ITP-UU-07/xx
SPIN-07/xx
arXiv:yymm.nnnn

INTRODUCTION TO GENERAL RELATIVITY

Gerard 't Hooft

Institute for Theoretical Physics
Utrecht University
and

Spinoza Institute
Postbox 80.195
3508 TD Utrecht, the Netherlands

e-mail: g.thooft@phys.uu.nl
internet: http://www.phys.uu.nl/~thooft/

Version 15/9/2007

Prologue

General relativity is a beautiful scheme for describing the gravitational field and the equations it obeys. Nowadays this theory is often used as a prototype for other, more intricate constructions to describe forces between elementary particles or other branches of fundamental physics. This is why in an introduction to general relativity it is of importance to separate as clearly as possible the various ingredients that together give shape to this paradigm. After explaining the physical motivations we first introduce curved coordinates, then add to this the notion of an affine connection field and only as a later step add to that the metric field. One then sees clearly how space and time get more and more structure, until finally all we have to do is deduce Einstein's field equations.

These notes materialized when I was asked to present some lectures on General Relativity. Small changes were made over the years. I decided to make them freely available on the web, via my home page. Some readers expressed their irritation over the fact that after 12 pages I switch notation: the i in the time components of vectors disappears, and the metric becomes the $-+++$ metric. Why this "inconsistency" in the notation?

There were two reasons for this. The transition is made where we proceed from *special* relativity to *general* relativity. In special relativity, the i has a considerable practical advantage: Lorentz transformations are orthogonal, and all inner products only come with $+$ signs. No confusion over signs remain. The use of a $-+++$ metric, or worse even, a $+---$ metric, inevitably leads to sign errors. In *general* relativity, however, the i is superfluous. Here, we need to work with the quantity g_{00} anyway. Choosing it to be negative rarely leads to sign errors or other problems.

But there is another pedagogical point. I see no reason to shield students against the phenomenon of changes of convention and notation. Such transitions are necessary whenever one switches from one field of research to another. They better get used to it.

As for applications of the theory, the usual ones such as the gravitational red shift, the Schwarzschild metric, the perihelion shift and light deflection are pretty standard. They can be found in the cited literature if one wants any further details. In this new version of my lecture notes, mainly chapter 14 was revised, partly due to the recent claims that the effects of a non-vanishing cosmological constant have been detected, but also because I found that the treatment could be adapted more to standard literature on cosmology and at the same time the exposition could be improved. Finally, I do pay extra attention to an application that may well become important in the near future: gravitational radiation. The derivations given are often tedious, but they can be produced rather elegantly using standard Lagrangian methods from field theory, which is what will be demonstrated. When teaching this material, I found that this last chapter is still a bit too technical for an elementary course, but I leave it there anyway, just because it is omitted from introductory text books a bit too often.

I thank A. van der Ven for a careful reading of the manuscript.

Literature

C.W. Misner, K.S. Thorne and J.A. Wheeler, "Gravitation", W.H. Freeman and Comp., San Francisco 1973, ISBN 0-7167-0344-0.

R. Adler, M. Bazin, M. Schiffer, "Introduction to General Relativity", Mc.Graw-Hill 1965.

R. M. Wald, "General Relativity", Univ. of Chicago Press 1984.

P.A.M. Dirac, "General Theory of Relativity", Wiley Interscience 1975.

S. Weinberg, "Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity", J. Wiley & Sons, 1972

S.W. Hawking, G.F.R. Ellis, "The large scale structure of space-time", Cambridge Univ. Press 1973.

S. Chandrasekhar, "The Mathematical Theory of Black Holes", Clarendon Press, Oxford Univ. Press, 1983

Dr. A.D. Fokker, "Relativiteitstheorie", P. Noordhoff, Groningen, 1929.

J.A. Wheeler, "A Journey into Gravity and Spacetime", Scientific American Library, New York, 1990, distr. by W.H. Freeman & Co, New York.

H. Stephani, "General Relativity: An introduction to the theory of the gravitational field", Cambridge University Press, 1990.

Prologue 1
Literature 2

Contents

1 Summary of the theory of Special Relativity. Notations. 4

2 The Eötvös experiments and the Equivalence Principle. 8

3 The constantly accelerated elevator. Rindler Space. 9

4 Curved coordinates. 14

5 The affine connection. Riemann curvature. 19

6 The metric tensor. 25

7 The perturbative expansion and Einstein's law of gravity. 30

8 The action principle. 34

9 Special coordinates. 39

10 Electromagnetism. 42

11 The Schwarzschild solution. 44

12 Mercury and light rays in the Schwarzschild metric. 51

13 Generalizations of the Schwarzschild solution. 55

14 The Robertson-Walker metric. 58

15 Gravitational radiation. 62

1. Summary of the theory of Special Relativity. Notations.

Special Relativity is the theory claiming that space and time exhibit a particular symmetry pattern. This statement contains two ingredients which we further explain:

(i) There is a transformation law, and these transformations form a group.

(ii) Consider a system in which a set of physical variables is described as being a correct solution to the laws of physics. Then if all these physical variables are transformed appropriately according to the given transformation law, one obtains a new solution to the laws of physics.

A "point-event" is a point in space, given by its three coordinates $\vec{x} = (x, y, z)$, at a given instant t in time. For short, we will call this a "point" in space-time, and it is a four component vector,

$$x = \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} . \tag{1.1}$$

Here c is the velocity of light. Clearly, space-time is a four dimensional space. These vectors are often written as x^μ, where μ is an index running from 0 to 3. It will however be convenient to use a slightly different notation, $x^\mu,\ \mu = 1, \dots, 4$, where $x^4 = ict$ and $i = \sqrt{-1}$. Note that we do this *only* in the sections 1 and 3, where special relativity in flat space-time is discussed (see the Prologue). The intermittent use of superscript indices ($\{\}^\mu$) and subscript indices ($\{\}_\mu$) is of no significance in these sections, but will become important later.

In Special Relativity, the transformation group is what one could call the "velocity transformations", or *Lorentz transformations*. It is the set of linear transformations,

$$(x^\mu)' = \sum_{\nu=1}^{4} L^\mu{}_\nu \, x^\nu \tag{1.2}$$

subject to the extra condition that the quantity σ defined by

$$\sigma^2 = \sum_{\mu=1}^{4} (x^\mu)^2 = |\vec{x}|^2 - c^2 t^2 \qquad (\sigma \geq 0) \tag{1.3}$$

remains invariant. This condition implies that the coefficients $L^\mu{}_\nu$ form an orthogonal matrix:

$$\begin{aligned} \sum_{\nu=1}^{4} L^\mu{}_\nu \, L^\alpha{}_\nu &= \delta^{\mu\alpha} ; \\ \sum_{\alpha=1}^{4} L^\alpha{}_\mu \, L^\alpha{}_\nu &= \delta_{\mu\nu} . \end{aligned} \tag{1.4}$$

Because of the i in the definition of x^4, the coefficients $L^i{}_4$ and $L^4{}_i$ must be purely imaginary. The quantities $\delta^{\mu\alpha}$ and $\delta_{\mu\nu}$ are Kronecker delta symbols:

$$\delta^{\mu\nu} = \delta_{\mu\nu} = 1 \quad \text{if} \quad \mu = \nu\,, \qquad \text{and} \;\; 0 \quad \text{otherwise.} \tag{1.5}$$

One can enlarge the invariance group with the *translations:*

$$(x^\mu)' = \sum_{\nu=1}^{4} L^\mu{}_\nu\, x^\nu + a^\mu\,, \tag{1.6}$$

in which case it is referred to as the *Poincaré group*.

We introduce *summation convention*:
If an index occurs exactly twice in a multiplication (at one side of the $=$ sign) it will automatically be summed over from 1 to 4 even if we do not indicate explicitly the summation symbol §. Thus, Eqs. (1.2)–(1.4) can be written as:

$$\begin{aligned} (x^\mu)' &= L^\mu{}_\nu\, x^\nu\,, & \sigma^2 &= x^\mu x^\mu = (x^\mu)^2\,, \\ L^\mu{}_\nu\, L^\alpha{}_\nu &= \delta^{\mu\alpha}\,, & L^\alpha{}_\mu\, L^\alpha{}_\nu &= \delta_{\mu\nu}\,. \end{aligned} \tag{1.7}$$

If we do *not* want to sum over an index that occurs twice, or if we want to sum over an index occurring three times, we put one of the indices between brackets so as to indicate that it does not participate in the summation convention. Greek indices $\mu, \nu, \ldots$ run from 1 to 4; Latin indices $i, j, \ldots$ indicate spacelike components only and hence run from 1 to 3.

A special element of the Lorentz group is

$$L^\mu{}_\nu = \begin{array}{c} \\ \\ \\ \downarrow \\ \mu \end{array} \begin{array}{r} \to \nu \qquad \\ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cosh\chi & i\sinh\chi \\ 0 & 0 & -i\sinh\chi & \cosh\chi \end{pmatrix} \end{array}, \tag{1.8}$$

where χ is a parameter. Or

$$\begin{aligned} x' &= x \quad ; \quad y' = y \quad ; \\ z' &= z\cosh\chi - ct\sinh\chi\,; \\ t' &= -\frac{z}{c}\sinh\chi + t\cosh\chi\,. \end{aligned} \tag{1.9}$$

This is a transformation from one coordinate frame to another with velocity

$$v/c = \tanh\chi \tag{1.10}$$

with respect to each other.

Units of length and time will henceforth be chosen such that

$$c = 1\,. \tag{1.11}$$

Note that the velocity v given in (1.10) will always be less than that of light. The light velocity itself is Lorentz-invariant. This indeed has been the requirement that lead to the introduction of the Lorentz group.

Many physical quantities are not *in*variant but *co*variant under Lorentz transformations. For instance, energy E and momentum p transform as a four-vector:

$$p^\mu = \begin{pmatrix} p_x \\ p_y \\ p_z \\ iE \end{pmatrix} \; ; \quad (p^\mu)' = L^\mu{}_\nu \, p^\nu \; . \tag{1.12}$$

Electro-magnetic fields transform as a tensor:

$$F^{\mu\nu} = \begin{array}{c c} & \overset{\longrightarrow \quad \nu}{} \\ \underset{\mu}{\downarrow} & \begin{pmatrix} 0 & B_3 & -B_2 & -iE_1 \\ -B_3 & 0 & B_1 & -iE_2 \\ B_2 & -B_1 & 0 & -iE_3 \\ iE_1 & iE_2 & iE_3 & 0 \end{pmatrix} \end{array} \; ; \qquad (F^{\mu\nu})' = L^\mu{}_\alpha \, L^\nu{}_\beta \, F^{\alpha\beta} \, . \tag{1.13}$$

It is of importance to realize what this implies: although we have the well-known postulate that an experimenter on a moving platform, when doing some experiment, will find the same outcomes as a colleague at rest, we must *rearrange* the results before comparing them. What could look like an electric field for one observer could be a superposition of an electric and a magnetic field for the other. And so on. This is what we mean with *covariance* as opposed to *invariance*. Much more symmetry groups could be found in Nature than the ones known, if only we knew how to rearrange the phenomena. The *transformation rule* could be very complicated.

We now have formulated the theory of Special Relativity in such a way that it has become very easy to check if some suspect Law of Nature actually obeys Lorentz invariance. Left- and right hand side of an equation must transform the same way, and this is guaranteed if they are written as vectors or tensors with Lorentz indices always transforming as follows:

$$(X'^{\mu\nu\ldots}{}_{\alpha\beta\ldots})' = L^\mu{}_\kappa \, L^\nu{}_\lambda \ldots L_\gamma^{\;\alpha} \, L_\delta^{\;\beta} \ldots X^{\kappa\lambda\ldots}{}_{\gamma\delta\ldots} \, . \tag{1.14}$$

Note that this transformation rule is just as if we were dealing with products of vectors $X^\mu Y^\nu$, etc. Quantities transforming as in Eq. (1.14) are called *tensors*. Due to the orthogonality (1.4) of $L^\mu{}_\nu$ one can multiply and contract tensors covariantly, e.g.:

$$X^\mu = Y_{\mu\alpha} Z^{\alpha\beta\beta} \tag{1.15}$$

is a "tensor" (a tensor with just one index is called a "vector"), if Y and Z are tensors.

The relativistically covariant form of Maxwell's equations is:

$$\begin{aligned} \partial_\mu F_{\mu\nu} &= -J_\nu \; ; & (1.16) \\ \partial_\alpha F_{\beta\gamma} + \partial_\beta F_{\gamma\alpha} + \partial_\gamma F_{\alpha\beta} &= 0 \, ; & (1.17) \\ F_{\mu\nu} &= \partial_\mu A_\nu - \partial_\nu A_\mu \; , & (1.18) \\ \partial_\mu J_\mu &= 0 \; . & (1.19) \end{aligned}$$

Here ∂_μ stands for $\partial/\partial x^\mu$, and the current four-vector J_μ is defined as $J_\mu(x) = (\vec{j}(x),\ ic\varrho(x))$, in units where μ_0 and ε_0 have been normalized to one. A special tensor is $\varepsilon_{\mu\nu\alpha\beta}$, which is defined by

$$\begin{aligned}
\varepsilon_{1234} &= 1\ ; \\
\varepsilon_{\mu\nu\alpha\beta} &= \varepsilon_{\mu\alpha\beta\nu} = -\varepsilon_{\nu\mu\alpha\beta}\ ; \\
\varepsilon_{\mu\nu\alpha\beta} &= 0 \quad \text{if any two of its indices are equal.}
\end{aligned} \tag{1.20}$$

This tensor is invariant under the set of *homogeneous* Lorentz transformations, in fact for all Lorentz transformations $L^\mu{}_\nu$ with $\det(L) = 1$. One can rewrite Eq. (1.17) as

$$\varepsilon_{\mu\nu\alpha\beta}\,\partial_\nu F_{\alpha\beta} = 0\,. \tag{1.21}$$

A particle with mass m and electric charge q moves along a curve $x^\mu(s)$, where s runs from $-\infty$ to $+\infty$, with

$$(\partial_s x^\mu)^2 = -1\ ; \tag{1.22}$$

$$m\,\partial_s^2 x^\mu = q\,F_{\mu\nu}\,\partial_s x^\nu\,. \tag{1.23}$$

The tensor $T^{\text{em}}_{\mu\nu}$ defined by[1]

$$T^{\text{em}}_{\mu\nu} = T^{\text{em}}_{\nu\mu} = F_{\mu\lambda}F_{\lambda\nu} + \tfrac{1}{4}\delta_{\mu\nu}F_{\lambda\sigma}F_{\lambda\sigma}\,, \tag{1.24}$$

describes the energy density, momentum density and mechanical tension of the fields $F_{\alpha\beta}$. In particular the energy density is

$$T^{\text{em}}_{44} = -\tfrac{1}{2}F_{4i}^2 + \tfrac{1}{4}F_{ij}F_{ij} = \tfrac{1}{2}(\vec{E}^2 + \vec{B}^2)\,, \tag{1.25}$$

where we remind the reader that Latin indices $i, j, \dots$ only take the values 1, 2 and 3. Energy and momentum conservation implies that, if at any given space-time point x, we add the contributions of *all* fields and particles to $T_{\mu\nu}(x)$, then for this total energy-momentum tensor, we have

$$\partial_\mu T_{\mu\nu} = 0\,. \tag{1.26}$$

The equation $\partial_0 T_{44} = -\partial_i T_{i0}$ may be regarded as a continuity equation, and so one one must regard the vector T_{i0} as the *energy current*. It is also the *momentum density*, and, in the case of electro-magnetism, it is usually called the Poynting vector. In turn, it obeys the equation $\partial_0 T_{i0} = \partial_j T_{ij}$, so that $-T_{ij}$ can be regarded as the *momentum flow*. However, the time derivative of the momentum is always equal to the force acting on a system, and therefore, T_{ij} can be seen as the force density, or more precisely: the tension, or the force F_i through a unit surface in the direction j. In a neutral gas with pressure p, we have

$$T_{ij} = -p\,\delta_{ij}\ . \tag{1.27}$$

[1] N.B. Sometimes $T_{\mu\nu}$ is defined in different units, so that extra factors 4π appear in the denominator.

2. The Eötvös experiments and the Equivalence Principle.

Suppose that objects made of different kinds of material would react slightly differently to the presence of a gravitational field $\vec{g}$, by having not exactly the same constant of proportionality between gravitational mass and inertial mass:

$$\begin{aligned}
\vec{F}^{(1)} &= M^{(1)}_{\text{inert}}\,\vec{a}^{(1)} = M^{(1)}_{\text{grav}}\,\vec{g}\,, \\
\vec{F}^{(2)} &= M^{(2)}_{\text{inert}}\,\vec{a}^{(2)} = M^{(2)}_{\text{grav}}\,\vec{g}\,; \\
\vec{a}^{(2)} &= \frac{M^{(2)}_{\text{grav}}}{M^{(2)}_{\text{inert}}}\,\vec{g} \neq \frac{M^{(1)}_{\text{grav}}}{M^{(1)}_{\text{inert}}}\,\vec{g} = \vec{a}^{(1)}\,.
\end{aligned} \tag{2.1}$$

These objects would show different accelerations $\vec{a}$ and this would lead to effects that can be detected very accurately. In a space ship, the acceleration would be determined by the material the space ship is made of; any other kind of material would be accelerated differently, and the relative acceleration would be experienced as a weak residual gravitational force. On earth we can also do such experiments. Consider for example a rotating platform with a parabolic surface. A spherical object would be pulled to the center by the earth's gravitational force but pushed to the rim by the centrifugal counter forces of the circular motion. If these two forces just balance out, the object could find stable positions anywhere on the surface, but an object made of different material could still feel a residual force.

Actually the Earth itself is such a rotating platform, and this enabled the Hungarian baron Loránd Eötvös to check extremely accurately the equivalence between inertial mass and gravitational mass (the "Equivalence Principle"). The gravitational force on an object on the Earth's surface is

$$\vec{F}_g = -G_N M_\oplus M_{\text{grav}}\frac{\vec{r}}{r^3}\,, \tag{2.2}$$

where G_N is Newton's constant of gravity, and $M_\oplus$ is the Earth's mass. The centrifugal force is

$$\vec{F}_\omega = M_{\text{inert}}\omega^2\vec{r}_{\text{axis}}\,, \tag{2.3}$$

where ω is the Earth's angular velocity and

$$\vec{r}_{\text{axis}} = \vec{r} - \frac{(\vec{\omega}\cdot\vec{r})\vec{\omega}}{\omega^2} \tag{2.4}$$

is the distance from the Earth's rotational axis. The combined force an object (i) feels on the surface is $\vec{F}^{(i)} = \vec{F}^{(i)}_g + \vec{F}^{(i)}_\omega$. If for two objects, (1) and (2), these forces, $\vec{F}^{(1)}$ and $\vec{F}^{(2)}$, are not exactly parallel, one could measure

$$\alpha = \frac{|\vec{F}^{(1)}\wedge\vec{F}^{(2)}|}{|F^{(1)}||F^{(2)}|} \approx \Big|\frac{M^{(1)}_{\text{inert}}}{M^{(1)}_{\text{grav}}} - \frac{M^{(2)}_{\text{inert}}}{M^{(2)}_{\text{grav}}}\Big|\frac{|\vec{\omega}\wedge\vec{r}|(\vec{\omega}\cdot\vec{r})r}{G_N M_\oplus} \tag{2.5}$$

where we assumed that the gravitational force is much stronger than the centrifugal one. Actually, for the Earth we have:

$$\frac{G_N M_\oplus}{\omega^2 r_\oplus^3} \approx 300\,. \tag{2.6}$$

From (2.5) we see that the misalignment α is given by

$$\alpha \approx (1/300)\cos\theta\sin\theta\left|\frac{M^{(1)}_{\mathrm{inert}}}{M^{(1)}_{\mathrm{grav}}}-\frac{M^{(2)}_{\mathrm{inert}}}{M^{(2)}_{\mathrm{grav}}}\right|, \tag{2.7}$$

where θ is the latitude of the laboratory in Hungary, fortunately sufficiently far from both the North Pole and the Equator.

Eötvös found no such effect, reaching an accuracy of about one part in 10^9 for the equivalence principle. By observing that the Earth also revolves around the Sun one can repeat the experiment using the Sun's gravitational field. The advantage one then has is that the effect one searches for fluctuates daily. This was R.H. Dicke's experiment, in which he established an accuracy of one part in 10^{11}. There are plans to launch a dedicated satellite named STEP (Satellite Test of the Equivalence Principle), to check the equivalence principle with an accuracy of one part in 10^{17}. One expects that there will be no observable deviation. In any case it will be important to formulate a theory of the gravitational force in which the equivalence principle is postulated to hold exactly. Since Special Relativity is also a theory from which never deviations have been detected it is natural to ask for our theory of the gravitational force also to obey the postulates of special relativity. The theory resulting from combining these two demands is the topic of these lectures.

