4-velocities $u^\mu$ are normalised according to $\eta_{\mu\nu}u^\mu u^\nu = c^2 > 0$ with the associated motion inside the light cone. Clearly, that normalisation is conserved under Lorentz-transforms, but one might be curious as to the possibility whether forces could accelerate a particle to super-luminous speeds: A classical argument would be that this would be energetically impossible due to relativistic mass increase (which is really only a consequence of proper time dilation), but there is a more elegant, geometric argument. Acting on a charged, massive particle with a Lorentz-force leads to the equation of motion

$$\frac{\mathrm{d}u^\mu}{\mathrm{d}\tau} = \frac{q}{m}F^{\mu\nu}u_\nu. \quad (A.83)$$

Multiplying both sides with $u_\mu$ then gives a relation how the normalisation of $u^\mu$ would change under the influence of a Lorentz-force:

$$u_\mu\frac{\mathrm{d}u^\mu}{\mathrm{d}\tau} = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}\tau}\left(u_\mu u^\mu\right) = \frac{q}{m}F^{\mu\nu}u_\mu u_\nu = 0, \quad (A.84)$$

where the last term is vanishing as a contraction between an antisymmetric tensor $F^{\mu\nu}$ and a symmetric one, $u_\mu u_\nu$, making sure that the normalisation $u_\mu u^\mu = c^2$ is conserved and the motion of a massive particle is restricted to the inside of the light cone: Electromagnetic forces can therefore not push a particle outside the light cone and it is impossible to achieve superluminal speeds.

At this point, the Lorentz-geometry arises because of the requirement that the speed of light was equal in al inertial frames, but one might ask if there is a more fundamental reason: As it is, the constancy of $c$ might just be an empirical observation. The truth is very far from that as the Lorentz-geometry is a natural way for Nature to construct hyperbolic partial differential equations as her field equations (where Maxwell's equations or even the gravitational field equation are just examples). Hyperbolic (partial) differential equations are peculiar because they $(i)$ realise a unique time evolution for specified initial conditions, $(ii)$ are perfectly time-invertible and $(iii)$ show causal propagation: There is a finite speed (in our case $c$) at which excitations of the fields travel, and the Lorentzian structure of spacetime makes sure that the light cones are in fact identical in all frames: In this way one can be sure that the initial conditions for the evolution of the fields at a given coordinate are identical in all frames!

That implies that the fundamental Lorentzian structure of spacetime is in fact compatible with the hyperbolicity of the field equations. This is reached by defining a partial differentiation $\partial^\mu$ with respect to the coordinates,

$$\partial^\mu = \begin{pmatrix} \partial^{ct} \\ -\partial^i \end{pmatrix} \tag{A.85}$$

where the minus-sign is added to make sure that the divergence of $x^\mu$ is equal to the dimensionality, i.e. 4:

$$\partial_\mu x^\mu = \frac{\partial x^\mu}{\partial x^\mu} = 4 = \partial_{ct}(ct) + \partial_i x^i = 1 + 3 = \eta_{\mu\nu}\, \partial^\mu x^\nu, \tag{A.86}$$

and the corresponding linear form is given by $\partial_\mu = \eta_{\mu\nu}\partial^\nu = (\partial_{ct}, \partial_i)$. Then, the d'Alembert-operator would be naturally Lorentz-invariant because it is defined as a Lorentz-scalar,

$$\Box = \partial_\mu \partial^\mu = \eta_{\mu\nu}\partial^\mu\partial^\nu = \partial_{ct}^2 - \Delta, \tag{A.87}$$

and typical wave equations like $\Box\Phi = 0$ would generate a light cone, as propagation of excitation proceeds with velocities $\pm c$:

$$\Box\, \Phi = \left(\partial_{ct}^2 - \partial_x^2\right) = \left(\partial_{ct} + \partial_x\right)\left(\partial_{ct} - \partial_x\right) \Phi = 0. \tag{A.88}$$