3. The constantly accelerated elevator. Rindler Space.

The equivalence principle implies a new symmetry and associated invariance. The realization of this symmetry and its subsequent exploitation will enable us to give a unique formulation of this gravity theory. This solution was first discovered by Einstein in 1915. We will now describe the modern ways to construct it.

Consider an idealized "elevator", that can make any kinds of vertical movements, including a free fall. When it makes a free fall, all objects inside it will be accelerated equally, according to the Equivalence Principle. This means that during the time the elevator makes a free fall, its inhabitants will not experience any gravitational field at all; they are weightless.

Conversely, we can consider a similar elevator in outer space, far away from any star or planet. Now give it a constant acceleration upward. All inhabitants will feel the pressure from the floor, just as if they were living in the gravitational field of the Earth or any other planet. Thus, we can construct an "artificial" gravitational field. Let us consider such an artificial gravitational field more closely. Suppose we want this artificial gravitational field to be constant in space[2] and time. The inhabitants will feel a constant acceleration.

An essential ingredient in relativity theory is the notion of a coordinate grid. So let us introduce a coordinate grid ξ^μ, $\mu = 1, \dots, 4$, inside the elevator, such that points on its walls are given by ξ^i constant, $i = 1, 2, 3$. The fourth coordinate, ξ^4, is i times the time as measured from the inside of the elevator. An observer in outer space uses a Cartesian grid (inertial frame) x^μ there. The motion of the elevator is described by the functions $x^\mu(\xi)$. Let the origin of the ξ coordinates be a point in the middle of the floor of the elevator, and let it coincide with the origin of the x coordinates. Suppose that we know the acceleration $\vec{g}$ as experienced by the inhabitants of the elevator. How do we determine the functions $x^\mu(\xi)$?

For simplicity, we shall assume that $\vec{g} = (0, 0, g)$, and that $g(\tau) = g$ is constant. We assumed that at $\tau = 0$ the ξ and x coordinates coincide, so

$$\begin{pmatrix} \vec{x}(\vec{\xi}, 0) \\ 0 \end{pmatrix} = \begin{pmatrix} \vec{\xi} \\ 0 \end{pmatrix} . \tag{3.1}$$

Now consider an infinitesimal time lapse, $\mathrm{d}\tau$. After that, the elevator has a velocity $\vec{v} = \vec{g}\,\mathrm{d}\tau$. The middle of the floor of the elevator is now at

$$\begin{pmatrix} \vec{x} \\ it \end{pmatrix} (\vec{0}, i\mathrm{d}\tau) = \begin{pmatrix} \vec{0} \\ i\mathrm{d}\tau \end{pmatrix} . \tag{3.2}$$

But the inhabitants of the elevator will see all other points Lorentz transformed, since they have velocity $\vec{v}$. The Lorentz transformation matrix is only infinitesimally different from the identity matrix:

$$\mathbb{I} + \delta L \;=\; \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -ig\,\mathrm{d}\tau \\ 0 & 0 & ig\,\mathrm{d}\tau & 1 \end{pmatrix} . \tag{3.3}$$

Therefore, the other points $(\vec{\xi}, i\mathrm{d}\tau)$ will be seen at the coordinates $(\vec{x}, it)$ given by

$$\begin{pmatrix} \vec{x} \\ it \end{pmatrix} - \begin{pmatrix} \vec{0} \\ i\mathrm{d}\tau \end{pmatrix} = (\mathbb{I} + \delta L) \begin{pmatrix} \vec{\xi} \\ 0 \end{pmatrix} . \tag{3.4}$$

Now, we perform a little trick. Eq. (3.4) is a Poincaré transformation, that is, a combination of a Lorentz transformation and a translation in time. In many instances (but not always), a Poincaré transformation can be rewritten as a pure Lorentz transformation with respect to a carefully chosen reference point as the origin. Here, we can find such a reference point, by observing that

$$\begin{pmatrix} \vec{0} \\ i\mathrm{d}\tau \end{pmatrix} = \delta L \begin{pmatrix} \vec{g}/g^2 \\ 0 \end{pmatrix} , \tag{3.5}$$

[2] We shall discover shortly, however, that the field we arrive at is constant in the x, y and t direction, but not constant in the direction of the field itself, the z direction.

so that

$$\begin{pmatrix} \vec{x} + \vec{g}/g^2 \\ it \end{pmatrix} = (\mathbb{I} + \delta L) \begin{pmatrix} \vec{\xi} + \vec{g}/g^2 \\ 0 \end{pmatrix} . \tag{3.6}$$

It is important to see what this equation means: after an infinitesimal lapse of time $\mathrm{d}\tau$ inside the elevator, the coordinates $(\vec{x}, it)$ are obtained from the previous set by means of an infinitesimal Lorentz transformation with the point $(-\vec{g}/g^2, 0)$ as its origin. The inhabitants of the elevator van identify this point. Now consider another lapse of time $\mathrm{d}\tau$. Since the elevator is assumed to feel a constant acceleration, the new position can then again be obtained from the old one by means of *the same* Lorentz transformation. So, at time $\tau = N\mathrm{d}\tau$, the coordinates $(\vec{x}, it)$ are given by

$$\begin{pmatrix} \vec{x} + \vec{g}/g^2 \\ it \end{pmatrix} = (\mathbb{I} + \delta L)^N \begin{pmatrix} \vec{\xi} + \vec{g}/g^2 \\ 0 \end{pmatrix} . \tag{3.7}$$

All that remains to be done is compute $(\mathbb{I} + \delta L)^N$. This is not hard:

$$\tau = N\mathrm{d}\tau \ , \quad L(\tau) = (\mathbb{I} + \delta L)^N \ ; \quad L(\tau + \mathrm{d}\tau) = (\mathbb{I} + \delta L) L(\tau) \ ; \tag{3.8}$$

$$\delta L = \begin{pmatrix} 0 & & 0 & \\ & 0 & & \\ 0 & & 0 & -ig \\ & & ig & 0 \end{pmatrix} \mathrm{d}\tau \ ; \qquad L(\tau) = \begin{pmatrix} 1 & & 0 & \\ & 1 & & \\ 0 & & A(\tau) & -iB(\tau) \\ & & iB(\tau) & A(\tau) \end{pmatrix} . \tag{3.9}$$

$$\begin{aligned} L(0) = \mathbb{I} \ ; \quad \mathrm{d}A/\mathrm{d}\tau &= gB \ , \quad \mathrm{d}B/\mathrm{d}\tau = gA \ ; \\ A = \cosh(g\tau) \ , &\quad B = \sinh(g\tau) \, . \end{aligned} \tag{3.10}$$

Combining all this, we derive

$$x^\mu(\vec{\xi}, \, i\tau) \ = \begin{pmatrix} \xi^1 \\ \xi^2 \\ \cosh(g\,\tau)\Big(\xi^3 + \frac{1}{g}\Big) - \frac{1}{g} \\ i\,\sinh(g\,\tau)\Big(\xi^3 + \frac{1}{g}\Big) \end{pmatrix} . \tag{3.11}$$

The 3, 4 components of the ξ coordinates, imbedded in the x coordinates, are pictured in Fig. 1. The description of a quadrant of space-time in terms of the ξ coordinates is called “Rindler space”. From Eq. (3.11) it should be clear that an observer inside the elevator feels no effects that depend explicitly on his time coordinate τ, since a transition from τ to τ' is nothing but a Lorentz transformation. We also notice some important effects:

(i) We see that the equal τ lines converge at the left. It follows that the local clock speed, which is given by $\varrho = \sqrt{-(\partial x^\mu/\partial\tau)^2}$, varies with height ξ^3:

$$\varrho \ = \ 1 + g\,\xi^3 \, , \tag{3.12}$$

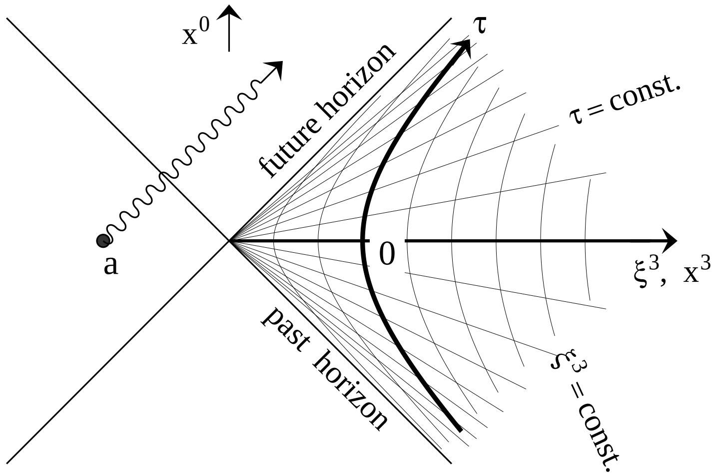

Figure 1: Rindler Space. The curved solid line represents the floor of the elevator, $\xi^3 = 0$. A signal emitted from point a can never be received by an inhabitant of Rindler Space, who lives in the quadrant at the right.

(ii) The gravitational field strength felt locally is $\varrho^{-2}\vec{g}(\xi)$, which is inversely proportional to the distance to the point $x^\mu = -A^\mu$. So even though our field is constant in the transverse direction and with time, it decreases with height.

(iii) The region of space-time described by the observer in the elevator is only part of all of space-time (the quadrant at the right in Fig. 1, where $x^3 + 1/g > |x^0|$). The boundary lines are called (past and future) *horizons*.

All these are typically *relativistic* effects. In the non-relativistic limit ($g \to 0$) Eq. (3.11) simply becomes:

$$x^3 = \xi^3 + \tfrac{1}{2}g\tau^2 \; ; \quad x^4 = i\tau = \xi^4 \, . \tag{3.13}$$

According to the equivalence principle the relativistic effects we discovered here should also be features of gravitational fields generated by matter. Let us inspect them one by one.

Observation (i) suggests that clocks will run slower if they are deep down a gravitational field. Indeed one may suspect that Eq. (3.12) generalizes into

$$\varrho = 1 + V(x) \, , \tag{3.14}$$

where $V(x)$ is the gravitational potential. Indeed this will turn out to be true, provided that the gravitational field is stationary. This effect is called the gravitational red shift.

(ii) is also a relativistic effect. It could have been predicted by the following argument. The *energy density* of a gravitational field is negative. Since the energy of two masses M_1 and M_2 at a distance r apart is $E = -G_N M_1 M_2 / r$ we can calculate the energy density of a field $\vec{g}$ as $T_{44} = -(1/8\pi G_N)\vec{g}^{\,2}$. Since we had normalized $c = 1$ this is also its mass density. But then this mass density in turn should generate a gravitational field! This would imply[3]

$$\vec{\partial}\cdot\vec{g}\ \stackrel{?}{=}\ 4\pi G_N T_{44}\ =\ -\tfrac{1}{2}\vec{g}^{\,2}\ , \tag{3.15}$$

so that indeed the field strength should decrease with height. However this reasoning is apparently too simplistic, since our field obeys a differential equation as Eq. (3.15) but without the coefficient $\frac{1}{2}$.

The possible emergence of *horizons*, our observation (iii), will turn out to be a very important new feature of gravitational fields. Under normal circumstances of course the fields are so weak that no horizon will be seen, but *gravitational collapse* may produce horizons. If this happens there will be regions in space-time from which no signals can be observed. In Fig. 1 we see that signals from a radio station at the point a will never reach an observer in Rindler space.

The most important conclusion to be drawn from this chapter is that in order to describe a gravitational field one may have to perform a transformation from the coordinates ξ^μ that were used inside the elevator where one feels the gravitational field, towards coordinates x^μ that describe empty space-time, in which freely falling objects move along straight lines. Now we know that in an empty space without gravitational fields the clock speeds, and the lengths of rulers, are described by a distance function σ as given in Eq. (1.3). We can rewrite it as

$$\mathrm{d}\sigma^2\ =\ g_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu\ ;\quad g_{\mu\nu}\ =\ \mathrm{diag}(1,1,1,1)\,, \tag{3.16}$$

We wrote here $\mathrm{d}\sigma$ and $\mathrm{d}x^\mu$ to indicate that we look at the infinitesimal distance between two points close together in space-time. In terms of the coordinates ξ^μ appropriate for the elevator we have for infinitesimal displacements $\mathrm{d}\xi^\mu$,

$$\begin{aligned}\mathrm{d}x^3\ &=\ \ \cosh(g\,\tau)\mathrm{d}\xi^3+(1+g\,\xi^3)\sinh(g\,\tau)\mathrm{d}\tau\ ,\\ \mathrm{d}x^4\ &=\ i\sinh(g\,\tau)\mathrm{d}\xi^3+i(1+g\,\xi^3)\cosh(g\,\tau)\mathrm{d}\tau\,.\end{aligned} \tag{3.17}$$

implying

$$\mathrm{d}\sigma^2\ =\ -(1+g\,\xi^3)^2\mathrm{d}\tau^2+(\mathrm{d}\vec{\xi}\,)^2\,. \tag{3.18}$$

If we write this as

$$\mathrm{d}\sigma^2\ =\ g_{\mu\nu}(\xi)\,\mathrm{d}\xi^\mu\mathrm{d}\xi^\nu\ =(\mathrm{d}\vec{\xi}\,)^2+(1+g\,\xi^3)^2(\mathrm{d}\xi^4)^2, \tag{3.19}$$

then we see that all effects that gravitational fields have on rulers and clocks can be described in terms of a space (and time) dependent field $g_{\mu\nu}(\xi)$. *Only in the gravitational field of a Rindler space* can one find coordinates x^μ such that in terms of these the function $g_{\mu\nu}$ takes the simple form of Eq. (3.16). We will see that $g_{\mu\nu}(\xi)$ is all we need to describe the gravitational field completely.

Spaces in which the infinitesimal distance $\mathrm{d}\sigma$ is described by a space(time) dependent function $g_{\mu\nu}(\xi)$ are called *curved* or *Riemann* spaces. Space-time is a Riemann space. We will now investigate such spaces more systematically.

[3]Temporarily we do not show the minus sign usually inserted to indicate that the field is pointed downward.

4. Curved coordinates.

Eq. (3.11) is a special case of a coordinate transformation relevant for inspecting the Equivalence Principle for gravitational fields. It is not a Lorentz transformation since it is not linear in τ. We see in Fig. 1 that the ξ^μ coordinates are curved. The empty space coordinates could be called "straight" because in terms of them all particles move in straight lines. However, such a straight coordinate frame will only exist if the gravitational field has the same Rindler form everywhere, whereas in the vicinity of stars and planets it takes much more complicated forms.

But in the latter case we can also use the Equivalence Principle: the laws of gravity should be formulated in such a way that *any* coordinate frame that uniquely describes the points in our four-dimensional space-time can be used in principle. None of these frames will be superior to any of the others since in any of these frames one will feel some sort of gravitational field[4]. Let us start with just one choice of coordinates $x^\mu = (t, x, y, z)$. From this chapter onwards it will no longer be useful to keep the factor i in the time component because it doesn't simplify things. It has become convention to define $x^0 = t$ and drop the x^4 which was it. So now μ runs from 0 to 3. It will be of importance now that the indices for the coordinates be indicated as *super*scripts $^\mu$, $^\nu$.

Let there now be some one-to-one mapping onto another set of coordinates u^μ,

$$u^\mu \Leftrightarrow x^\mu \; ; \quad x = x(u) \,. \tag{4.1}$$

Quantities depending on these coordinates will simply be called "fields". A *scalar field* ϕ is a quantity that depends on x but does not undergo further transformations, so that in the new coordinate frame (we distinguish the functions of the new coordinates u from the functions of x by using the tilde, $\tilde{}$)

$$\phi = \tilde\phi(u) = \phi(x(u)) \,. \tag{4.2}$$

Now define the gradient (and note that we use a *sub*script index)

$$\phi_\mu(x) = \frac{\partial}{\partial x^\mu}\,\phi(x)\Big|_{x^\nu \text{ constant, for } \nu\neq\mu} \,. \tag{4.3}$$

Remember that the partial derivative is defined by using an infinitesimal displacement $\mathrm{d}x^\mu$,

$$\phi(x + \mathrm{d}x) = \phi(x) + \phi_\mu \mathrm{d}x^\mu + \mathcal{O}(\mathrm{d}x^2) \,. \tag{4.4}$$

We derive

$$\tilde\phi(u + \mathrm{d}u) = \tilde\phi(u) + \frac{\partial x^\mu}{\partial u^\nu}\,\phi_\mu \mathrm{d}u^\nu + \mathcal{O}(\mathrm{d}u^2) = \tilde\phi(u) + \tilde\phi_\nu(u)\mathrm{d}u^\nu \,. \tag{4.5}$$

Therefore in the new coordinate frame the gradient is

$$\tilde\phi_\nu(u) = x^\mu{}_{,\nu}\,\phi_\mu(x(u)) \,, \tag{4.6}$$

[4]There will be some limitations in the sense of continuity and differentiability as we will see.

where we use the notation

$$x^{\mu}{}_{,\nu} \stackrel{\text{def}}{=} \frac{\partial}{\partial u^{\nu}}\, x^{\mu}(u)\Big|_{u^{\alpha\neq\nu}\ \text{constant}}\,, \tag{4.7}$$

so the comma denotes partial derivation.

Notice that in all these equations superscript indices and subscript indices always keep their position and they are used in such a way that in the summation convention one subscript and one superscript occur:

$$\sum_{\mu}(\ldots)_{\mu}(\ldots)^{\mu}$$

Of course one can transform back from the x to the u coordinates:

$$\phi_{\mu}(x) \;=\; u^{\nu}{}_{,\mu}\,\tilde{\phi}_{\nu}(u(x))\,. \tag{4.8}$$

Indeed,

$$u^{\nu}{}_{,\mu}\,x^{\mu}{}_{,\alpha} \;=\; \delta^{\nu}_{\alpha}\,, \tag{4.9}$$

(the matrix $u^{\nu}{}_{,\mu}$ is the inverse of $x^{\mu}{}_{,\alpha}$) A special case would be if the matrix $x^{\mu}{}_{,\alpha}$ would be an element of the Lorentz group. The Lorentz group is just a subgroup of the much larger set of coordinate transformations considered here. We see that $\phi_{\mu}(x)$ transforms as a *vector*. All fields $A_{\mu}(x)$ that transform just like the gradients $\phi_{\mu}(x)$, that is,

$$\tilde{A}_{\nu}(u) \;=\; x^{\mu}{}_{,\nu}\,A_{\mu}(x(u))\,, \tag{4.10}$$

will be called covariant vector fields, co-vector for short, even if they cannot be written as the gradient of a scalar field.

Note that the product of a scalar field ϕ and a co-vector A_{μ} transforms again as a co-vector:

$$\begin{aligned} B_{\mu} &= \phi A_{\mu}\;; \\ \tilde{B}_{\nu}(u) &= \tilde{\phi}(u)\tilde{A}_{\nu}(u) \;=\; \phi(x(u))x^{\mu}{}_{,\nu}A_{\mu}(x(u)) \\ &= x^{\mu}{}_{,\nu}\,B_{\mu}(x(u))\;. \end{aligned} \tag{4.11}$$

Now consider the direct product $B_{\mu\nu} \;=\; A^{(1)}_{\mu}A^{(2)}_{\nu}$. It transforms as follows:

$$\tilde{B}_{\mu\nu}(u) \;=\; x^{\alpha}{}_{,\mu}x^{\beta}{}_{,\nu}\,B_{\alpha\beta}(x(u))\,. \tag{4.12}$$

A collection of field components that can be characterized with a certain number of indices μ, ν, ... and that transforms according to (4.12) is called a *covariant tensor.*

Warning: In a tensor such as $B_{\mu\nu}$ one may not sum over repeated indices to obtain a scalar field. This is because the matrices $x^{\alpha}{}_{,\mu}$ in general do not obey the orthogonality conditions (1.4) of the Lorentz transformations L^{α}_{μ}. One is not advised to sum over two repeated subscript indices. Nevertheless we would like to formulate things such as

Maxwell's equations in General Relativity, and there of course inner products of vectors do occur. To enable us to do this we introduce another type of vectors: the so-called *contravariant* vectors and tensors. Since a contravariant vector transforms differently from a covariant vector we have to indicate this somehow. This we do by putting its indices upstairs: $F^\mu(x)$. The transformation rule for such a superscript index is postulated to be

$$\tilde{F}^\mu(u) = u^\mu{}_{,\alpha}\, F^\alpha(x(u))\,, \tag{4.13}$$

as opposed to the rules (4.10), (4.12) for subscript indices; and contravariant tensors $F^{\mu\nu\alpha\dots}$ transform as products

$$F^{(1)\mu}\, F^{(2)\nu}\, F^{(3)\alpha}\, \dots\,. \tag{4.14}$$

We will also see *mixed* tensors having both upper (superscript) and lower (subscript) indices. They transform as the corresponding products.

> *Exercise:* check that the transformation rules (4.10) and (4.13) form groups, *i.e.* the transformation $x \to u$ yields the same tensor as the sequence $x \to v \to u$. Make use of the fact that partial differentiation obeys

$$\frac{\partial x^\mu}{\partial u^\nu} = \frac{\partial x^\mu}{\partial v^\alpha}\frac{\partial v^\alpha}{\partial u^\nu}\,. \tag{4.15}$$

Summation over repeated indices is admitted if one of the indices is a superscript and one is a subscript:

$$\tilde{F}^\mu(u)\tilde{A}_\mu(u) = u^\mu{}_{,\alpha}\, F^\alpha(x(u))\, x^\beta{}_{,\mu}\, A_\beta(x(u))\,, \tag{4.16}$$

and since the matrix $u^\nu{}_{,\alpha}$ is the inverse of $x^\beta{}_{,\mu}$ (according to 4.9), we have

$$u^\mu{}_{,\alpha}\, x^\beta{}_{,\mu} = \delta^\beta_\alpha\,, \tag{4.17}$$

so that the product $F^\mu A_\mu$ indeed transforms as a scalar:

$$\tilde{F}^\mu(u)\tilde{A}_\mu(u) = F^\alpha(x(u))A_\alpha(x(u))\,. \tag{4.18}$$

Note that since the summation convention makes us sum over repeated indices with the same name, we must ensure in formulae such as (4.16) that indices *not* summed over are each given a *different* name.

We recognize that in Eqs. (4.4) and (4.5) the infinitesimal displacement $\mathrm{d}x^\mu$ of a coordinate transforms as a contravariant vector. This is why coordinates are given superscript indices. Eq. (4.17) also tells us that the Kronecker delta symbol (provided it has one subscript and one superscript index) is an *invariant* tensor: it has the same form in all coordinate grids.

Gradients of tensors

The gradient of a scalar field ϕ transforms as a covariant vector. Are gradients of covariant vectors and tensors again covariant tensors? Unfortunately no. Let us from now on indicate partial differentiation $\partial/\partial x^\mu$ simply as ∂_μ. Sometimes we will use an even shorter notation:

$$\frac{\partial}{\partial x^\mu}\phi = \partial_\mu\phi = \phi_{,\mu}\,. \tag{4.19}$$

From (4.10) we find

$$\begin{aligned}\partial_\alpha \tilde{A}_\nu(u) &= \frac{\partial}{\partial u^\alpha}\tilde{A}_\nu(u) = \frac{\partial}{\partial u^\alpha}\Big(\frac{\partial x^\mu}{\partial u^\nu}A_\mu(x(u))\Big)\\ &= \frac{\partial x^\mu}{\partial u^\nu}\frac{\partial x^\beta}{\partial u^\alpha}\frac{\partial}{\partial x^\beta}A_\mu(x(u)) + \frac{\partial^2 x^\mu}{\partial u^\alpha\partial u^\nu}A_\mu(x(u))\\ &= x^\mu{}_{,\nu}x^\beta{}_{,\alpha}\,\partial_\beta A_\mu(x(u)) + x^\mu{}_{,\alpha,\nu}\,A_\mu(x(u))\,.\end{aligned} \tag{4.20}$$

The last term here deviates from the postulated tensor transformation rule (4.12).

Now notice that

$$x^\mu{}_{,\alpha,\nu} = x^\mu{}_{,\nu,\alpha}\,, \tag{4.21}$$

which always holds for ordinary partial differentiations. From this it follows that the *antisymmetric part* of $\partial_\alpha A_\mu$ *is* a covariant tensor:

$$\begin{aligned}F_{\alpha\mu} &= \partial_\alpha A_\mu - \partial_\mu A_\alpha\,;\\ \tilde{F}_{\alpha\mu}(u) &= x^\beta{}_{,\alpha}x^\nu{}_{,\mu}\,F_{\beta\nu}(x(u))\,.\end{aligned} \tag{4.22}$$

This is an essential ingredient in the mathematical theory of differential forms. We can continue this way: if $A_{\alpha\beta} = -A_{\beta\alpha}$ then

$$F_{\alpha\beta\gamma} = \partial_\alpha A_{\beta\gamma} + \partial_\beta A_{\gamma\alpha} + \partial_\gamma A_{\alpha\beta} \tag{4.23}$$

is a fully antisymmetric covariant tensor.

Next, consider a fully antisymmetric tensor $g_{\mu\nu\alpha\beta}$ having as many indices as the dimensionality of space-time (let's keep space-time four-dimensional). Then one can write

$$g_{\mu\nu\alpha\beta} = \omega\,\varepsilon_{\mu\nu\alpha\beta}\,, \tag{4.24}$$

(see the definition of ε in Eq. (1.20)) since the antisymmetry condition fixes the values of all coefficients of $g_{\mu\nu\alpha\beta}$ apart from one common factor ω. Although ω carries no indices it will turn out not to transform as a scalar field. Instead, we find:

$$\tilde{\omega}(u) = \det(x^\mu{}_{,\nu})\,\omega(x(u))\,. \tag{4.25}$$

A quantity transforming this way will be called a *density*.

The determinant in (4.25) can act as the Jacobian of a transformation in an integral. If $\phi(x)$ is some scalar field (or the inner product of tensors with matching superscript and subscript indices) then the integral

$$\int \omega(x)\phi(x)\mathrm{d}^4 x \tag{4.26}$$

is independent of the choice of coordinates, because

$$\int \mathrm{d}^4 x \ldots = \int \mathrm{d}^4 u \cdot \det(\partial x^\mu / \partial u^\nu) \ldots \ . \tag{4.27}$$

This can also be seen from the definition (4.24):

$$\begin{aligned} \textstyle\int \tilde{g}_{\mu\nu\alpha\beta}\, \mathrm{d}u^\mu \wedge \mathrm{d}u^\nu \wedge \mathrm{d}u^\alpha \wedge \mathrm{d}u^\beta \ = \\ \textstyle\int g_{\kappa\lambda\gamma\delta}\, \mathrm{d}x^\kappa \wedge \mathrm{d}x^\lambda \wedge \mathrm{d}x^\gamma \wedge \mathrm{d}x^\delta \ . \end{aligned} \tag{4.28}$$

Two important properties of tensors are:

1) The decomposition theorem.
Every tensor $X^{\mu\nu\alpha\beta\ldots}_{\kappa\lambda\sigma\tau\ldots}$ can be written as a finite sum of products of covariant and contravariant vectors:

$$X^{\mu\nu\ldots}_{\kappa\lambda\ldots} = \sum_{t=1}^{N} A^\mu_{(t)} B^\nu_{(t)} \ \ldots \ P^{(t)}_\kappa Q^{(t)}_\lambda \ \ldots \ . \tag{4.29}$$

The number of terms, N, does not have to be larger than the number of components of the tensor[5]. By choosing in one coordinate frame the vectors A, B, ... each such that they are non vanishing for only one value of the index the proof can easily be given.

2) The quotient theorem.
Let there be given an arbitrary set of components $X^{\mu\nu\ldots\alpha\beta\ldots}_{\kappa\lambda\ldots\sigma\tau\ldots}$. Let it be known that for *all* tensors $A^{\sigma\tau\ldots}_{\alpha\beta\ldots}$ (with a *given, fixed* number of superscript and/or subscript indices) the quantity

$$B^{\mu\nu\ldots}_{\kappa\lambda\ldots} = X^{\mu\nu\ldots\alpha\beta\ldots}_{\kappa\lambda\ldots\sigma\tau\ldots} A^{\sigma\tau\ldots}_{\alpha\beta\ldots}$$

transforms as a tensor. Then it follows that X itself also transforms as a tensor.

The proof can be given by induction. First one chooses A to have just one index. Then in one coordinate frame we choose it to have just one non-vanishing component. One then uses (4.9) or (4.17). If A has several indices one decomposes it using the decomposition theorem.

[5] If n is the dimensionality of spacetime, and r the number of indices (the *rank* of the tensor), then one needs at most $N \le n^{r-1}$ terms.

What has been achieved in this chapter is that we learned to work with tensors in curved coordinate frames. They can be differentiated and integrated. But before we can construct physically interesting theories in curved spaces two more obstacles will have to be overcome:

(i) Thus far we have only been able to differentiate *antisymmetrically*, otherwise the resulting gradients do not transform as tensors.

(ii) There still are two types of indices. Summation is only permitted if one index is a superscript and one is a subscript index. This is too much of a limitation for constructing covariant formulations of the existing laws of nature, such as the Maxwell laws. We shall deal with these obstacles one by one.

5. The affine connection. Riemann curvature.

The space described in the previous chapter does not yet have enough structure to formulate all known physical laws in it. For a good understanding of the structure now to be added we first must define the notion of "affine connection". Only in the next chapter we will define distances in time and space.

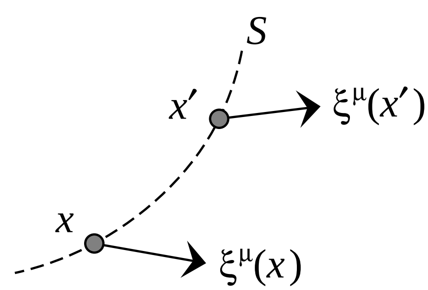

Figure 2: Two contravariant vectors close to each other on a curve S .

Let $\xi^\mu(x)$ be a contravariant vector field, and let $x^\mu(\tau)$ be the space-time trajectory S of an observer. We now assume that the observer has a way to establish whether $\xi^\mu(x)$ is constant or varies as his eigentime τ goes by. Let us indicate the observed time derivative by a dot:

$$\dot{\xi}^\mu = \frac{\mathrm{d}}{\mathrm{d}\tau}\xi^\mu(x(\tau))\,. \tag{5.1}$$

The observer will have used a coordinate frame x where he stays at the origin O of three-space. What will equation (5.1) be like in some other coordinate frame u?

$$\begin{aligned} \xi^\mu(x) &= x^\mu{}_{,\nu}\,\tilde{\xi}^\nu(u(x))\,; \\ x^\mu{}_{,\nu}\dot{\tilde{\xi}}^\nu &\stackrel{\text{def}}{=} \frac{\mathrm{d}}{\mathrm{d}\tau}\xi^\mu(x(\tau)) = x^\mu{}_{,\nu}\frac{\mathrm{d}}{\mathrm{d}\tau}\tilde{\xi}^\nu\Big(u(x(\tau))\Big) + x^\mu{}_{,\nu,\lambda}\frac{\mathrm{d}u^\lambda}{\mathrm{d}\tau}\cdot\tilde{\xi}^\nu(u)\,. \end{aligned} \tag{5.2}$$

Thus, if we wish to define a quantity $\dot{\xi}^\nu$ that transforms as a contravector then in a general coordinate frame this is to be written as

$$\dot{\xi}^\nu(u(\tau)) \stackrel{\text{def}}{=} \frac{\mathrm{d}}{\mathrm{d}\tau}\xi^\nu(u(\tau)) + \Gamma^\nu_{\kappa\lambda}\frac{\mathrm{d}u^\lambda}{\mathrm{d}\tau}\xi^\kappa(u(\tau))\,. \tag{5.3}$$

Here, $\Gamma^\nu_{\lambda\kappa}$ is a new field, and near the point u the local observer can use a "preferred coordinate frame" x such that

$$u^\nu{}_{,\mu} x^\mu{}_{,\kappa,\lambda} = \Gamma^\nu_{\kappa\lambda}\,. \tag{5.4}$$

In his preferred coordinate frame, Γ will vanish, but *only on his curve* S *!* In general it will not be possible to find a coordinate frame such that Γ vanishes everywhere. Eq. (5.3) defines the *parallel displacement of a contravariant vector* along a curve S. To do this a new field was introduced, $\Gamma^\mu_{\lambda\kappa}(u)$, called "affine connection field" by Levi-Civita. It is a field, but *not* a tensor field, since it transforms as

$$\tilde{\Gamma}^\nu_{\kappa\lambda}(u(x)) \;=\; u^\nu{}_{,\mu}\Big[x^\alpha{}_{,\kappa}x^\beta{}_{,\lambda}\Gamma^\mu_{\alpha\beta}(x) + x^\mu{}_{,\kappa,\lambda}\Big]\,. \tag{5.5}$$

> *Exercise:* Prove (5.5) and show that two successive transformations of this type again produces a transformation of the form (5.5).

We now observe that Eq. (5.4) implies

$$\Gamma^\nu_{\lambda\kappa} \;=\; \Gamma^\nu_{\kappa\lambda}\,, \tag{5.6}$$

and since

$$x^\mu{}_{,\kappa,\lambda} \;=\; x^\mu{}_{,\lambda,\kappa}\,, \tag{5.7}$$

this symmetry will also hold in any other coordinate frame. Now, in principle, one can consider spaces with a parallel displacement according to (5.3) where Γ does not obey (5.6). In this case there are no local inertial frames where in some given point x one has $\Gamma^\mu_{\lambda\kappa} = 0$. This is called *torsion.* We will not pursue this, apart from noting that the antisymmetric part of $\Gamma^\mu_{\kappa\lambda}$ would be an ordinary tensor field, which could always be added to our models at a later stage. So we limit ourselves now to the case that Eq. (5.6) always holds.

A *geodesic* is a curve $x^\mu(\sigma)$ that obeys

$$\frac{\mathrm{d}^2}{\mathrm{d}\sigma^2}x^\mu(\sigma) + \Gamma^\mu_{\kappa\lambda}\frac{\mathrm{d}x^\kappa}{\mathrm{d}\sigma}\frac{\mathrm{d}x^\lambda}{\mathrm{d}\sigma} \;=\; 0\,. \tag{5.8}$$

Since $\mathrm{d}x^\mu/\mathrm{d}\sigma$ is a contravariant vector this is a special case of Eq. (5.3) and the equation for the curve will look the same in all coordinate frames.

N.B. If one chooses an arbitrary, different parametrization of the curve (5.8), using a parameter $\tilde{\sigma}$ that is an arbitrary differentiable function of σ, one obtains a different equation,

$$\frac{\mathrm{d}^2}{\mathrm{d}\tilde{\sigma}^2}x^\mu(\tilde{\sigma}) + \alpha(\tilde{\sigma})\frac{\mathrm{d}}{\mathrm{d}\tilde{\sigma}}x^\mu(\tilde{\sigma}) + \Gamma^\mu_{\kappa\lambda}\frac{\mathrm{d}x^\kappa}{\mathrm{d}\tilde{\sigma}}\frac{\mathrm{d}x^\lambda}{\mathrm{d}\tilde{\sigma}} \;=\; 0\;. \tag{5.8a}$$

where $\alpha(\tilde{\sigma})$ can be any function of $\tilde{\sigma}$. Apparently the shape of the curve in coordinate space does not depend on the function $\alpha(\tilde{\sigma})$.

Exercise: check Eq. (5.8a).

Curves described by Eq. (5.8) could be defined to be the space-time trajectories of particles moving in a gravitational field. Indeed, in every point x there exists a coordinate frame such that Γ vanishes there, so that the trajectory goes straight (the coordinate frame of the freely falling elevator). In an accelerated elevator, the trajectories look curved, and an observer inside the elevator can attribute this curvature to a gravitational field. The gravitational field is hereby identified as an affine connection field.

Since now we have a field that transforms according to Eq. (5.5) we can use it to eliminate the offending last term in Eq. (4.20). We define a *covariant derivative* of a co-vector field:

$$D_\alpha A_\mu \;=\; \partial_\alpha A_\mu - \Gamma^\nu_{\alpha\mu} A_\nu \,. \tag{5.9}$$

This quantity $D_\alpha A_\mu$ neatly transforms as a tensor:

$$D_\alpha \tilde{A}_\nu(u) \;=\; x^\mu{}_{,\nu} x^\beta{}_{,\alpha}\, D_\beta A_\mu(x)\,. \tag{5.10}$$

Notice that

$$D_\alpha A_\mu - D_\mu A_\alpha \;=\; \partial_\alpha A_\mu - \partial_\mu A_\alpha\,, \tag{5.11}$$

so that Eq. (4.22) is kept unchanged.

Similarly one can now define the covariant derivative of a contravariant vector:

$$D_\alpha A^\mu \;=\; \partial_\alpha A^\mu + \Gamma^\mu_{\alpha\beta} A^\beta\,. \tag{5.12}$$

(notice the differences with (5.9)!) It is not difficult now to define covariant derivatives of all other tensors:

$$\begin{aligned} D_\alpha X^{\mu\nu\dots}_{\kappa\lambda\dots} \;=\;& \partial_\alpha X^{\mu\nu\dots}_{\kappa\lambda\dots} + \Gamma^\mu_{\alpha\beta} X^{\beta\nu\dots}_{\kappa\lambda\dots} + \Gamma^\nu_{\alpha\beta} X^{\mu\beta\dots}_{\kappa\lambda\dots} \;\dots \\ & - \Gamma^\beta_{\kappa\alpha} X^{\mu\nu\dots}_{\beta\lambda\dots} - \Gamma^\beta_{\lambda\alpha} X^{\mu\nu\dots}_{\kappa\beta\dots} \;\dots\,. \end{aligned} \tag{5.13}$$

Expressions (5.12) and (5.13) also transform as tensors.

We also easily verify a "product rule". Let the tensor Z be the product of two tensors X and Y:

$$Z^{\kappa\lambda\dots\pi\varrho\dots}_{\mu\nu\dots\alpha\beta\dots} \;=\; X^{\kappa\lambda\dots}_{\mu\nu\dots}\, Y^{\pi\varrho\dots}_{\alpha\beta\dots}\,. \tag{5.14}$$

Then one has (in a notation where we temporarily suppress the indices)

$$D_\alpha Z \;=\; (D_\alpha X)Y + X(D_\alpha Y)\,. \tag{5.15}$$

Furthermore, if one sums over repeated indices (one subscript and one superscript, we will call this a *contraction* of indices):

$$(D_\alpha X)^{\mu\kappa\dots}_{\mu\beta\dots} \;=\; D_\alpha (X^{\mu\kappa\dots}_{\mu\beta\dots})\,, \tag{5.16}$$

so that we can just as well omit the brackets in (5.16). Eqs. (5.15) and (5.16) can easily be proven to hold in any point x, by choosing the reference frame where Γ vanishes at that point x.

The covariant derivative of a scalar field ϕ is the ordinary derivative:

$$D_\alpha \phi \;=\; \partial_\alpha \phi\,, \tag{5.17}$$

but this does not hold for a density function ω (see Eq. (4.24),

$$D_\alpha \omega \;=\; \partial_\alpha \omega - \Gamma^\mu_{\mu\alpha}\omega\,. \tag{5.18}$$

$D_\alpha \omega$ is a density times a covector. This one derives from (4.24) and

$$\varepsilon^{\alpha\mu\nu\lambda}\varepsilon_{\beta\mu\nu\lambda} \;=\; 6\,\delta^\alpha_\beta\,. \tag{5.19}$$

Thus we have found that if one introduces in a space or space-time a field $\Gamma^\mu_{\nu\lambda}$ that transforms according to Eq. (5.5), called 'affine connection', then one can define: 1) geodesic curves such as the trajectories of freely falling particles, and 2) the covariant derivative of any vector and tensor field. But what we do *not* yet have is (i) a unique definition of *distance* between points and (ii) a way to identify *co* vectors with *contra* vectors. Summation over repeated indices only makes sense if one of them is a superscript and the other is a subscript index.