The same property is reflected by the wave vectors being null: $\Phi = \exp\left(\pm\, \mathrm{i} k_\alpha x^\alpha\right)$ solves the wave equation

$$\Box\, \Phi = \eta^{\mu\nu}\, \partial_\mu \partial_\nu \Phi = 0 \tag{A.89}$$

only of $\eta_{\mu\nu} k^\mu k^\nu = 0$, which holds again in every frame. Giving the components of the wave vector $k^\mu$ the interpretation of the angular frequency $\omega$ and the spatial wave vector $k^i$,

$$k^\mu = \begin{pmatrix} \frac{\omega}{c} \\ k^i \end{pmatrix} \tag{A.90}$$

shows first of all the dispersion-free propagation along the light cone, as the normalisation $\eta_{\mu\nu}k^\mu k^n u = 0$ implies that $\omega^2/c^2 - k^2 = 0$ and therefore a proportionality $\omega = \pm ck$, such that the phase velocity $\omega/k$ and the group velocity $d\omega/dk$ are identical and dispersion is not taking place. Secondly, the (relativistic) Doppler-effect can be derived by projecting $k^\mu$ onto an observer's 4-velocity $u^\mu$. At rest, $u^\mu$ has only a temporal nonzero component of $c$, such that $\omega = \eta_{\mu\nu}u^\mu k^\mu$, but for a moving observer with $u'^\mu$ one obtains

*$k^\mu$ being a null-vector and dispersion-free propagation are equivalent.*

$$\omega' = \eta_{\mu\nu}u'^\mu k^\nu = \gamma(\omega - v_i k^i). \tag{A.91}$$

### A.4.3 Non-relativistic motion in weak gravitational potentials

In anticipation of general relativity we should have a look at changing the geometry of spacetime and to move away from a Lorentzian space. And we need to make sure that the relativistic line element is in fact the relativistic generalisation of the classical Lagrange-function.

Weak gravitational potentials $\Phi = -GM/r$ sourced by a mass M at distance $r$ perturb the line element

$$ds^2 = \left(1 + \frac{2\Phi}{c^2}\right)c^2dt^2 - \left(1 - \frac{2\Phi}{c^2}\right)\delta_{ij}\,dx^i dx^j \tag{A.92}$$

such that one recovers in the Minkowski-metric at large distances $r \gg 2GM/c^2$. If that is the case, the passage of proper time of a stationary observer where $dx^i = 0$ would be dilated

$$ds^2 = c^2d\tau^2 \simeq \left(1 + \frac{2\Phi}{c^2}\right)c^2dt^2 \tag{A.93}$$

and proper time would in fact depend on the presence of gravitational potentials! That would then imply that the variational principle should find a different trajectory if $\Phi$ is nonzero compared to the case $\Phi = 0$. The action would again be given as the line element, but now derived from the actual perturbed metric $g_{\mu\nu}$ instead of the Minkowski-metric $\eta_{\mu\nu}$:

$$S = -mc\int_A^B ds = -mc\int_A^B d\tau\cdot\sqrt{g_{\mu\nu}\,u^\mu u^\nu} \quad \text{using} \quad ds^2 = g_{\mu\nu}\,dx^\mu dx^\nu \tag{A.94}$$

Substituting the 4-velocity $u^\mu$ with the spatial component $\gamma v^i$ then yields for the action

$$S = -mc\int_A^B d\tau\cdot\gamma\sqrt{\left(1 + \frac{2\Phi}{c^2}\right)c^2 - \left(1 - \frac{2\Phi}{c^2}\right)\cdot\delta_{ij}\,v^i v^j} \tag{A.95}$$

which is then approximated to give

$$S \simeq -mc\int_A^B dt\sqrt{c^2\cdot\left(1 + \frac{2\Phi}{c^2} - \delta_{ij}\,\beta^i\beta^j\right)}. \tag{A.96}$$

Finally, a Taylor-expansion yields

$$S =\simeq -mc^2 \int_A^B \mathrm{d}t \left(1 + \frac{\Phi}{c^2} - \frac{\delta_{ij}\,\beta^i\beta^j}{2}\right) \tag{A.97}$$

so that we finally arrive at

$$S = \int_A^B \mathrm{d}t \left(m\delta_{ij}\,\frac{1}{2}\upsilon^i\upsilon^j - m\Phi\right) = \int_A^B \mathrm{d}t\, L \tag{A.98}$$

where we recognise the classical Lagrange function in the integrand, with a kinetic and a potential term.