Curvature

Now again consider a curve S as in Fig. 2, but close it (Fig. 3). Let us have a contravector field $\xi^\nu(x)$ with

$$\dot\xi^\nu(x(\tau)) \;=\; 0\,; \tag{5.20}$$

We take the curve to be very small[6] so that we can write

$$\xi^\nu(x) \;=\; \xi^\nu + \xi^\nu_{\;,\mu}x^\mu + \mathcal{O}(x^2)\,. \tag{5.21}$$

Will this contravector return to its original value if we follow it while going around the curve one full loop? According to (5.3) it certainly will if the connection field vanishes: $\Gamma = 0$. But if there is a strong gravity field there might be a deviation $\delta\xi^\nu$. We find:

$$\begin{aligned}
\oint \mathrm{d}\tau\dot\xi \;&=\; 0\,;\\
\delta\xi^\nu \;&=\; \oint \mathrm{d}\tau\frac{\mathrm{d}}{\mathrm{d}\tau}\xi^\nu(x(\tau)) \;=\; -\oint \Gamma^\nu_{\kappa\lambda}\frac{\mathrm{d}x^\lambda}{\mathrm{d}\tau}\xi^\kappa(x(\tau))\mathrm{d}\tau\\
&=\; -\oint \mathrm{d}\tau\Big(\Gamma^\nu_{\kappa\lambda} + \Gamma^\nu_{\kappa\lambda,\,\alpha}x^\alpha\Big)\frac{\mathrm{d}x^\lambda}{\mathrm{d}\tau}\Big(\xi^\kappa + \xi^\kappa_{\;,\mu}x^\mu\Big)\,.
\end{aligned} \tag{5.22}$$

[6] In an affine space without metric the words 'small' and 'large' appear to be meaningless. However, since differentiability is required, the small size limit *is* well defined. Thus, it is more precise to state that the curve is *in*finitesimally small.

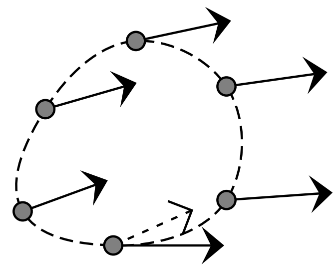

Figure 3: Parallel displacement along a closed curve in a curved space.

where we chose the function $x(\tau)$ to be very small, so that terms $\mathcal{O}(x^2)$ could be neglected. We have a closed curve, so

$$\oint \mathrm{d}\tau \tfrac{\mathrm{d}x^\lambda}{\mathrm{d}\tau} = 0 \qquad \text{and}$$
$$D_\mu \xi^\kappa \approx 0 \;\rightarrow\; \xi^\kappa_{\ ,\mu} \approx -\Gamma^\kappa_{\ \mu\beta}\xi^\beta\,, \tag{5.23}$$

so that Eq. (5.22) becomes

$$\delta\xi^\nu = \tfrac{1}{2}\Big(\oint x^\alpha \frac{\mathrm{d}x^\lambda}{\mathrm{d}\tau}\mathrm{d}\tau\Big) R^\nu_{\ \kappa\lambda\alpha}\xi^\kappa \;+\; \text{higher orders in } x\,. \tag{5.24}$$

Since

$$\oint x^\alpha \frac{\mathrm{d}x^\lambda}{\mathrm{d}\tau}\mathrm{d}\tau + \oint x^\lambda \frac{\mathrm{d}x^\alpha}{\mathrm{d}\tau}\mathrm{d}\tau = 0\,, \tag{5.25}$$

only the antisymmetric part of R matters. We choose

$$R^\nu_{\ \kappa\lambda\alpha} = -R^\nu_{\ \kappa\alpha\lambda} \tag{5.26}$$

(the factor $\frac{1}{2}$ in (5.24) is conventionally chosen this way). Thus we find:

$$R^\nu_{\ \kappa\lambda\alpha} = \partial_\lambda \Gamma^\nu_{\ \kappa\alpha} - \partial_\alpha \Gamma^\nu_{\ \kappa\lambda} + \Gamma^\nu_{\ \lambda\sigma}\Gamma^\sigma_{\ \kappa\alpha} - \Gamma^\nu_{\ \alpha\sigma}\Gamma^\sigma_{\ \kappa\lambda}\,. \tag{5.27}$$

We now claim that this quantity must transform as a true tensor. This should be surprising since Γ itself is not a tensor, and since there are ordinary derivatives ∂_λ instead of covariant derivatives. The argument goes as follows. In Eq. (5.24) the l.h.s., $\delta\xi^\nu$ is a true contravector, and also the quantity

$$S^{\alpha\lambda} = \oint x^\alpha \frac{\mathrm{d}x^\lambda}{\mathrm{d}\tau}\mathrm{d}\tau\,, \tag{5.28}$$

transforms as a tensor. Now we can choose ξ^κ any way we want and also the surface elements $S^{\alpha\lambda}$ may be chosen freely. Therefore we may use the quotient theorem (expanded to cover the case of antisymmetric tensors) to conclude that in that case the set of coefficients $R^\nu_{\ \kappa\lambda\alpha}$ must also transform as a genuine tensor. Of course we can check explicitly by using (5.5) that the combination (5.27) indeed transforms as a tensor, showing that the inhomogeneous terms cancel out.

$R^{\nu}{}_{\kappa\lambda\alpha}$ tells us something about the extent to which this space is curved. It is called the *Riemann curvature tensor.* From (5.27) we derive

$$R^{\nu}{}_{\kappa\lambda\alpha}+R^{\nu}{}_{\lambda\alpha\kappa}+R^{\nu}{}_{\alpha\kappa\lambda}\;=\;0\,, \tag{5.29}$$

and

$$D_{\alpha}R^{\nu}{}_{\kappa\beta\gamma}+D_{\beta}R^{\nu}{}_{\kappa\gamma\alpha}+D_{\gamma}R^{\nu}{}_{\kappa\alpha\beta}\;=\;0\,. \tag{5.30}$$

The latter equation, called Bianchi identity, can be derived most easily by noting that for every point x a coordinate frame exists such that at that point x one has $\Gamma^{\nu}_{\kappa\alpha}=0$ (though its derivative $\partial\Gamma$ cannot be tuned to zero). One then only needs to take into account those terms of Eq. (5.30) that are linear in $\partial\Gamma$.

Partial derivatives ∂_μ have the property that the order may be interchanged, $\partial_\mu\partial_\nu=\partial_\nu\partial_\mu$. This is no longer true for covariant derivatives. For any covector field $A_\mu(x)$ we find

$$D_\mu D_\nu A_\alpha-D_\nu D_\mu A_\alpha\;=\;-R^{\lambda}{}_{\alpha\mu\nu}A_\lambda\,, \tag{5.31}$$

and for any contravector field A^α:

$$D_\mu D_\nu A^\alpha-D_\nu D_\mu A^\alpha\;=\;R^{\alpha}{}_{\lambda\mu\nu}A^\lambda\,, \tag{5.32}$$

which we can verify directly from the definition of $R^{\lambda}{}_{\alpha\mu\nu}$. These equations also show clearly why the Riemann curvature transforms as a true tensor; (5.31) and (5.32) hold for all A_λ and A^λ and the l.h.s. transform as tensors.

An important theorem is that the Riemann tensor *completely* specifies the extent to which space or space-time is curved, if this space-time is simply connected. We shall not give a mathematically rigorous proof of this, but an acceptable argument can be found as follows. Assume that $R^{\nu}{}_{\kappa\lambda\alpha}=0$ everywhere. Consider then a point x and a coordinate frame such that $\Gamma^{\nu}_{\kappa\lambda}(x)=0$. We assume our manifold to be C_∞ at the point x. Then consider a Taylor expansion of Γ around x:

$$\Gamma^{\nu}_{\kappa\lambda}(x')=\Gamma^{[1]\nu}_{\kappa\lambda,\,\alpha}(x'-x)^\alpha+\tfrac{1}{2}\Gamma^{[2]\nu}_{\kappa\lambda,\,\alpha\beta}(x'-x)^\alpha(x'-x)^\beta\;\ldots\,, \tag{5.33}$$

From the fact that (5.27) vanishes we deduce that $\Gamma^{[1]\nu}_{\kappa\lambda,\,\alpha}$ is symmetric:

$$\Gamma^{[1]\nu}_{\kappa\lambda,\,\alpha}=\Gamma^{[1]\nu}_{\kappa\alpha,\lambda}\,, \tag{5.34}$$

and furthermore, from the symmetry (5.6) we have

$$\Gamma^{[1]\nu}_{\kappa\lambda,\,\alpha}=\Gamma^{[1]\nu}_{\lambda\kappa,\,\alpha}\,, \tag{5.35}$$

so that there is complete symmetry in the lower indices. From this we derive that

$$\Gamma^{\nu}_{\kappa\lambda}=\partial_\lambda\partial_k Y^\nu+\mathcal{O}(x'-x)^2\,, \tag{5.36}$$

with

$$Y^\nu = \tfrac{1}{6}\Gamma^{[1]\nu}_{\kappa\lambda,\alpha}(x'-x)^\alpha(x'-x)^\lambda(x'-x)^\kappa\,. \tag{5.37}$$

If now we turn to the coordinates $u^\mu = x^\mu + Y^\mu$ then, according to the transformation rule (5.5), Γ vanishes in these coordinates up to terms of order $(x'-x)^2$. So, here, the coefficients $\Gamma^{[1]}$ vanish.

The argument can now be repeated to prove that, in (5.33), all coefficients $\Gamma^{[i]}$ can be made to vanish by choosing suitable coordinates. Unless our space-time were extremely singular at the point x, one finds a domain this way around x where, given suitable coordinates, Γ vanish completely. All domains treated this way can be glued together, and only if there is an obstruction because our space-time isn't simply-connected, this leads to coordinates where the Γ vanish everywhere.

Thus we see that if the Riemann curvature vanishes a coordinate frame can be constructed in terms of which all geodesics are straight lines and all covariant derivatives are ordinary derivatives. This is a flat space.

Warning: there is no universal agreement in the literature about sign conventions in the definitions of $\mathrm{d}\sigma^2$, $\Gamma^\nu_{\kappa\lambda}$, $R^\nu{}_{\kappa\lambda\alpha}$, $T_{\mu\nu}$ and the field $g_{\mu\nu}$ of the next chapter. This should be no impediment against studying other literature. One frequently has to adjust signs and pre-factors.

6. The metric tensor.

In a space with affine connection we have geodesics, but no clocks and rulers. These we will introduce now. In Chapter 3 we saw that in flat space one has a matrix

$$g_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{6.1}$$

so that for the Lorentz invariant distance σ we can write

$$\sigma^2 \;=\; -t^2 + \vec{x}^{\,2} \;=\; g_{\mu\nu}x^\mu x^\nu\,. \tag{6.2}$$

(time will be the zeroth coordinate, which is agreed upon to be the convention if all coordinates are chosen to stay real numbers). For a particle running along a timelike curve $C = \{x(\sigma)\}$ the increase in eigentime T is

$$\begin{aligned} T = \int_C \mathrm{d}T\,, \quad \text{with} \quad \mathrm{d}T^2 \;&=\; -g_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\sigma}\frac{\mathrm{d}x^\nu}{\mathrm{d}\sigma}\cdot\mathrm{d}\sigma^2 \\ &\stackrel{\text{def}}{=}\; -\,g_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu\,. \end{aligned} \tag{6.3}$$

This expression is coordinate independent, provided that $g_{\mu\nu}$ is treated as a co-tensor with two subscript indices. It is symmetric under interchange of these. In curved coordinates we get

$$g_{\mu\nu} \;=\; g_{\nu\mu} \;=\; g_{\mu\nu}(x)\,. \tag{6.4}$$

This is the metric tensor field. Only far away from stars and planets we can find coordinates such that it will coincide with (6.1) everywhere. In general it will deviate from this slightly, but usually not very much. In particular we will demand that upon diagonalization one will always find three positive and one negative eigenvalue. This property can be shown to be unchanged under coordinate transformations. The inverse of $g_{\mu\nu}$ which we will simply refer to as $g^{\mu\nu}$ is uniquely defined by

$$g_{\mu\nu}g^{\nu\alpha} = \delta^\alpha_\mu \,. \tag{6.5}$$

This inverse is also symmetric under interchange of its indices.

It now turns out that the introduction of such a two-index co-tensor field gives space-time *more structure* than the three-index affine connection of the previous chapter. First of all, the tensor $g_{\mu\nu}$ *induces one special choice for the affine connection field.* Let us elucidate this first by using a physical argument. Consider a freely falling elevator (or spaceship). Assume that the elevator is so small that the gravitational pull from stars and planets surrounding it appears to be the same everywhere inside the elevator. Then an observer inside the elevator will not experience any gravitational field anywhere inside the elevator. He or she should be able to introduce a Cartesian coordinate grid inside the elevator, as if gravitational forces did not exist. He or she could use as metric tensor $g_{\mu\nu} = \mathrm{diag}(-1, 1, 1, 1)$. Since there is no gravitational field, clocks run equally fast everywhere, and rulers show the same lengths everywhere (as long as we stay inside the elevator). Therefore, the inhabitant must conclude that $\partial_\alpha g_{\mu\nu} = 0$. Since there is no need of curved coordinates, one would also have $\Gamma^\lambda_{\mu\nu} = 0$ at the location of the elevator. Note: the *gradient* of Γ, and the *second derivative* of $g_{\mu\nu}$ would be difficult to detect, so we put no constraints on those.

Clearly, we conclude that, at the location of the elevator, the covariant derivative of $g_{\mu\nu}$ should vanish:

$$D_\alpha g_{\mu\nu} = 0 \,. \tag{6.6}$$

In fact, we shall now argue that Eq. (6.6) can be used as a *definition* of the affine connection Γ for a space or space-time where a metric tensor $g_{\mu\nu}(x)$ is given. This argument goes as follows.

From (6.6) we see:

$$\partial_\alpha g_{\mu\nu} = \Gamma^\lambda_{\alpha\mu} g_{\lambda\nu} + \Gamma^\lambda_{\alpha\nu} g_{\mu\lambda} \,. \tag{6.7}$$

Write

$$\Gamma_{\lambda\alpha\mu} = g_{\lambda\nu}\Gamma^\nu_{\alpha\mu} \,, \tag{6.8}$$

$$\Gamma_{\lambda\alpha\mu} = \Gamma_{\lambda\mu\alpha} \,. \tag{6.9}$$

Then one finds from (6.7)

$$\tfrac{1}{2}(\partial_\mu g_{\lambda\nu} + \partial_\nu g_{\lambda\mu} - \partial_\lambda g_{\mu\nu}) = \Gamma_{\lambda\mu\nu} \,, \tag{6.10}$$

$$\Gamma^\lambda_{\mu\nu} = g^{\lambda\alpha}\Gamma_{\alpha\mu\nu} \,. \tag{6.11}$$

These equations now *define* an affine connection field. Indeed Eq. (6.6) follows from (6.10), (6.11). In the literature one also finds the "Christoffel symbol" $\left\{ {\mu \atop \kappa\lambda} \right\}$ which means the same thing. The convention used here is that of Hawking and Ellis. Since

$$D_\alpha \delta^\lambda_\mu \,=\, \partial_\alpha \delta^\lambda_\mu \,=\, 0\,, \tag{6.12}$$

we also have for the inverse of $g_{\mu\nu}$

$$D_\alpha g^{\mu\nu} \,=\, 0\,, \tag{6.13}$$

which follows from (6.5) in combination with the product rule (5.15).

But the metric tensor $g_{\mu\nu}$ not only gives us an affine connection field, it now also enables us to replace subscript indices by superscript indices and back. For every covector $A_\mu(x)$ we define a contravector $A^\nu(x)$ by

$$A_\mu(x) \,=\, g_{\mu\nu}(x) A^\nu(x)\ ; \qquad A^\nu \,=\, g^{\nu\mu} A_\mu\,. \tag{6.14}$$

Very important is what is implied by the product rule (5.15), together with (6.6) and (6.13):

$$\begin{aligned} D_\alpha A^\mu &= g^{\mu\nu} D_\alpha A_\nu\ , \\ D_\alpha A_\mu &= g_{\mu\nu} D_\alpha A^\nu\ . \end{aligned} \tag{6.15}$$

It follows that raising or lowering indices by multiplication with $g_{\mu\nu}$ or $g^{\mu\nu}$ can be done before or after covariant differentiation.

The metric tensor also generates a density function ω:

$$\omega \,=\, \sqrt{-\det(g_{\mu\nu})}\,. \tag{6.16}$$

It transforms according to Eq. (4.25). This can be understood by observing that in a coordinate frame with in some point x

$$g_{\mu\nu}(x) \,=\, \mathrm{diag}(-a, b, c, d)\,, \tag{6.17}$$

the volume element is given by $\sqrt{abcd}$.

The space of the previous chapter is called an "affine space". In the present chapter we have a subclass of the affine spaces called a metric space or Riemann space; indeed we can call it a Riemann space-time. The presence of a time coordinate is betrayed by the one negative eigenvalue of $g_{\mu\nu}$.

The geodesics

Consider two arbitrary points X and Y in our metric space. For every curve $C = \{x^\mu(\sigma)\}$ that has X and Y as its end points,

$$x^\mu(0) \,=\, X^\mu\ ; \qquad x^\mu(1) \,=\, Y^\mu\,, \tag{6.18}$$

we consider the integral

$$\ell = \int_{C\ \sigma=0}^{\sigma=1} \mathrm{d}s\,, \tag{6.19}$$

with either

$$\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu\,, \tag{6.20}$$

when the curve is spacelike, or

$$\mathrm{d}s^2 = -g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu\,, \tag{6.21}$$

wherever the curve is timelike. For simplicity we choose the curve to be spacelike, Eq. (6.20). The timelike case goes exactly analogously.

Consider now an infinitesimal *displacement* of the curve, keeping however X and Y in their places:

$$\begin{aligned} x'^\mu(\sigma) &= x^\mu(\sigma) + \eta^\mu(\sigma)\,, \quad \eta \text{ infinitesimal,} \\ \eta^\mu(0) &= \eta^\mu(1) = 0\,, \end{aligned} \tag{6.22}$$

then what is the infinitesimal change in ℓ ?

$$\begin{aligned} \delta\ell &= \int \delta \mathrm{d}s\,; \\ 2\mathrm{d}s\delta\mathrm{d}s &= (\delta g_{\mu\nu})\mathrm{d}x^\mu\mathrm{d}x^\nu + 2g_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}\eta^\nu + \mathcal{O}(\mathrm{d}\eta^2) \\ &= (\partial_\alpha g_{\mu\nu})\eta^\alpha\mathrm{d}x^\mu\mathrm{d}x^\nu + 2g_{\mu\nu}\mathrm{d}x^\mu\frac{\mathrm{d}\eta^\nu}{\mathrm{d}\sigma}\mathrm{d}\sigma\,. \end{aligned} \tag{6.23}$$

Now we make a restriction for the *original* curve:

$$\frac{\mathrm{d}s}{\mathrm{d}\sigma} = 1\,, \tag{6.24}$$

which one can always realize by choosing an appropriate parametrization of the curve. (6.23) then reads

$$\delta\ell = \int \mathrm{d}\sigma\Big(\tfrac{1}{2}\eta^\alpha g_{\mu\nu,\alpha}\frac{\mathrm{d}x^\mu}{\mathrm{d}\sigma}\frac{\mathrm{d}x^\nu}{\mathrm{d}\sigma} + g_{\mu\alpha}\frac{\mathrm{d}x^\mu}{\mathrm{d}\sigma}\frac{\mathrm{d}\eta^\alpha}{\mathrm{d}\sigma}\Big)\,. \tag{6.25}$$

We can take care of the $\mathrm{d}\eta/\mathrm{d}\sigma$ term by partial integration; using

$$\frac{\mathrm{d}}{\mathrm{d}\sigma}g_{\mu\alpha} = g_{\mu\alpha,\lambda}\frac{\mathrm{d}x^\lambda}{\mathrm{d}\sigma}\,, \tag{6.26}$$

we get

$$\begin{aligned} \delta\ell &= \int \mathrm{d}\sigma\Big(\eta^\alpha\Big(\tfrac{1}{2}g_{\mu\nu,\alpha}\frac{\mathrm{d}x^\mu}{\mathrm{d}\sigma}\frac{\mathrm{d}x^\nu}{\mathrm{d}\sigma} - g_{\mu\alpha,\lambda}\frac{\mathrm{d}x^\lambda}{\mathrm{d}\sigma}\frac{\mathrm{d}x^\mu}{\mathrm{d}\sigma} - g_{\mu\alpha}\frac{\mathrm{d}^2x^\mu}{\mathrm{d}\sigma^2}\Big) + \frac{\mathrm{d}}{\mathrm{d}\sigma}\Big(g_{\mu\alpha}\frac{\mathrm{d}x^\mu}{\mathrm{d}\sigma}\eta^\alpha\Big)\Big)\,. \\ &= -\int \mathrm{d}\sigma\,\eta^\alpha(\sigma)g_{\mu\alpha}\Big(\frac{\mathrm{d}^2x^\mu}{\mathrm{d}\sigma^2} + \Gamma^\mu_{\kappa\lambda}\frac{\mathrm{d}x^\kappa}{\mathrm{d}\sigma}\frac{\mathrm{d}x^\lambda}{\mathrm{d}\sigma}\Big)\,. \end{aligned} \tag{6.27}$$

The pure derivative term vanishes since we require η to vanish at the end points, Eq. (6.22). We used symmetry under interchange of the indices λ and μ in the first line and the definitions (6.10) and (6.11) for Γ. Now, strictly following standard procedure in mathematical physics, we can demand that $\delta\ell$ vanishes for all choices of the infinitesimal function $\eta^\alpha(\sigma)$ obeying the boundary condition. We obtain exactly the equation for geodesics, (5.8). If we hadn't imposed Eq. (6.24) we would have obtained Eq. (5.8a).

We have *spacelike* geodesics (with Eq. (6.20) and *timelike* geodesics (with Eq. (6.21). One can show that for timelike geodesics ℓ is a relative *maximum*. For spacelike geodesics it is on a saddle point. Only in spaces with a positive definite $g_{\mu\nu}$ the length ℓ of the path is a minimum for the geodesic.

Curvature

As for the Riemann curvature tensor defined in the previous chapter, we can now raise and lower all its indices:

$$R_{\mu\nu\alpha\beta} = g_{\mu\lambda}R^\lambda{}_{\nu\alpha\beta}\,, \tag{6.28}$$

and we can check if there are any further symmetries, apart from (5.26), (5.29) and (5.30). By writing down the full expressions for the curvature in terms of $g_{\mu\nu}$ one finds

$$R_{\mu\nu\alpha\beta} = -R_{\nu\mu\alpha\beta} = R_{\alpha\beta\mu\nu}\,. \tag{6.29}$$

By contracting two indices one obtains the *Ricci tensor:*

$$R_{\mu\nu} = R^\lambda{}_{\mu\lambda\nu}\,, \tag{6.30}$$

It now obeys

$$R_{\mu\nu} = R_{\nu\mu}\,, \tag{6.31}$$

We can contract further to obtain the *Ricci scalar,*

$$R = g^{\mu\nu}R_{\mu\nu} = R^\mu_\mu\,. \tag{6.32}$$

Now that we have the metric tensor $g_{\mu\nu}$, we may use a generalized version of the summation convention: *If there is a repeated subscript index, it means that one of them must be raised using the metric tensor $g^{\mu\nu}$, after which we sum over the values.* Similarly, repeated superscript indices can now be summed over:

$$A_m\,B_\mu \equiv A_\mu\,B^\mu \equiv A^\mu\,B_\mu \equiv A_\mu\,B_\nu\,g_{\mu\nu}\,. \tag{6.33}$$

The Bianchi identity (5.30) implies for the Ricci tensor:

$$D_\mu R_{\mu\nu} - \tfrac{1}{2}D_\nu R = 0\,. \tag{6.34}$$

We define the *Einstein tensor* $G_{\mu\nu}(x)$ as

$$G_{\mu\nu} = R_{\mu\nu} - \tfrac{1}{2} R g_{\mu\nu} \,, \quad D_\mu G_{\mu\nu} = 0 \,. \tag{6.35}$$

The formalism developed in this chapter can be used to describe any kind of curved space or space-time. Every choice for the metric $g_{\mu\nu}$ (under certain constraints concerning its eigenvalues) can be considered. We obtain the trajectories – geodesics – of particles moving in gravitational fields. However so-far we have not discussed the equations that determine the gravity field configurations given some configuration of stars and planets in space and time. This will be done in the next chapters.

7. The perturbative expansion and Einstein's law of gravity.

We have a law of gravity if we have some prescription to pin down the values of the curvature tensor $R^\mu{}_{\alpha\beta\gamma}$ near a given matter distribution in space and time. To obtain such a prescription we want to make use of the given fact that Newton's law of gravity holds whenever the non-relativistic approximation is justified. This will be the case in any region of space and time that is sufficiently small so that a coordinate frame can be devised there that is approximately flat. The gravitational fields are then sufficiently weak and then at that spot we not only know fairly well how to describe the laws of matter, but we also know how these *weak* gravitational fields are determined by the matter distribution there. In our small region of space-time we write

$$g_{\mu\nu}(x) = \eta_{\mu\nu} + h_{\mu\nu} \,, \tag{7.1}$$

where

$$\eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} , \tag{7.2}$$

and $h_{\mu\nu}$ is a small perturbation. We find (see (6.10):

$$\Gamma_{\lambda\mu\nu} = \tfrac{1}{2}(\partial_\mu h_{\lambda\nu} + \partial_\nu h_{\lambda\mu} - \partial_\lambda h_{\mu\nu}) \,; \tag{7.3}$$

$$g^{\mu\nu} = \eta_{\mu\nu} - h^{\mu\nu} + h^\mu{}_\alpha h^{\alpha\nu} - \dots \,. \tag{7.4}$$

In this latter expression the indices were raised and lowered using $\eta^{\mu\nu}$ and $\eta_{\mu\nu}$ instead of the $g^{\mu\nu}$ and $g_{\mu\nu}$. This is a revised index- and summation convention that we only apply on expressions containing $h_{\mu\nu}$. Note that the indices in $\eta_{\mu\nu}$ need not be raised or lowered.

$$\Gamma^\alpha_{\mu\nu} = \eta^{\alpha\lambda}\Gamma_{\lambda\mu\nu} + \mathcal{O}(h^2) \,. \tag{7.5}$$

The curvature tensor is

$$R^\alpha{}_{\beta\gamma\delta} = \partial_\gamma \Gamma^\alpha_{\beta\delta} - \partial_\delta \Gamma^\alpha_{\beta\gamma} + \mathcal{O}(h^2) \,, \tag{7.6}$$

and the Ricci tensor

$$\begin{aligned} R_{\mu\nu} &= \partial_\alpha \Gamma^\alpha_{\mu\nu} - \partial_\mu \Gamma^\alpha_{\nu\alpha} + \mathcal{O}(h^2) \\ &= \tfrac{1}{2}\big(-\partial^2 h_{\mu\nu} + \partial_\alpha \partial_\mu h^\alpha_\nu + \partial_\alpha \partial_\nu h^\alpha_\mu - \partial_\mu \partial_\nu h^\alpha{}_\alpha \big) + \mathcal{O}(h^2)\ . \end{aligned} \tag{7.7}$$

The Ricci scalar is

$$R = -\partial^2 h_{\mu\mu} + \partial_\mu \partial_\nu h_{\mu\nu} + \mathcal{O}(h^2)\,. \tag{7.8}$$

A slowly moving particle has

$$\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} \approx (1,\,0,\,0,\,0)\,, \tag{7.9}$$

so that the geodesic equation (5.8) becomes

$$\frac{\mathrm{d}^2}{\mathrm{d}\tau^2} x^i(\tau) = -\Gamma^i_{00}\,. \tag{7.10}$$

Apparently, $\Gamma^i = -\Gamma^i_{00}$ is to identified with the gravitational field. Now in a stationary system one may ignore time derivatives ∂_0. Therefore Eq. (7.3) for the gravitational field reduces to

$$\Gamma_i = -\Gamma_{i00} = \tfrac{1}{2}\partial_i h_{00}\,, \tag{7.11}$$

so that one may identify $-\frac{1}{2}h_{00}$ as the gravitational potential. This confirms the suspicion expressed in Chapter 3 that the local clock speed, which is $\varrho = \sqrt{-g_{00}} \approx 1 - \frac{1}{2}h_{00}$, can be identified with the gravitational potential, Eq. (3.18) (apart from an additive constant, of course).

Now let $T_{\mu\nu}$ be the energy-momentum-stress-tensor; $T_{44} = -T_{00}$ is the mass-energy density and since in our coordinate frame the distinction between covariant derivative and ordinary derivatives is negligible, Eq. (1.26) for energy-momentum conservation reads

$$D_\mu T_{\mu\nu} = 0 \tag{7.12}$$

In other coordinate frames this deviates from ordinary energy-momentum conservation just because the gravitational fields can carry away energy and momentum; the $T_{\mu\nu}$ we work with presently will be only the contribution from stars and planets, not their gravitational fields. Now Newton's equations for slowly moving matter imply

$$\begin{aligned} \Gamma^i &= -\Gamma^i_{00} = -\partial_i V(x) = \tfrac{1}{2}\partial_i h_{00}\ ; \\ \partial_i \Gamma_i &= -4\pi G_N T_{44} = 4\pi G_N T_{00}\ ; \\ \vec{\partial}^2 h_{00} &= 8\pi G_N T_{00}\ . \end{aligned} \tag{7.13}$$

This we now wish to rewrite in a way that is invariant under general coordinate transformations. This is a very important step in the theory. Instead of having one component of the $T_{\mu\nu}$ depend on certain partial derivatives of the connection fields Γ

we want a relation between covariant tensors. The energy momentum density for matter, $T_{\mu\nu}$, satisfying Eq. (7.12), is clearly a covariant tensor. The only covariant tensors one can build from the expressions in Eq. (7.13) are the Ricci tensor $R_{\mu\nu}$ and the scalar R. The two independent components that are scalars under spacelike rotations are

$$R_{00} \;=\; -\tfrac{1}{2}\vec{\partial}^2\, h_{00}\ ; \tag{7.14}$$

$$\text{and} \qquad R \;=\; \partial_i\partial_j h_{ij} + \vec{\partial}^2(h_{00} - h_{ii})\ . \tag{7.15}$$

Now these equations strongly suggest a relationship between the tensors $T_{\mu\nu}$ and $R_{\mu\nu}$, but we now have to be careful. Eq. (7.15) cannot be used since it is not a priori clear whether we can neglect the spacelike components of h_{ij} (we cannot). The most general tensor relation one can expect of this type would be

$$R_{\mu\nu} \;=\; AT_{\mu\nu} + Bg_{\mu\nu}T^\alpha_\alpha\,, \tag{7.16}$$

where A and B are constants yet to be determined. Here the trace of the energy momentum tensor is, in the non-relativistic approximation

$$T^\alpha_\alpha \;=\; -T_{00} + T_{ii}\,. \tag{7.17}$$

so the 00 component can be written as

$$R_{00} \;=\; -\tfrac{1}{2}\vec{\partial}^2 h_{00} \;=\; (A+B)T_{00} - BT_{ii}\ , \tag{7.18}$$

to be compared with (7.13). It is of importance to realize that in the Newtonian limit the T_{ii} term (the pressure p) vanishes, not only because the pressure of ordinary (non-relativistic) matter is very small, but also because it averages out to zero as a source: in the stationary case we have

$$0 \;=\; \partial_\mu T_{\mu i} \;=\; \partial_j T_{ji}\,, \tag{7.19}$$

$$\frac{\mathrm{d}}{\mathrm{d}x^1}\int T_{11}\mathrm{d}x^2\mathrm{d}x^3 \;=\; -\int \mathrm{d}x^2\mathrm{d}x^3(\partial_2 T_{21} + \partial_3 T_{31}) \;=\; 0\,, \tag{7.20}$$

and therefore, if our source is surrounded by a vacuum, we must have

$$\int T_{11}\,\mathrm{d}x^2\mathrm{d}x^3 \;=\; 0 \quad\rightarrow\quad \int \mathrm{d}^3\vec{x}\,T_{11} \;=\; 0\ ,$$

$$\text{and similarly,}\qquad \int \mathrm{d}^3\vec{x}\,T_{22} \;=\; \int \mathrm{d}^3\vec{x}\,T_{33} \;=\; 0\ . \tag{7.21}$$

We must conclude that all one can deduce from (7.18) and (7.13) is

$$A+B \;=\; -4\pi G_N\,. \tag{7.22}$$

Fortunately we have another piece of information. The trace of (7.16) is $R = (A+4B)T^\alpha_\alpha$. The quantity $G_{\mu\nu}$ in Eq. (6.35) is then

$$G_{\mu\nu} \;=\; AT_{\mu\nu} - (\tfrac{1}{2}A + B)T^\alpha_\alpha\, g_{\mu\nu}\,, \tag{7.23}$$

and since we have both the Bianchi identity (6.35) and the energy conservation law (7.12) we get (using the modified summation convention, Eq. (6.33))

$$D_\mu G_{\mu\nu} = 0\,; \qquad D_\mu T_{\mu\nu} = 0\,; \qquad \text{therefore} \qquad (\tfrac{1}{2}A + B)\partial_\nu(T^\alpha_\alpha) = 0\,. \tag{7.24}$$

Now T^α_α , the trace of the energy-momentum tensor, is dominated by $-T_{00}$. This will in general not be space-time independent. So our theory would be inconsistent unless

$$B = -\tfrac{1}{2}A\,; \qquad A = -8\pi G_N\,, \tag{7.25}$$

using (7.22). We conclude that the only tensor equation consistent with Newton's equation in a locally flat coordinate frame is

$$R_{\mu\nu} - \tfrac{1}{2}Rg_{\mu\nu} = -8\pi G_N T_{\mu\nu}\,, \tag{7.26}$$

where the sign of the energy-momentum tensor is defined by (ϱ is the energy density)

$$T_{44} = -T_{00} = T^0_0 = \varrho\,. \tag{7.27}$$

This is Einstein's celebrated law of gravitation. From the equivalence principle it follows that if this law holds in a locally flat coordinate frame it should hold in any other frame as well.

Since both left and right of Eq. (7.26) are symmetric under interchange of the indices we have here 10 equations. We know however that both sides obey the conservation law

$$D_\mu G_{\mu\nu} = 0\,. \tag{7.28}$$

These are 4 equations that are automatically satisfied. This leaves 6 non-trivial equations. They should determine the 10 components of the metric tensor $g_{\mu\nu}$, so one expects a remaining freedom of 4 equations. Indeed the coordinate transformations are as yet undetermined, and there are 4 coordinates. Counting degrees of freedom this way suggests that Einstein's gravity equations should indeed determine the space-time metric uniquely (apart from coordinate transformations) and could replace Newton's gravity law. However one has to be extremely careful with arguments of this sort. In the next chapter we show that the equations are associated with an action principle, and this is a much better way to get some feeling for the internal self-consistency of the equations. Fundamental difficulties are not completely resolved, in particular regarding the possible emergence of singularities in the solutions.

Note that (7.26) implies

$$\begin{aligned} 8\pi G_N T^\mu_\mu &= R\,; \\ R_{\mu\nu} &= -8\pi G_N(T_{\mu\nu} - \tfrac{1}{2}T^\alpha_\alpha g_{\mu\nu})\,. \end{aligned} \tag{7.29}$$

therefore in parts of space-time where no matter is present one has

$$R_{\mu\nu} = 0\,, \tag{7.30}$$

but the complete Riemann tensor $R^\alpha{}_{\beta\gamma\delta}$ will not vanish.

The *Weyl tensor* is defined by subtracting from $R_{\alpha\beta\gamma\delta}$ a part in such a way that all contractions of any pair of indices gives zero:

$$C_{\alpha\beta\gamma\delta} = R_{\alpha\beta\gamma\delta} + \tfrac{1}{2}\Big[g_{\alpha\delta}R_{\gamma\beta} + g_{\beta\gamma}R_{\alpha\delta} + \tfrac{1}{3}R\, g_{\alpha\gamma}\, g_{\beta\delta} - (\gamma \Leftrightarrow \delta)\Big]. \tag{7.31}$$

This construction is such that $C_{\alpha\beta\gamma\delta}$ has the same symmetry properties (5.26), (5.29) and (6.29) and furthermore

$$C^\mu_{\ \beta\mu\gamma} = 0\,. \tag{7.32}$$

If one carefully counts the number of independent components one finds in a given point x that $R_{\alpha\beta\gamma\delta}$ has 20 degrees of freedom, and $R_{\mu\nu}$ and $C_{\alpha\beta\gamma\delta}$ each 10.

The cosmological constant

We have seen that Eq. (7.26) can be derived uniquely; there is no room for correction terms if we insist that both the equivalence principle and the Newtonian limit are valid. But if we allow for a small deviation from Newton's law then another term can be imagined. Apart from (7.28) we also have

$$D_\mu\, g_{\mu\nu} = 0\,, \tag{7.33}$$

and therefore one might replace (7.26) by

$$R_{\mu\nu} - \tfrac{1}{2}R\, g_{\mu\nu} + \Lambda\, g_{\mu\nu} = -8\pi G_N\, T_{\mu\nu}\,, \tag{7.34}$$

where Λ is a constant of Nature, with a very small numerical value, called the *cosmological constant*. The extra term may also be regarded as a 'renormalization':

$$\delta T_{\mu\nu} \propto g_{\mu\nu}\,, \tag{7.35}$$

implying some residual energy and pressure in the vacuum. Einstein first introduced such a term in order to obtain interesting solutions, but later "regretted this". In any case, a residual gravitational field emanating from the vacuum, if it exists at all, must be extraordinarily weak. For a long time, it was presumed that the cosmological constant $\Lambda = 0$. Only very recently, strong indications were reported for a tiny, positive value of Λ. Whether or not the term exists, it is very mysterious why Λ should be so close to zero. In modern field theories it is difficult to understand why the energy and momentum density of the vacuum state (which just happens to be the state with lowest energy content) are tuned to zero. So we do not know why $\Lambda = 0$, exactly or approximately, with or without Einstein's regrets.

8. The action principle.

We saw that a particle's trajectory in a space-time with a gravitational field is determined by the geodesic equation (5.8), but also by postulating that the quantity

$$\ell = \int \mathrm{d}s\,, \quad \text{with} \quad (\mathrm{d}s)^2 = -g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu\,, \tag{8.1}$$

is stationary under infinitesimal displacements $x^\mu(\tau) \to x^\mu(\tau) + \delta x^\mu(\tau)$:

$$\delta\ell = 0\,. \tag{8.2}$$

This is an example of an action principle, ℓ being the action for the particle's motion in its orbit. The advantage of this action principle is its simplicity as well as the fact that the expressions are manifestly covariant so that we see immediately that they will give the same results in any coordinate frame. Furthermore the *existence* of solutions of (8.2) is very plausible in particular if the expression for this action is bounded. For example, for most timelike geodesics ℓ is an absolute maximum.

Now let

$$g \stackrel{\text{def}}{=} \det(g_{\mu\nu})\,. \tag{8.3}$$

Then consider in some volume V of 4 dimensional space-time the so-called *Einstein-Hilbert action:*

$$I = \int_V \sqrt{-g}\, R \mathrm{d}^4 x\,, \tag{8.4}$$

where R is the Ricci scalar (6.32). We saw in chapters 4 and 6 that with this factor $\sqrt{-g}$ the integral (8.4) is invariant under coordinate transformations, but if we keep V finite then of course the boundary should be kept unaffected. Consider now an *infinitesimal* variation of the metric tensor $g_{\mu\nu}$:

$$\tilde g_{\mu\nu} = g_{\mu\nu} + \delta g_{\mu\nu}\,, \tag{8.5}$$

so that its inverse, $g^{\mu\nu}$ changes as

$$\tilde g^{\mu\nu} = g^{\mu\nu} - \delta g^{\mu\nu}\,. \tag{8.6}$$

We impose that $\delta g_{\mu\nu}$ and its first derivatives *vanish* on the boundary of V. What effect does this have on the Ricci tensor $R_{\mu\nu}$ and the Ricci scalar R?

First, compute to lowest order in $\delta g_{\mu\nu}$ the variation $\delta\Gamma^\lambda_{\mu\nu}$ of the connection field

$$\tilde\Gamma^\lambda_{\mu\nu} = \Gamma^\lambda_{\mu\nu} + \delta\Gamma^\lambda_{\mu\nu}\,.$$

Using this, and Eqs. (6.8), (6.10) and (6.11), we find :

$$\delta\Gamma^\lambda_{\mu\nu} = \tfrac{1}{2} g^{\lambda\alpha}(\partial_\mu \delta g_{\alpha\nu} + \partial_\nu \delta g_{\alpha\mu} - \partial_\alpha \delta g_{\mu\nu}) - \delta g^{\alpha\lambda}\Gamma_{\alpha\mu\nu}\,.$$

Now, we make an important observation. Since $\delta\Gamma^\lambda_{\mu\nu}$ is the *difference* between two connection fields, it transforms as a true tensor. Therefore, this last expression can be written in such a way that we see only covariant derivatives:

$$\delta\Gamma^\lambda_{\mu\nu} = \tfrac{1}{2} g^{\lambda\alpha}(D_\mu \delta g_{\alpha\nu} + D_\nu \delta g_{\alpha\mu} - D_\alpha \delta g_{\mu\nu})\,.$$

This, of course, we can check explicitly. Similarly, again using the fact that these expressions must transform as true tensors, we derive (see Eq. (5.27):

$$\tilde{R}^{\nu}_{\kappa\lambda\alpha} = R^{\nu}_{\kappa\lambda\alpha} + D_\lambda\delta\Gamma^{\nu}_{\kappa\alpha} - D_\alpha\delta\Gamma^{\nu}_{\kappa\lambda} \ ,$$

so that the variation in the Ricci tensor $R_{\mu\nu}$ to lowest order in $\delta g_{\mu\nu}$ is given by

$$\tilde{R}_{\mu\nu} \ = \ R_{\mu\nu} + \tfrac{1}{2}\Big(- D^2\delta g_{\mu\nu} + D_\alpha D_\mu \delta g^\alpha_\nu + D_\alpha D_\nu \delta g^\alpha_\mu - D_\mu D_\nu \delta g^\alpha_\alpha \Big) \,, \tag{8.7}$$

Exercise: check the derivation of Eq. (8.7).