### A.4.4 Photon propagation on the Lorentzian spacetime

Up to this point, have shown that the archetypical hyperbolic wave equation $\Box\Phi = 0$ is solved in fact by plane waves $\Phi \sim \exp(\pm i\eta_{\mu\nu}\,k^\mu x^\nu)$ with a wave vector $k^\mu$ which is null, $\eta_{\mu\nu}k^\mu k^\nu = 0$. The same should be true for the propagation of electromagnetic waves, so we need to make sure that Maxwell's equations provide a pathway to obtain a hyperbolic wave equation for the field tensor $F^{\mu\nu}$. Specifically, the homogenous Maxwell-equation (or the Bianchi-identiy) should be the relevant here, as electromagnetic waves are vacuum solutions.

$$\partial^\lambda F^{\mu\nu} + \partial^\mu F^{\nu\lambda} + \partial^\nu F^{\lambda\mu} = 0, \tag{A.99}$$

to which one can apply the differentiation $\partial_\lambda$ to obtain

$$\partial_\lambda\partial^\lambda F^{\mu\nu} + \partial_\lambda\partial^\mu F^{\nu\lambda} + \partial_\lambda\partial^\nu F^{\lambda\mu} = 0. \tag{A.100}$$

Identifying the d'Alembert operator $\partial_\lambda\partial^\lambda = \Box$ and using commutativity of partial derivatives, $\partial_\lambda\partial^\mu = \partial^\mu\partial_\lambda$ as well as the antisymmetry of the field tensor, $F^{\nu\lambda} = -F^{\lambda\nu}$ and $\partial_\lambda\partial^\nu = \partial^\nu\partial_\lambda$ this becomes

$$\Box F^{\mu\nu} - \partial^\mu\partial_\lambda F^{\lambda\nu} + \partial^\nu\partial_\lambda F^{\lambda\mu} = 0. \tag{A.101}$$

Now in vacuum, i.e. in the absence of a source $\jmath^\mu = 0$, the field equation is $\partial_\mu F^{\mu\nu} = 0$ and in fact there is a wave-equation with for the field tensor,

$$\Box F^{\mu\nu} = \eta^{\alpha\beta}\,\partial_\alpha\partial_\beta\, F^{\mu\nu} = 0. \tag{A.102}$$

Analogously to the case of a scalar field one expects a plane wave of the type $F^{\mu\nu} \simeq \exp(\pm i\eta_{\gamma\delta}\,k^\gamma x^\delta)$ to solve this equation. Doing that, it is a good idea to use different indices for the differentiation and for the quadratic form $\eta_{\gamma\delta}\,k^\gamma x^\delta$ and to rename the indices with the Kronecker-$\delta$ appearing through $\partial_\alpha x^\mu = \partial x^\mu/\partial x^\alpha = \delta^\mu_\alpha$.

$$\Box F^{\mu\nu} = \eta^{\alpha\beta}\,\partial_\alpha\partial_\beta \exp(\pm i\eta_{\gamma\delta}\,k^\gamma k^\delta) = (\pm i)^2 \cdot \exp\left(\pm i\eta_{\gamma\delta}\,k^\gamma x^\delta\right)\eta_{\gamma\beta}\,k^\gamma k^\beta = 0, \tag{A.103}$$

recovering the null-condition $\eta_{\gamma\beta}\,k^\gamma k^\beta = 0$, confirming that excitations of the electromagnetic field do in fact travel along null-lines, which implies that the Maxwell-equations respect the fundamental Lorentzian structure of spacetime.