With $\tilde{R} = \tilde{g}^{\mu\nu}\tilde{R}_{\mu\nu}$ we have

$$\tilde{R} \ = \ R - R_{\mu\nu}\delta g^{\mu\nu} + (D_\mu D_\nu \delta g^{\mu\nu} - D^2\delta g^\alpha_\alpha) \,. \tag{8.8}$$

Finally, the determinant of $\tilde{g}_{\mu\nu}$ is obtained by

$$\det(\tilde{g}_{\mu\nu}) = \det\left(g_{\mu\lambda}(\delta^\lambda_\nu + g^{\lambda\alpha}\delta g_{\alpha\nu})\right) \ = \ \det(g_{\mu\nu})\det(\delta^\mu_\nu + g^{\mu\alpha}\delta g_{\alpha\nu}) = g(1+\delta g^\mu_\mu) \ ; \tag{8.9}$$

$$\sqrt{-\tilde{g}} \ = \ \sqrt{-g}\,(1 + \tfrac{1}{2}\delta g^\mu_\mu) \ . \tag{8.10}$$

and so we find for the variation of the integral I as a consequence of the variation (8.5):

$$\tilde{I} \ = \ I + \int_V \sqrt{-g}\,(- R^{\mu\nu} + \tfrac{1}{2}R\, g^{\mu\nu})\delta g_{\mu\nu} + \int_V \sqrt{-g}\,(D_\mu D_\nu - g_{\mu\nu}D^2)\delta g^{\mu\nu} \,. \tag{8.11}$$

However,

$$\sqrt{-g}\, D_\mu X^\mu \ = \ \partial_\mu(\sqrt{-g}\, X^\mu) \,, \tag{8.12}$$

and therefore the second half in (8.11) is an integral over a pure derivative and since we demanded that $\delta g_{\mu\nu}$ (and its derivatives) vanish at the boundary the second half of Eq. (8.11) vanishes. So we find

$$\delta I \ = \ - \int_V \sqrt{-g}\, G^{\mu\nu}\delta g_{\mu\nu} \,, \tag{8.13}$$

with $G_{\mu\nu}$ as defined in (6.35). Note that in these derivations we mixed superscript and subscript indices. Only in (8.12) it is essential that X^μ is a contra-vector since we insist in having an ordinary rather than a covariant derivative in order to be able to do partial integration. Here we see that partial integration using covariant derivatives works out fine provided we have the factor $\sqrt{-g}$ inside the integral as indicated.

We read off from Eq. (8.13) that Einstein's equations for the vacuum, $G_{\mu\nu} = 0$, are equivalent with demanding that

$$\delta I \ = \ 0 \,, \tag{8.14}$$

for all smooth variations $\delta g_{\mu\nu}(x)$. In the previous chapter a connection was suggested between the gauge freedom in choosing the coordinates on the one hand and the conservation law (Bianchi identity) for $G_{\mu\nu}$ on the other. We can now expatiate on this.

For any system, even if it does not obey Einstein's equations, I will be invariant under infinitesimal coordinate transformations:

$$\begin{aligned}
\tilde{x}^\mu &= x^\mu + u^\mu(x)\ , \\
\tilde{g}_{\mu\nu}(x) &= \frac{\partial\tilde{x}^\alpha}{\partial x^\mu}\frac{\partial\tilde{x}^\beta}{\partial x^\nu}\,g_{\alpha\beta}(\tilde{x})\ ; \\
g_{\alpha\beta}(\tilde{x}) &= g_{\alpha\beta}(x) + u^\lambda\partial_\lambda g_{\alpha\beta}(x) + \mathcal{O}(u^2)\ ; \\
\frac{\partial\tilde{x}^\alpha}{\partial x^\mu} &= \delta^\alpha_\mu + u^\alpha{}_{,\mu} + \mathcal{O}(u^2)\ ,
\end{aligned} \tag{8.15}$$

so that

$$\tilde{g}_{\mu\nu}(x) = g_{\mu\nu} + u^\alpha\partial_\alpha g_{\mu\nu} + g_{\alpha\nu}u^\alpha{}_{,\mu} + g_{\mu\alpha}u^\alpha{}_{,\nu} + \mathcal{O}(u^2)\,. \tag{8.16}$$

This combination precisely produces the *covariant* derivatives of u^α. Again the reason is that all other tensors in the equation are true tensors so that non-covariant derivatives are outlawed. And so we find that the variation in $g_{\mu\nu}$ is

$$\tilde{g}_{\mu\nu} = g_{\mu\nu} + D_\mu u_\nu + D_\nu u_\mu\,. \tag{8.17}$$

This leaves I always invariant:

$$\delta I = -2\int\sqrt{-g}\,G^{\mu\nu}D_\mu u_\nu = 0\,; \tag{8.18}$$

for any $u_\nu(x)$. By partial integration one finds that the equation

$$\sqrt{-g}\,u_\nu D_\mu G^{\mu\nu} = 0 \tag{8.19}$$

is automatically obeyed for all $u_\nu(x)$. This is why the Bianchi identity $D_\mu G_{\mu\nu} = 0$, Eq. (6.35) is always automatically obeyed.

The action principle can be expanded for the case that matter is present. Take for instance scalar fields $\phi(x)$. In ordinary flat space-time these obey the Klein-Gordon equation:

$$(\partial^2 - m^2)\phi = 0\,. \tag{8.20}$$

In a gravitational field this will have to be replaced by the covariant expression

$$(D^2 - m^2)\phi = (g^{\mu\nu}D_\mu D_\nu - m^2)\phi = 0\,. \tag{8.21}$$

It is not difficult to verify that this equation also follows by demanding that

$$\begin{aligned}
\delta J &= 0\ ; \\
J = \tfrac{1}{2}\int\sqrt{-g}\,\mathrm{d}^4x\,\phi(D^2 - m^2)\phi &= \int\sqrt{-g}\,\mathrm{d}^4x\Big(-\tfrac{1}{2}(D_\mu\phi)^2 - \tfrac{1}{2}m^2\phi^2\Big)\ ,
\end{aligned} \tag{8.22}$$

for all infinitesimal variations $\delta\phi$ in ϕ (Note that (8.21) follows from (8.22) via partial integrations which are allowed for covariant derivatives in the presence of the $\sqrt{-g}$ term).

Now consider the sum

$$S \;=\; \frac{1}{16\pi G_N} I + J \;=\; \int_V \sqrt{-g}\,\mathrm{d}^4x\Big(\frac{R}{16\pi G_N} - \tfrac{1}{2}(D_\mu\phi)^2 - \tfrac{1}{2}m^2\phi^2\Big)\,, \tag{8.23}$$

and remember that

$$(D_\mu\phi)^2 \;=\; g^{\mu\nu}\partial_\mu\phi\,\partial_\nu\phi\,. \tag{8.24}$$

Then variation in ϕ will yield the Klein-Gordon equation (8.21) for ϕ as usual. Variation in $g_{\mu\nu}$ now gives

$$\delta S \;=\; \int_V \sqrt{-g}\,\mathrm{d}^4x\Big(- \frac{G^{\mu\nu}}{16\pi G_N} + \tfrac{1}{2}D^\mu\phi D^\nu\phi - \tfrac{1}{4}((D_\alpha\phi)^2 + m^2\phi^2)g^{\mu\nu}\Big)\delta g_{\mu\nu}\,. \tag{8.25}$$

So we have

$$G^{\mu\nu} \;=\; -8\pi G_N T^{\mu\nu}\,, \tag{8.26}$$

if we write

$$T_{\mu\nu} \;=\; -D_\mu\phi D_\nu\phi + \tfrac{1}{2}\Big((D_\alpha\phi)^2 + m^2\phi^2\Big)g_{\mu\nu}\,. \tag{8.27}$$

Now since J is invariant under coordinate transformations, Eqs. (8.15), it must obey a continuity equation just as (8.18), (8.19):

$$D_\mu T_{\mu\nu} \;=\; 0\,. \tag{8.28}$$

This equation holds only if the matter field(s) $\phi(x)$ obey the matter field equations. That is because we should add to Eqs. (8.15) the transformation rule for these fields:

$$\tilde\phi(x) = \phi(x) + u^\lambda\partial_\lambda\phi(x) + \mathcal{O}(u^2)\ .$$

Precisely if the fields obey the field equations, the action is stationary under such variations of these fields, so that we could omit this contribution and use an equation similar to (8.18) to derive (8.28). It is important to observe that, by varying the action with respect to the metric tensor $g_{\mu\nu}$, as is done in Eq. (8.25), we can always find a symmetric tensor $T_{\mu\nu}(x)$ that obeys a conservation law (8.28) as soon as the field equations are obeyed.

Since we also have

$$T_{44} \;=\; \tfrac{1}{2}(\vec D\phi)^2 + \tfrac{1}{2}m^2\phi^2 + \tfrac{1}{2}(D_0\phi)^2 \;=\; \mathcal{H}(x)\,, \tag{8.29}$$

which can be identified as the energy density for the field ϕ, the $\{i0\}$ components of (8.28) must represent the energy flow, which is the momentum density, and this implies that this $T_{\mu\nu}$ has to coincide exactly with the ordinary energy-momentum density for the scalar field. In conclusion, demanding (8.25) to vanish also for all infinitesimal variations in $g_{\mu\nu}$ indeed gives us the correct Einstein equation (8.26).

Finally, there is room for a cosmological term in the action:

$$S = \int_V \sqrt{-g}\Big(\frac{R-2\Lambda}{16\pi G_N} - \tfrac{1}{2}(D_\mu\phi)^2 - \tfrac{1}{2}m^2\phi^2\Big)\,. \tag{8.30}$$

This example with the scalar field ϕ can immediately be extended to other kinds of matter such as other fields, fields with further interaction terms (such as $\lambda\phi^4$), and electromagnetism, and even liquids and free point particles. Every time, all we need is the classical action S which we rewrite in a covariant way: $S_{\mathrm{matter}} = \int \sqrt{-g}\,\mathcal{L}_{\mathrm{matter}}$, to which we then add the Einstein-Hilbert action:

$$S = \int_V \sqrt{-g}\Big(\frac{R-2\Lambda}{16\pi G_N} + \mathcal{L}_{\mathrm{matter}}\Big)\,. \tag{8.31}$$

Of course we will often omit the Λ term. Unless stated otherwise the integral symbol will stand short for $\int \mathrm{d}^4x$.

9. Special coordinates.

In the preceding chapters no restrictions were made concerning the choice of coordinate frame. Every choice is equivalent to any other choice (provided the mapping is one-to-one and differentiable). Complete invariance was ensured. However, when one wishes to calculate in detail the properties of some particular solution such as space-time surrounding a point particle or the history of the universe, one is forced to make a choice. Since we have a four-fold freedom for the use of coordinates we can in general formulate four equations and then try to choose our coordinates such a way that these equations are obeyed. Such equations are called "gauge conditions". Of course one should choose the gauge conditions such a way that one can easily see how to obey them, and demonstrate that coordinates obeying these equations exist. We discuss some examples.

1) The *temporal gauge.*

Choose

$$g_{00} = -1\;; \tag{9.1}$$

$$g_{0i} = 0\,, \qquad (i=1,2,3)\,. \tag{9.2}$$

At first sight it seems easy to show that one can always obey these. If in an arbitrary coordinate frame the equations (9.1) and (9.2) are not obeyed, one writes

$$\tilde g_{00} = g_{00} + 2D_0u_0 = -1\;,, \tag{9.3}$$

$$\tilde g_{0i} = g_{0i} + D_iu_0 + D_0u_i = 0\;,. \tag{9.4}$$

$u_0(\vec x,t)$ can be solved from eq. (9.3) by integrating (9.3) in the time direction, after which we can find u_i by integrating (9.4) with respect to time. We then apply Eq. (8.17) to observe that $\tau g_{\mu\nu}(x-u)$ obeys the equations (9.1) and (9.2) up to terms or oder $(u)^2$ (note that Eqs. (9.3) and (9.4) only correspond to coordinate transformations when u is infinitesimal). Iterating the procedure, it seems easy to obey (9.1) and (9.2) with increasing accuracy. Will such an iteration procedure converge? These are coordinates in which there is no gravitational field (only space, not space-time, is curved), hence all lines of the form $\vec{x}(t) =$constant are actually geodesics, as one can easily check (in Eq. (5.8), $\Gamma^i_{00} = 0$). Therefore they are "freely falling" coordinates, but of course freely falling objects in general will go into orbits and hence either wander away from or collide against each other, at which instances these coordinates generate singularities.

2) The gauge:

$$\partial_\mu g_{\mu\nu} = 0\,. \tag{9.5}$$

This gauge has the advantage of being Lorentz invariant. The equations for infinitesimal u_μ become

$$\partial_\mu \tilde{g}_{\mu\nu} = \partial_\mu g_{\mu\nu} + \partial_\mu D_\mu u_\nu + \partial_\mu D_\nu u_\mu = 0\ . \tag{9.6}$$

(Note that ordinary and covariant derivatives must now be distinguished carefully) In an iterative procedure we first solve for $\partial_\nu u_\nu$. Let ∂_ν act on (9.6):

$$2\partial^2 \partial_\nu u_\nu = -\partial_\nu \partial_\mu g_{\mu\nu} + \text{ higher orders,} \tag{9.7}$$

after which

$$\partial^2 u_\nu = -\partial_\mu g_{\mu\nu} - \partial_\nu(\partial_\mu u_\mu) + \text{ higher orders.} \tag{9.8}$$

These are d'Alembert equations of which the solutions are less singular than those of Eqs. (9.3) and (9.4).

A smarter choice is

3) the *harmonic* or *De Donder gauge:*

$$g^{\mu\nu}\Gamma^\lambda_{\mu\nu} = 0\ . \tag{9.9}$$

Coordinates obeying this condition are called harmonic coordinates, for the following reason. Consider a *scalar* field V obeying

$$D^2 V = 0\ , \tag{9.10}$$

$$\text{or} \qquad g^{\mu\nu}\Big(\partial_\mu\partial_\nu V - \Gamma^\lambda_{\mu\nu}\partial_\lambda V\Big) = 0\ . \tag{9.11}$$

Now let us choose four coordinates $x^{1,\dots,4}$ that obey this equation. Note that these then are not covariant equations because the index α of x^α is not participating:

$$g_{\mu\nu}\Big(\partial_\mu\partial_\nu x^\alpha - \Gamma^\lambda_{\mu\nu}\partial_\lambda x^\alpha\Big) = 0\,. \tag{9.12}$$

Now of course, in the gauge (9.9),

$$\partial_\mu \partial_\nu x^\alpha \;=\; 0\;; \qquad \partial_\lambda x^\alpha \;=\; \delta^\alpha_\lambda\,. \tag{9.13}$$

Hence, *in these coordinates,* the equations (9.12) imply (9.9). Eq. (9.10) can be solved quite generally (it helps a lot that the equation is linear!) For

$$g_{\mu\nu} \;=\; \eta_{\mu\nu} + h_{\mu\nu} \tag{9.14}$$

with infinitesimal $h_{\mu\nu}$ this gauge differs slightly from gauge # 2:

$$f_\nu \;=\; \partial_\mu h_{\mu\nu} - \tfrac{1}{2}\partial_\nu h_{\mu\mu} \;=\; 0\,, \tag{9.15}$$

and for infinitesimal u_ν we have

$$\begin{aligned} \tilde{f}_\nu \;&=\; f_\nu + \partial^2 u_\nu + \partial_\mu\partial_\nu u_\mu - \partial_\nu\partial_\mu u_\mu \\ &=\; f_\nu + \partial^2 u_\nu \;=\; 0 \quad \text{(apart from higher orders)} \end{aligned} \tag{9.16}$$

so (of course) we get directly a d'Alembert equation for u_ν. Observe also that the equation (9.10) is the massless Klein-Gordon equation that extremises the action J of Eq. (8.22) when $m = 0$. In this gauge the infinitesimal expression (7.7) for $R_{\mu\nu}$ simplifies into

$$R_{\mu\nu} \;=\; -\tfrac{1}{2}\partial^2 h_{\mu\nu}\,, \tag{9.17}$$

which simplifies practical calculations.

The action principle for Einstein's equations can be extended such that the gauge condition also follows from varying the same action as the one that generates the field equations. This can be done various ways. Suppose the gauge condition is phrased as

$$f_\mu(\{g_{\alpha\beta}\}, x) \;=\; 0\,, \tag{9.18}$$

and that it has been shown that a coordinate choice that obeys (9.18) always exists. Then one adds to the invariant action (8.23), which we now call $S_{\text{inv.}}$:

$$S_{\text{gauge}} \;=\; \int \sqrt{-g}\,\lambda^\mu(x) f_\mu(g,x)\mathrm{d}^4x\;, \tag{9.19}$$

$$S_{\text{total}} \;=\; S_{\text{inv}} + S_{\text{gauge}}\;, \tag{9.20}$$

where $\lambda^\mu(x)$ is a new dynamical variable, called a *Lagrange multiplier.* Variation $\lambda \to \lambda + \delta\lambda$ immediately yields (9.18) as Euler-Lagrange equation. However, we can also consider as a variation the gauge transformation

$$\tilde{g}_{\mu\nu}(x) \;=\; \tilde{x}^\alpha{}_{,\mu}\tilde{x}^\beta{}_{,\nu}\, g_{\alpha\beta}(\tilde{x}(x))\,. \tag{9.21}$$

Then

$$\delta S_{\text{inv}} \;=\; 0\;, \tag{9.22}$$

$$\delta S_{\text{gauge}} \;=\; \int \lambda^\mu \delta f_\mu \overset{?}{=} 0\;. \tag{9.23}$$

Now we must assume that there exists a gauge transformation that produces

$$\delta f_\mu(x) \;=\; \delta^\alpha_\mu\, \delta(x - x^{(1)})\,, \tag{9.24}$$

for any choice of the point $x^{(1)}$ and the index α. This is precisely the assumption that under any circumstance a gauge transformation exists that can tune f_μ to zero. Then the Euler-Lagrange equation tells us that

$$\delta S_{\text{gauge}} = \lambda^\alpha(x^{(1)}) \;\rightarrow\; \lambda^\alpha(x^{(1)}) \;=\; 0\,. \tag{9.25}$$

All other variations of $g_{\mu\nu}$ that are not coordinate transformations then produce the usual equations as described in the previous chapter.

A technical detail: often Eq. (9.24) cannot be realized by gauge transformations that vanish everywhere on the boundary. Therefore we must allow δf_μ also to be non-vanishing on the boundary. if now we impose $\lambda = 0$ on the boundary then this insures (9.25): $\lambda = 0$ everywhere. This means that the equations generated by the action (9.20) may generate solutions with $\lambda \neq 0$ that have to be discarded. There will always be solutions with $\lambda = 0$ everywhere, and these are the solutions we want.

Another way to implement the gauge condition in the Lagrangian is by choosing

$$S_{\text{gauge}} = \int -\tfrac{1}{2}\sqrt{-g}\; g^{\mu\nu} f_\mu f_\nu\,. \tag{9.26}$$

Let us write this as $\int -\frac{1}{2}(\tilde f_\alpha)^2$, where $\tilde f_\alpha$ is defined as $(\sqrt{\sqrt{-g}\; g^{\cdot\cdot}})^{\alpha\mu} f_\mu$. If now we perform an infinitesimal gauge transformation (8.17), and again assume that it can be done such that Eq. (9.24) is realized for $\delta\tilde f_a$, we find

$$\delta S_{\text{total}} = \delta S_{\text{gauge}} = -\tilde f_\alpha(x^{(1)})\,. \tag{9.27}$$

Requiring S_{total} to be stationary then implies $f_\mu(x^{(1)}) = 0$, and all other equations can be seen to be compatible with the ones from S_{inv} alone.

Here, one must impose $f_\mu(x) = 0$ on the boundary, which then will guarantee that $f_\mu = 0$ everywhere in space-time. By choosing to fix the gauge this way, one can often realize that S_{total} has a simpler form than S_{inv}, so that calculations at a later stage simplify, for instance when gravitational radiation is considered (Chapter 15).

10. Electromagnetism.

We write the Lagrangian for the Maxwell equations as[7]

$$\mathcal{L} \;=\; -\tfrac{1}{4} F_{\mu\nu} F_{\mu\nu} + J_\mu A_\mu\,, \tag{10.1}$$

[7]Note that conventions used here differ from others such as Jackson, Classical Electrodynamics by factors such as 4π. The reader may have to adapt the expressions here to his or her own notation. Again the modified summation convention of Eq. (6.33) is implied.

with

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \,; \tag{10.2}$$

This means that for any variation

$$A_\mu \to A_\mu + \delta A_\mu \,, \tag{10.3}$$

the action

$$S = \int \mathcal{L} \mathrm{d}^4 x \,, \tag{10.4}$$

should be stationary when the Maxwell equations are obeyed. We see indeed that, if δA_ν vanishes on the boundary,

$$\begin{aligned} \delta S &= \int \Big(-F_{\mu\nu}\,\partial_\mu \delta A_\nu + J_\mu \delta A_\mu \Big) \mathrm{d}^4 x \\ &= \int \mathrm{d}^4 x \, \delta A_\nu (\partial_\mu F_{\mu\nu} + J_\nu) \,, \end{aligned} \tag{10.5}$$

using partial integration. Therefore (in our simplified units)

$$\partial_\mu F_{\mu\nu} = -J_\nu \,. \tag{10.6}$$

Describing now the interactions of the Maxwell field with the gravitational field is easy. We first have to make S covariant:

$$\begin{aligned} S_{\mathrm{Max}} &= \textstyle\int \mathrm{d}^4 x \sqrt{-g} \Big(-\frac{1}{4} g^{\mu\alpha} g^{\nu\beta}\, F_{\mu\nu} F_{\alpha\beta} + g^{\mu\nu} J_\mu A_\nu \Big) \,, && (a) \\ r F_{\mu\nu} &= \partial_\mu A_\nu - \partial_\nu A_\mu \qquad \text{(unchanged)} \,, && (b) \end{aligned} \tag{10.7}$$

and

$$S = \int \sqrt{-g} \Big(\frac{R - 2\Lambda}{16\pi G_N} \Big) + S_{\mathrm{Max}} \,. \tag{10.8}$$

Indices may be raised or lowered with the usual conventions.

The energy-momentum tensor can be read off from (10.8) by varying with respect to $g_{\mu\nu}$ (and multiplying by 2):

$$T_{\mu\nu} = -F_{\mu\alpha} F_\nu{}^\alpha + (\tfrac{1}{4} F_{\alpha\beta} F^{\alpha\beta} - J^\alpha A_\alpha) g_{\mu\nu} \,; \tag{10.9}$$

here J^α (with the superscript index) was kept as an external fixed source. We have, in flat space-time, the energy density

$$\varrho = -T_{00} = \tfrac{1}{2} (\vec{E}^2 + \vec{B}^2) - J^\alpha A_\alpha \,, \tag{10.10}$$

as usual.

We also see that:

1) The interaction of the Maxwell field with gravitation is unique, there is no freedom to add an as yet unknown term.

2) The Maxwell field is a source of gravitational fields via its energy-momentum tensor, as was to be expected.

3) The *homogeneous* equation in Maxwell's laws, which follows from Eq. (10.7b),

$$\partial_\gamma F_{\alpha\beta} + \partial_\alpha F_{\beta\gamma} + \partial_\beta F_{\gamma\alpha} \;=\; 0\,, \tag{10.11}$$

remains unchanged.

4) Varying A_μ, we find that the inhomogeneous equation becomes

$$D_\mu F_{\mu\nu} \;=\; g^{\alpha\beta} D_\alpha F_{\beta\nu} \;=\; -J_\nu\,, \tag{10.12}$$

and hence receives a contribution from the gravitational field $\Gamma^\lambda_{\mu\nu}$ and the potential $g^{\alpha\beta}$.

Exercise: show, both with formal arguments and explicitly, that Eq. (10.11) does not change if we replace the derivatives by covariant derivatives.

Exercise: show that Eq. (10.12) can also be written as

$$\partial_\mu(\sqrt{-g}\,F^{\mu\nu}) \;=\; -\sqrt{-g}\,J^\nu\,, \tag{10.13}$$

and that

$$\partial_\mu(\sqrt{-g}\,J^\mu) \;=\; 0\,. \tag{10.14}$$

Thus $\sqrt{-g}\,J^\mu$ is the real conserved current, and Eq. (10.13) implies that $\sqrt{-g}$ acts as the dielectric constant of the vacuum.

11. The Schwarzschild solution.

Einstein's equation, (7.26), should be exactly valid. Therefore it is interesting to search for exact solutions. The simplest and most important one is empty space surrounding a static star or planet. There, one has

$$T_{\mu\nu} \;=\; 0\,. \tag{11.1}$$

If the planet does not rotate very fast, the effects of this rotation (which do exist!) may be ignored. Then there is spherical symmetry. Take spherical coordinates,

$$(x^0,\, x^1,\, x^2,\, x^3) \;=\; (t,\, r,\, \theta,\, \varphi)\,. \tag{11.2}$$

Spherical symmetry then implies

$$g_{02} \;=\; g_{03} \;=\; g_{12} \;=\; g_{13} \;=\; g_{23} \;=\; 0\,, \tag{11.3}$$

as well as

$$g_{33} = \sin^2\theta\, g_{22}\ , \tag{11.4}$$

and time-reversal symmetry

$$g_{01} \;=\; 0\,. \tag{11.5}$$

The metric tensor is then specified by writing down the length $\mathrm{d}s$ of the infinitesimal line element:

$$\mathrm{d}s^2 \;=\; -A\mathrm{d}t^2 + B\mathrm{d}r^2 + Cr^2\left(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2\right)\ , \tag{11.6}$$

where A, B, and C are positive functions depending only on r. At large distance from the source we expect:

$$r \;\rightarrow\; \infty\ ; \qquad A,\, B,\, C \;\rightarrow\; 1\,. \tag{11.7}$$

Our freedom to choose the coordinates can be used to choose a new r coordinate:

$$\tilde{r} \;=\; \sqrt{C(r)}\,r\ , \qquad \text{so that} \qquad Cr^2 \;=\; \tilde{r}^2\,. \tag{11.8}$$

We then have

$$B\mathrm{d}r^2 \;=\; B\Big(\sqrt{C} + \frac{r}{2\sqrt{C}}\frac{\mathrm{d}C}{\mathrm{d}r}\Big)^{-2}\mathrm{d}\tilde{r}^2 \;\stackrel{\mathrm{def}}{=}\; \tilde{B}\mathrm{d}\tilde{r}^2\,. \tag{11.9}$$

In the new coordinate one has (henceforth omitting the tilde $\tilde{}$):

$$\mathrm{d}s^2 \;=\; -A\mathrm{d}t^2 + B\mathrm{d}r^2 + r^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2)\,, \tag{11.10}$$

where $A,\, B \rightarrow 1$ as $r \rightarrow \infty$. The signature of this metric must be $(-,+,+,+)$, so that

$$A > 0 \qquad \text{and} \qquad B > 0\,. \tag{11.11}$$

Now for general A and B we must find the affine connection Γ they generate. There is a method that saves us space in writing (but does not save us from having to do the calculations), because many of its coefficients will be zero. If we know all geodesics

$$\ddot{x}^\mu + \Gamma^\mu_{\kappa\lambda}\dot{x}^\kappa\dot{x}^\lambda \;=\; 0\,, \tag{11.12}$$

then they uniquely determine all Γ coefficients. The variational principle for a geodesic is

$$0 \;=\; \delta\int\mathrm{d}s \;=\; \delta\int\sqrt{g_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\sigma}\frac{\mathrm{d}x^\nu}{\mathrm{d}\sigma}}\,\mathrm{d}\sigma\,, \tag{11.13}$$

where σ is an arbitrary parametrization of the curve. In chapter 6 we saw that the original curve is chosen to have

$$\sigma \;=\; s\,. \tag{11.14}$$

The square root is then one, and Eq. (6.23) then corresponds to

$$\tfrac{1}{2}\delta\int g_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}s}\frac{\mathrm{d}x^\nu}{\mathrm{d}s}\mathrm{d}s \;=\; 0\,. \tag{11.15}$$

We write

$$-A\dot{t}^2+B\dot{r}^2+r^2\dot{\theta}^2+r^2\sin^2\theta\dot{\varphi}^2 \;\stackrel{\mathrm{def}}{=}\; F(s)\;; \qquad \delta\int F\mathrm{d}s=\,0\,. \tag{11.16}$$

The dot stands for differentiation with respect to s.

(11.16) generates the Lagrange equation

$$\frac{\mathrm{d}}{\mathrm{d}s}\frac{\partial F}{\partial\dot{x}^\mu} \;=\; \frac{\partial F}{\partial x^\mu}\;. \tag{11.17}$$

For $\mu=0$ this is

$$\frac{\mathrm{d}}{\mathrm{d}s}(-2A\dot{t}) \;=\; 0\,, \tag{11.18}$$

or

$$\ddot{t}+\frac{1}{A}\Big(\frac{\partial A}{\partial r}\,\cdot\dot{r}\Big)\dot{t} \;=\; 0\,. \tag{11.19}$$

Comparing (11.12) we see that all $\Gamma^0_{\ \mu\nu}$ vanish except

$$\Gamma^0_{\ 10} \;=\; \Gamma^0_{\ 01} \;=\; A'/2A \tag{11.20}$$

(the accent, $'$, stands for differentiation with respect to r; the 2 comes from symmetrization of the subscript indices 0 and 1. For $\mu=1$ Eq. (11.17) implies

$$\ddot{r}+\frac{B'}{2B}\dot{r}^2+\frac{A'}{2B}\dot{t}^2-\frac{r}{B}\dot{\theta}^2-\frac{r}{B}\sin^2\theta\,\dot{\varphi}^2 \;=\; 0\,, \tag{11.21}$$

so that all $\Gamma^1_{\ \mu\nu}$ are zero except

$$\begin{aligned}\Gamma^1_{\ 00} \;&=\; A'/2B \quad ; \qquad \Gamma^1_{\ 11} \;=\; B'/2B\;;\\ \Gamma^1_{\ 22} \;&=\; -r/B \quad ; \qquad \Gamma^1_{\ 33} \;=\; -(r/B)\sin^2\theta\;.\end{aligned} \tag{11.22}$$

For $\mu=2$ and 3 we find similarly:

$$\begin{aligned}\Gamma^2_{\ 21} \;=\; \Gamma^2_{\ 12} \;&=\; 1/r \quad ; \qquad \Gamma^2_{\ 33} \;=\; -\sin\theta\cos\theta\;;\\ \Gamma^3_{\ 23} \;=\; \Gamma^3_{\ 32} \;&=\; \cot\theta \quad ; \qquad \Gamma^3_{\ 13} \;=\; \Gamma^3_{\ 31} \;=\; 1/r\;.\end{aligned} \tag{11.23}$$

Furthermore we have

$$\sqrt{-g} \;=\; r^2\sin\theta\,\sqrt{AB}\,. \tag{11.24}$$

and from Eq. (5.18)

$$\Gamma^{\mu}_{\ \mu\beta} \;=\; (\partial_\beta\sqrt{-g})/\sqrt{-g} \;=\; \partial_\beta \log\sqrt{-g}\,. \tag{11.25}$$

Therefore

$$\begin{aligned}\Gamma^{\mu}_{\ \mu 1} &= A'/2A + B'/2B + 2/r\ ,\\ \Gamma^{\mu}_{\ \mu 2} &= \cot\theta\ .\end{aligned} \tag{11.26}$$

The equation

$$R_{\mu\nu} \;=\; 0\ , \tag{11.27}$$

now becomes (see (5.27))

$$R_{\mu\nu} \;=\; -(\log\sqrt{-g})_{,\mu,\,\nu} + \Gamma^{\alpha}_{\ \mu\nu,\,\alpha} - \Gamma^{\beta}_{\ \alpha\mu}\Gamma^{\alpha}_{\ \beta\nu} + \Gamma^{\alpha}_{\ \mu\nu}(\log\sqrt{-g})_{,\,\alpha} \;=\; 0\,. \tag{11.28}$$

Explicitly:

$$\begin{aligned}R_{00} &= \Gamma^{1}_{\ 00,1} - 2\Gamma^{1}_{\ 00}\Gamma^{0}_{\ 01} + \Gamma^{1}_{\ 00}(\log\sqrt{-g})_{,1}\\ &= (A'/2B)' - A'^2/2AB + (A'/2B)\Big(\frac{A'}{2A} + \frac{B'}{2B} + \frac{2}{r}\Big)\\ &= \frac{1}{2B}\Big(A'' - \frac{A'B'}{2B} - \frac{A'^2}{2A} + \frac{2A'}{r}\Big) \;=\; 0\ ,\end{aligned} \tag{11.29}$$

and

$$\begin{aligned}R_{11} \;&=\; -(\log\sqrt{-g})_{,1,1} + \Gamma^{1}_{\ 11,1} - \Gamma^{0}_{\ 10}\Gamma^{0}_{\ 10} - \Gamma^{1}_{\ 11}\Gamma^{1}_{\ 11}\\ &\quad -\Gamma^{2}_{\ 21}\Gamma^{2}_{\ 21} - \Gamma^{3}_{\ 31}\Gamma^{3}_{\ 31} + \Gamma^{1}_{\ 11}(\log\sqrt{-g})_{,1} \;=\; 0\ .\end{aligned} \tag{11.30}$$

This produces

$$\frac{1}{2A}\Big(-A'' + \frac{A'B'}{2B} + \frac{A'^2}{2A} + \frac{2AB'}{rB}\Big) \;=\; 0\,. \tag{11.31}$$

Combining (11.29) and (11.31) we obtain

$$\frac{2}{rB}(AB)' \;=\; 0\,. \tag{11.32}$$

Therefore $AB=$ constant. Since at $r\to\infty$ we have A and $B\to 1$ we conclude

$$B \;=\; 1/A\,. \tag{11.33}$$

In the $\theta\theta$ direction one has

$$\begin{aligned}R_{22} \;&=\; (-\log\sqrt{-g})_{,2,2} + \Gamma^{1}_{\ 22,1} - 2\Gamma^{1}_{\ 22}\Gamma^{2}_{\ 21}\\ &\quad -\Gamma^{3}_{\ 23}\Gamma^{3}_{\ 23} + \Gamma^{1}_{\ 22}(\log\sqrt{-g})_{,1} \;=\; 0\ .\end{aligned} \tag{11.34}$$

This becomes

$$R_{22} \;=\; -\frac{\partial}{\partial\theta}\cot\theta - (\frac{r}{B})' + \frac{2}{B} - \cot^2\theta - \frac{r}{B}\Big(\frac{2}{r} + \frac{(AB)'}{2AB}\Big) \;=\; 0\;. \tag{11.35}$$

Using (11.32) one obtains

$$(r/B)' \;=\; 1\,. \tag{11.36}$$

Upon integration,

$$r/B \;=\; r - 2M\;, \tag{11.37}$$

$$A \;=\; 1 - \frac{2M}{r}\;; \qquad B \;=\; \Big(1 - \frac{2M}{r}\Big)^{-1}\;. \tag{11.38}$$

Here $2M$ is an integration constant. We found the solution even though we did not yet use all equations $R_{\mu\nu} = 0$ available to us (and only a linear combination of R_{00} and R_{11} was used). It is not hard to convince oneself that indeed all equations $R_{\mu\nu} = 0$ are satisfied, first by substituting (11.38) in (11.29) or (11.31), and then spherical symmetry with (11.35) will also ensure that $R_{33} = 0$. The reason why the equations are over-determined is the Bianchi identity:

$$D_\mu G_{\mu\nu} \;=\; 0\,. \tag{11.39}$$

It will always be obeyed automatically, and implies that if most components of $G_{\mu\nu}$ have been set equal to zero the remainder will be forced to be zero too.

The solution we found is the Schwarzschild solution (Schwarzschild, 1916):

$$\mathrm{d}s^2 \;=\; -\Big(1 - \frac{2M}{r}\Big)\mathrm{d}t^2 + \frac{\mathrm{d}r^2}{1 - \dfrac{2M}{r}} + r^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2)\,. \tag{11.40}$$

In (11.37) we inserted $2M$ as an arbitrary integration constant. We see that far from the origin,

$$-g_{00} \;=\; 1 - \frac{2M}{r} \;\to\; 1 + 2V(\vec{x})\,. \tag{11.41}$$

So the gravitational potential $V(\vec{x})$ goes to $-M/r$, as near an object with mass m, if

$$M \;=\; G_N\, m \qquad (c = 1)\,. \tag{11.42}$$

Often we will normalize mass units such that $G_N = 1$.

The Schwarzschild solution[8] is singular at $r = 2M$, but this can be seen to be an artifact of our coordinate choice. By studying the geodesics in this region one can discover

[8]In his original paper, using a slightly different notation, Karl Schwarzschild replaced $\sqrt[3]{r^3 - (2M)^3}$ by a new coordinate r that vanishes at the horizon, since he insisted that what he saw as a singularity should be at the origin, claiming that only this way the solution becomes "eindeutig" (unique), so that you can calculate phenomena such as the perihelion movement (see Chapter 12) unambiguously. The substitution had to be of this form as he was using the equation that only holds if $g = 1$. He did not know that one may choose the coordinates freely, nor that the singularity is not a true singularity at all. This was 1916. The fact that he was the first to get the analytic form, justifies the name Schwarzschild solution.

different coordinate frames in terms of which no singularity is seen. We here give the result of such a procedure. Introduce new coordinates ("Kruskal coordinates")

$$(t,r,\theta,\varphi) \rightarrow (x,y,\theta,\varphi)\,, \tag{11.43}$$

defined by

$$\begin{aligned} \left(\tfrac{r}{2M}-1\right) e^{r/2M} &= xy\,, & (a)\\ e^{t/2M} &= x/y\,, & (b) \end{aligned} \tag{11.44}$$

so that

$$\begin{aligned} \frac{\mathrm{d}x}{x}+\frac{\mathrm{d}y}{y} &= \frac{\mathrm{d}r}{2M(1-2M/r)}\,;\\ \frac{\mathrm{d}x}{x}-\frac{\mathrm{d}y}{y} &= \frac{\mathrm{d}t}{2M}\,. \end{aligned} \tag{11.45}$$

The Schwarzschild line element is now given by

$$\begin{aligned} \mathrm{d}s^2 &= 16M^2\Big(1-\frac{2M}{r}\Big)\frac{\mathrm{d}x\mathrm{d}y}{xy}+r^2\mathrm{d}\Omega^2\\ &= \frac{32M^3}{r}\,e^{-r/2M}\mathrm{d}x\mathrm{d}y+r^2\mathrm{d}\Omega^2 \end{aligned} \tag{11.46}$$

with

$$\mathrm{d}\Omega^2 \stackrel{\text{def}}{=} \mathrm{d}\theta^2+\sin^2\theta\,\mathrm{d}\varphi^2\,. \tag{11.47}$$

The singularity at $r=2M$ disappeared. Remark that Eqs. (11.44) possess *two* solutions (x,y) for every r,t. This implies that the completely extended vacuum solution (= solution with no matter present as a source of gravitational fields) consists of *two* universes connected to each other at the center. Apart from a rotation over 45° the relation between Kruskal coordinates x,y and Schwarzschild coordinates r,t close to the point $r=2M$ can be seen to be exactly as the one between the flat space coordinates x^3, x^0 and the *Rindler* coordinates ξ^3, τ as discussed in chapter 3.

The points $r=0$ however remain singular in the Schwarzschild solution. The regular region of the "universe" has the line

$$xy = -1 \tag{11.48}$$

as its boundary. The region $x>0$, $y>0$ will be identified with the "ordinary world" extending far from our source. The second universe, the region of space-time with $x<0$ and $y<0$ has the same metric as the first one. It is connected to the first one by something one could call a "wormhole". The physical significance of this extended region however is very limited, because:

1) "ordinary" stars and planets contain matter ($T_{\mu\nu}\neq 0$) within a certain radius $r>2M$, so that for them the validity of the Schwarzschild solution stops there.