The field equation makes sure that the excitations of the fields are perpendicular to the propagation direction and that the wave is indeed transverse: Again, using the ansatz $F^{\mu\nu} = F^{(0),\mu\nu}\exp(\pm i\eta_{\gamma\delta}\,k^\gamma x^\delta))$ one immediately convinces oneself that

$$\partial_\mu F^{\mu\nu} = F^{(0),\mu\nu}\cdot\partial_\mu\exp(\pm i\eta_{\gamma\delta}\,k^\gamma x^\delta) = (\pm i)\exp(\pm i\eta_{\gamma\delta}\,k^\gamma x^\delta)\cdot\eta_{\gamma\mu}F^{(0),\mu\nu}k^\gamma = 0 \quad \text{(A.104)}$$

and therefore $\eta_{\gamma\mu}F^{(0),\mu\nu}k^\gamma = 0$. In terms of the field components of the electric field $E^i$ this means that $\delta_{ij}k^i E^j = 0$, and the analogous statement for the magnetic field $B^i$ would be obtained from the dual field tensor $\eta_{\gamma\mu}\tilde{F}^{(0),\mu\nu}k^\gamma = 0$, as a consequence of electromagnetic duality in vacuum.

Photons move along null-lines, so the arc length measured along their trajectory $x^\mu$ will always come out as zero: That means that one can not work with the proper time $\tau$. Using a new affine parameter $\lambda$ to address the points along the trajectory $x^\mu(\lambda)$ suggests the definition of the wave vector $k^\mu = dx^\mu/d\lambda$, because

$$ds^2 = \eta_{\mu\nu}\,dx^\mu dx^\nu = \eta_{\mu\nu}\,\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}\cdot d\lambda^2 = \eta_{\mu\nu}\,k^\mu k^\nu d\lambda^2 = 0 \quad \text{(A.105)}$$

At this point, we should start to be careful not to link the Lorentz-geometry to any particular coordinate choice. When considering light cone coordinates, $du = c dt + dx$ and $dv = c dt - dx$ the line element is given by

$$ds^2 = \eta_{\mu\nu}\,dx^\mu dx^\nu = c^2 dt^2 - dx^2 = (c dt + dx)(c dt - dx) = du\cdot dv, \quad \text{(A.106)}$$

and the corresponding Lorentzian metric is represented by the matrix

$$\eta_{\mu\nu} = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{(A.107)}$$

in these coordinates. Surely, the geometry is identical and has not been changed by the new definition of coordinates, and the spectrum of eigenvalues of the new metric is identical.

### A.4.5 Photon propagation through weak gravitational fields

At this point we should derive a puzzling result, which was in fact the first proper prediction of general relativity: that gravitational fields have a stronger effect on the motion of relativistic particles such as photons compared to non-relativistic particles. We start by introducing a weak perturbation to the Minkowski-metric and define the line element

$$ds^2 = g_{\mu\nu}\,dx^\mu dx^\nu \quad \text{(A.108)}$$

with $g_{\mu\nu}$ being the metric tensor. For fixed Cartesian coordinates and a weak gravitational potential $\Phi$ with $|\Phi| \ll c^2$ the line element becomes

$$ds^2 = \left(1 + \frac{2\Phi}{c^2}\right)c^2 dt^2 - \left(1 - \frac{2\Phi}{c^2}\right)\delta_{ij}\,dx^i dx^j. \quad \text{(A.109)}$$

It will be the case that for slow motion $\left|\dot{x}^i\right| \ll c$ the classical equation of motion is valid and will come out as $\ddot{x}^i = -\partial^i\Phi$, as expected, by variation of $\int \mathrm{d}\tau$. There is some intuition to this result because a non-relativistic particle moves essentially only along the $ct$-axis of the coordinate frame, so that $\mathrm{d}\tau$ is approximately equal to $(1 + \Phi/c^2)\mathrm{d}t$.

Photons need an entirely different argumentation, because they always follow null-lines, $\mathrm{d}s^2 = 0$. For that case we can define an effective speed of propagation

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \mathrm{c} \cdot \sqrt{\frac{1 + \frac{2\Phi}{\mathrm{c}^2}}{1 - \frac{2\Phi}{\mathrm{c}^2}}} \cong \mathrm{c} \cdot \left(1 + \frac{2\Phi}{\mathrm{c}^2}\right) \tag{A.110}$$

such that we can define an index of refraction, which is proportional to $2\Phi$ instead of $\Phi$! That realisation prompted A. Eddington in 1919 to measure gravitational light deflection during a Solar eclipse and the deflection angle was indeed twice as large as expected from a Newtonian theory.