2) Even if further gravitational contraction produces a "black hole" one finds that there will still be imploding matter around ($T_{\mu\nu} \neq 0$) that will cut off the second "universe" completely from the first.

3) even if there were no imploding matter present the second universe could only be reached by moving faster than the local speed of light.

Exercise: Check these statements by drawing an xy diagram and indicating where the two universes are and how matter and space travellers can move about. Show that also *signals* cannot be exchanged between the two universes.

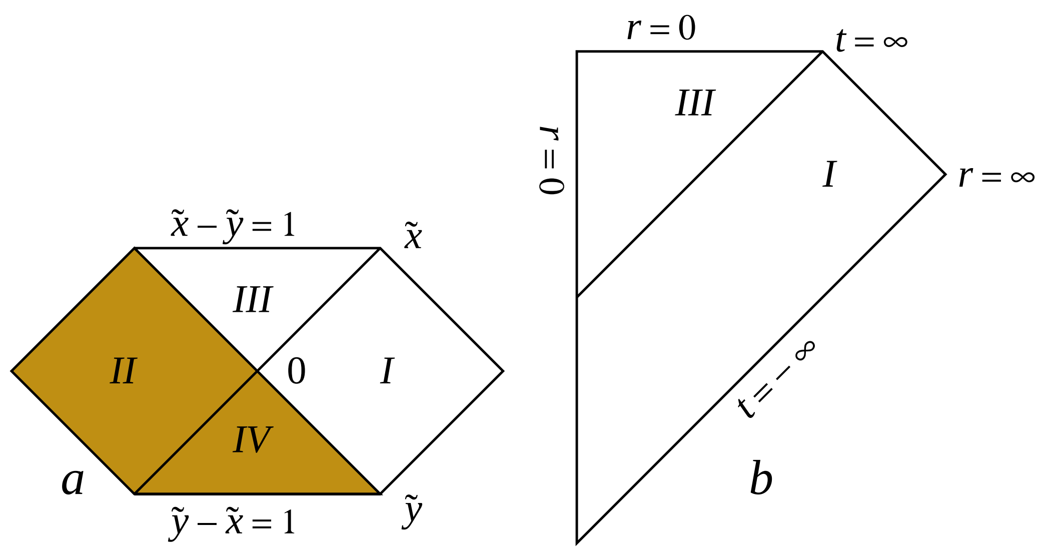

Figure 4: Penrose diagrams. (a) The Penrose diagram for the Schwarzschild metric. The shaded region does not exist in black holes with a collapse in their past; (b) A black hole after collapse. The shaded region is where the collapsing matter is. lightrays moving radially ($\dot\theta = \dot\phi = 0$) here always move at $45°$.

If one draws an "imploding star" in the $x\,y$ diagram one notices that the future horizon may be physically relevant. One then has the so-called black hole solution.

We define the *Penrose coordinates*, $\tilde{x}$ and $\tilde{y}$, by

$$x = \tan(\tfrac{1}{2}\pi\tilde{x}) \ ; \qquad y = \tan(\tfrac{1}{2}\pi\tilde{y}) \ . \tag{11.49}$$

In these coordinates, we see that

i. the lightcone is again at $45°$;

ii. the allowed values for $\tilde{x}$ and $\tilde{y}$ are:

$$|\tilde{x}| < 1 \ , \qquad |\tilde{y}| < 1 \ , \qquad |x - y| < 1 \ . \tag{11.50}$$

This region is sketched in Fig. 4a. We call this a *Penrose diagram*. The shaded part is not accessible if the black hole has a collapsing object in its distant past. Then the appropriate Penrose diagram is the one of Fig. 4b.

12. Mercury and light rays in the Schwarzschild metric.

Historically the orbital motion of the planet Mercury in the Sun's gravitational field has played an important role as a test for the validity of General Relativity (although Einstein would have launched his theory also if such tests had not been available)

To describe this motion we have the variation equation (11.16) for the functions $t(\tau)$, $r(\tau)$, $\theta(\tau)$ and $\varphi(\tau)$, where τ parametrizes the space-time trajectory. Writing $\dot r = \mathrm{d}r/\mathrm{d}\tau$, etc. we have

$$\delta\int\Big\{-\Big(1-\frac{2M}{r}\Big)\dot t^2+\Big(1-\frac{2M}{r}\Big)^{-1}\dot r^2+r^2\Big(\dot\theta^2+\sin^2\theta\,\dot\varphi^2\Big)\Big\}\,\mathrm{d}\tau\;=\;0\,, \tag{12.1}$$

in which we put $\mathrm{d}s^2/\mathrm{d}\tau^2 = -1$ because the trajectory is timelike. The equations of motion follow as Lagrange equations:

$$\frac{\mathrm{d}}{\mathrm{d}\tau}(r^2\dot\theta) \;=\; r^2\sin\theta\cos\theta\dot\varphi^2\;; \tag{12.2}$$

$$\frac{\mathrm{d}}{\mathrm{d}\tau}(r^2\sin^2\theta\,\dot\varphi) \;=\; 0\;; \tag{12.3}$$

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\Big[\Big(1-\frac{2M}{r}\Big)\dot t\Big] \;=\; 0\;. \tag{12.4}$$

We did not yet write the equation for $\ddot r$. Instead of that it is more convenient to divide Eq. (11.40) by $-\mathrm{d}s^2$:

$$1\;=\;\Big(1-\frac{2M}{r}\Big)\dot t^2-\Big(1-\frac{2M}{r}\Big)^{-1}\dot r^2-r^2(\dot\theta^2+\sin^2\theta\,\dot\varphi^2)\;. \tag{12.5}$$

Now even in the completely relativistic metric of the Schwarzschild solution all orbits will be in flat planes through the origin, since spherical symmetry allows us to choose as our initial condition

$$\theta\;=\;\pi/2\;;\qquad \dot\theta\;=\;0\,. \tag{12.6}$$

and then this will remain valid throughout because of Eq. (12.2). Eqs. (12.3) and (12.4) tell us:

$$r^2\dot\varphi\;=\;J\;=\;\text{constant.} \tag{12.7}$$

and

$$\Big(1-\frac{2M}{r}\Big)\dot t\;=\;E\;=\;\text{constant.} \tag{12.8}$$

Eq. (12.5) then becomes

$$1 \;=\; \Big(1-\frac{2M}{r}\Big)^{-1}E^2-\Big(1-\frac{2M}{r}\Big)^{-1}\dot r^2-J^2/r^2\,. \tag{12.9}$$

Just as in the Kepler problem it is convenient to treat r as a function of φ. t has already been eliminated. We now also eliminate s. Let us, for the remainder of this chapter, write differentiation with respect to φ with an accent:

$$r' \;=\; \dot r/\dot\varphi\,. \tag{12.10}$$

From (12.7) and (12.9) one derives:

$$1-2M/r \;=\; E^2-J^2r'^2/r^4-J^2\Big(1-\frac{2M}{r}\Big)/r^2\,. \tag{12.11}$$

Notice that we can interpret E as energy and J as angular momentum. Write, just as in the Kepler problem:

$$r \;=\; 1/u\;, \qquad r' \;=\; -u'/u^2\;; \tag{12.12}$$

$$1-2Mu \;=\; E^2-J^2u'^2-J^2u^2(1-2Mu)\;. \tag{12.13}$$

From this we find

$$\frac{\mathrm{d}u}{\mathrm{d}\varphi} \;=\; \sqrt{(2Mu-1)\Big(u^2+\frac{1}{J^2}\Big)+E^2/J^2}\;. \tag{12.14}$$

The formal solution is

$$\varphi-\varphi_0 \;=\; \int_{u_0}^{u}\mathrm{d}u\Big(\frac{E^2-1}{J^2}+\frac{2Mu}{J^2}-u^2+2Mu^3\Big)^{-\frac12}\;. \tag{12.15}$$

Exercise: show that in the Newtonian limit the u^3 term can be neglected and then compute the integral.

The relativistic perihelion shift will be the extent to which the complete integral from u_{min} to u_{max} (two roots of the third degree polynomial), multiplied by two, differs from 2π.

A neat way to obtain the perihelion shift is by differentiating Eq. (12.13) once more with respect to φ:

$$\frac{2M}{J^2}u'-2u'u''-2uu'+6Mu^2u' \;=\; 0\,. \tag{12.16}$$

Now of course

$$u' \;=\; 0 \tag{12.17}$$

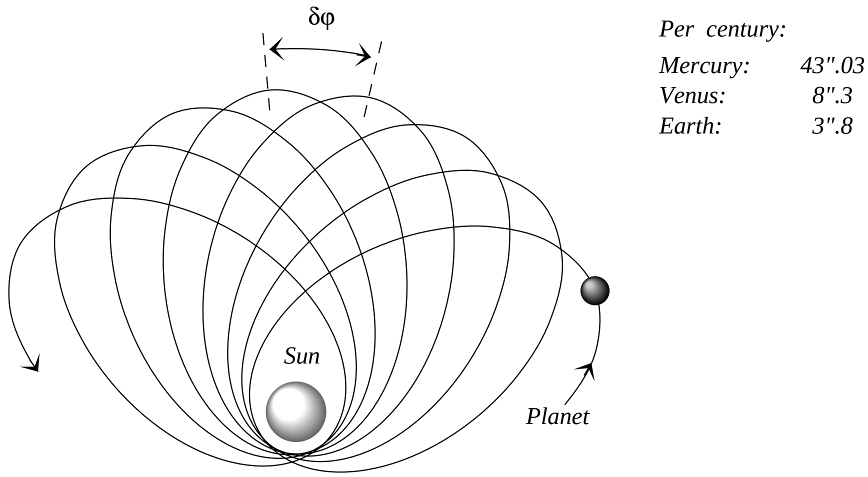

Figure 5: Perihelion shift of a planet in its orbit around a central star.

can be a solution (the circular orbit). If $u' \neq 0$ we divide by u':

$$u'' + u = \frac{M}{J^2} + 3Mu^2 \,. \tag{12.18}$$

The last term is the relativistic correction. Suppose it is small. Then we have a well-known problem in mathematical physics:

$$u'' + u = A + \varepsilon u^2 \,. \tag{12.19}$$

One could expand u as a perturbative expansion in powers of ε, but we wish an expansion that converges for all values of the independent variable φ. Note that Eq. (12.13) allows for every value of u only two possible values for u' so that the solution has to be periodic in φ. The unperturbed period is 2π. But with the u^2 term present we do not know the period exactly. Assume that it can be written as

$$2\pi(1 + \alpha\varepsilon + \mathcal{O}(\varepsilon^2)) \,. \tag{12.20}$$

Write

$$\begin{aligned}
u &= A + B\cos\left[(1-\alpha\varepsilon)\varphi\right] + \varepsilon u_1(\varphi) + \mathcal{O}(\varepsilon^2) \,, && (12.21)\\
u'' &= -B(1-2\alpha\varepsilon)\cos\left[(1-\alpha\varepsilon)\varphi\right] + \varepsilon u_1''(\varphi) + \mathcal{O}(\varepsilon^2) \,; && (12.22)\\
\varepsilon u^2 &= \varepsilon\Big(A^2 + 2AB\cos\left[(1-\alpha\varepsilon)\varphi\right] + B^2\cos^2\left[(1-\alpha\varepsilon)\varphi\right]\Big) + \mathcal{O}(\varepsilon^2) \,. && (12.23)
\end{aligned}$$

We find for u_1:

$$u_1'' + u_1 = (-2\alpha B + 2AB)\cos\varphi + B^2\cos^2\varphi + A^2 \,, \tag{12.24}$$

where now the $\mathcal{O}(\varepsilon)$ terms were omitted since they do not play any further role. This is just the equation for a forced pendulum. If we do not want that the pendulum oscillates with an ever increasing period (u_1 must stay small for all values of φ) then the external force is not allowed to have a Fourier component with the same periodicity as the pendulum itself. Now the term with $\cos\varphi$ in (12.24) is exactly in resonance[9] unless we choose $\alpha = A$. Then one has

$$u_1'' + u_1 \;=\; \tfrac{1}{2}B^2(cos2\varphi + 1) + A^2\ , \tag{12.25}$$

$$u_1 \;=\; \tfrac{1}{2}B^2\Big(1 - \frac{1}{2^2-1}\cos 2\varphi\Big) + A^2\ , \tag{12.26}$$

which is exactly periodic. Apparently one has to choose the period to be $2\pi(1 + A\varepsilon)$ if the orbit is to be periodic in φ. We find that after every passage through the perihelion its position is shifted by

$$\delta\varphi \;=\; 2\pi A\varepsilon \;=\; 2\pi\frac{3M^2}{J^2}\ , \tag{12.27}$$

(plus higher order corrections) in the direction of the planet itself (see Fig. 5).

Now we wish to compute the trajectory of a light ray. It is also a geodesic. Now however $\mathrm{d}s = 0$. In this limit we still have (12.1) – (12.4), but now we set

$$\mathrm{d}s/\mathrm{d}\tau \;=\; 0\ ,$$

so that Eq. (12.5) becomes

$$0 \;=\; \Big(1 - \frac{2M}{r}\Big)\dot t^2 - \Big(1 - \frac{2M}{r}\Big)^{-1}\dot r^2 - r^2(\dot\theta^2 + \sin^2\theta\,\dot\varphi^2)\,. \tag{12.28}$$

Since now the parameter τ is determined up to an arbitrary multiplicative constant, only the ratio J/E will be relevant. Call this j. Then Eq. (12.15) becomes

$$\varphi \;=\; \varphi_0 + \int_{u_0}^{u}\mathrm{d}u(j^{-2} - u^2 + 2Mu^3)^{-\frac{1}{2}}\,. \tag{12.29}$$

As the left hand side of Eq. (12.13) must now be replaced by zero, Eq. (12.18) becomes

$$u'' + u \;=\; 3Mu^2\,. \tag{12.30}$$

An expansion in powers of M is now permitted (because the angle φ is now confined within an interval a little larger than π):

$$u \;=\; A\cos\varphi + v\ , \tag{12.31}$$

$$v'' + v \;=\; 3MA^2\cos^2\varphi \;=\; \frac{3}{2}MA^2(1 + \cos 2\varphi)\ , \tag{12.32}$$

$$v \;=\; \frac{3}{2}MA^2(1 - \tfrac{1}{3}\cos 2\varphi) \;=\; MA^2(2 - \cos^2\varphi)\ . \tag{12.33}$$

[9]Note here and in the following that the solution of an equation of the form $u'' + u = \sum_i A_i\cos\omega_i\varphi$ is $u = \sum_i A_i\cos\omega_i\varphi\,/(1-\omega_i^2) + C_1\cos\varphi + C_2\sin\varphi$. This is singular when $\omega \to 1$.

So we have for small M

$$\frac{1}{r} = u = A\cos\varphi + MA^2(2-\cos^2\varphi)\,. \tag{12.34}$$

The angles φ at which the ray enters and exits are determined by

$$1/r = 0\,, \qquad \cos\varphi = \frac{1\pm\sqrt{1+8M^2A^2}}{2MA}\,. \tag{12.35}$$

Since M is a small expansion parameter and $|\cos\varphi|\le 1$ we must choose the minus sign:

$$\cos\varphi \approx -2MA = -2M/r_0\,, \tag{12.36}$$

$$\varphi \approx \pm\Big(\tfrac{\pi}{2}+2M/r_0\Big)\,, \tag{12.37}$$

where r_0 is the smallest distance of the light ray to the central source. In total the angle of deflection between in- and outgoing ray is in lowest order:

$$\Delta = 4M/r_0\,. \tag{12.38}$$

In conventional units this equation reads

$$\Delta = \frac{4G_N m_\odot}{r_0\,c^2}\,. \tag{12.39}$$

$m_\odot$ is the mass of the central star.

Exercise: show that this is twice what one would expect if a light ray could be regarded as a non-relativistic particle in a hyperbolic orbit around the star.

Exercise: show that expression (12.27) in ordinary units reads as

$$\delta\varphi = \frac{6\pi G_N m_\odot}{a(1-\varepsilon^2)\,c^2}\,, \tag{12.40}$$

where a is the major axis of the orbit, ε its excentricity and c the velocity of light.

13. Generalizations of the Schwarzschild solution.

a). The Reissner-Nordström solution.

Spherical symmetry can still be used as a starting point for the construction of a solution of the combined Einstein-Maxwell equations for the fields surrounding a "planet" with electric charge Q and mass m. Just as Eq. (11.10) we choose

$$\mathrm{d}s^2 = -A\mathrm{d}t^2 + B\mathrm{d}r^2 + r^2(\mathrm{d}\theta^2+\sin^2\theta\,\mathrm{d}\varphi^2)\,, \tag{13.1}$$

but now also a static electric field:

$$E_r = E(r)\,; \qquad E_\theta = E_\varphi = 0\,; \qquad \vec{B} = 0\,. \tag{13.2}$$

This implies that $F_{01} = -F_{10} = E(r)$ and all other components of $F_{\mu\nu}$ are zero. Let us assume that the source J^μ of this field is inside the planet and we are only interested in the solution outside the planet. So there we have

$$J^\mu = 0\,. \tag{13.3}$$

If we move the indices upstairs we get

$$F^{10} = E(r)/AB\,, \tag{13.4}$$

and using

$$\sqrt{-g} = \sqrt{AB}\,r^2\sin\theta\,, \tag{13.5}$$

we find that according to (10.13)

$$\partial_r\Big(\frac{E(r)r^2}{\sqrt{AB}}\Big) = 0\,. \tag{13.6}$$

Thus the inhomogeneous Maxwell law tells us that

$$E(r) = \frac{Q\sqrt{AB}}{4\pi r^2}\,, \tag{13.7}$$

where Q is an integration constant, to be identified with electric charge since at $r\to\infty$ both A and B tend to 1.

The homogeneous Maxwell law (10.11) is automatically obeyed because there is a field A_0 (potential field) with

$$E_r = -\partial_r A_0\,. \tag{13.8}$$

The field (13.7) contributes to $T_{\mu\nu}$:

$$T_{00} = -\,E^2/2B = -AQ^2/32\pi^2r^4\,; \tag{13.9}$$

$$T_{11} = E^2/2A = BQ^2/32\pi^2r^4\,; \tag{13.10}$$

$$T_{22} = -E^2r^2/2AB = -Q^2/32\pi^2r^2\,, \tag{13.11}$$

$$T_{33} = T_{22}\sin^2\theta = -Q^2\sin^2\theta\,/32\pi^2r^2\,. \tag{13.12}$$

We find

$$T^\mu_\mu = g^{\mu\nu}T_{\mu\nu} = 0\,; \qquad R = 0\,, \tag{13.13}$$

a general property of the free Maxwell field. In this case we have ($G_N = 1$)

$$R_{\mu\nu} = -8\pi\,T_{\mu\nu}\,. \tag{13.14}$$

Herewith the equations (11.29) – (11.31) become

$$\begin{aligned} A'' - \frac{A'B'}{2B} - \frac{A'^2}{2A} + \frac{2A'}{r} &= ABQ^2/2\pi r^4 \ , \\ -A'' + \frac{A'B'}{2B} + \frac{A'^2}{2A} + \frac{2AB'}{rB} &= -ABQ^2/2\pi r^4 \ . \end{aligned} \tag{13.15}$$

We find that Eq. (11.32) still holds so that here also

$$B = 1/A\,. \tag{13.16}$$

Eq. (11.36) is now replaced by

$$(r/B)' - 1 = -Q^2/4\pi r^2\,. \tag{13.17}$$

This gives upon integration

$$r/B = r - 2M + Q^2/4\pi r\,. \tag{13.18}$$

So now we have instead of Eq. (11.38),

$$A = 1 - \frac{2M}{r} + \frac{Q^2}{4\pi r^2}\ ; \qquad B = 1/A\,. \tag{13.19}$$

This is the Reissner-Nordstrøm solution (1916, 1918).

If we choose $Q^2/4\pi < M^2$ there are two “horizons”, the roots of the equation $A = 0$:

$$r = r_\pm = M \pm \sqrt{M^2 - Q^2/4\pi}\,. \tag{13.20}$$

Again these singularities are artifacts of our coordinate choice and can be removed by generalizations of the Kruskal coordinates. Now one finds that there would be an infinite sequence of ghost universes connected to ours, if the horizons hadn't been blocked by imploding matter. See Hawking and Ellis for a much more detailed description.

b) The Kerr solution

A fast rotating planet has a gravitational field that is no longer spherically symmetric but only cylindrically. We here only give the solution:

$$\begin{aligned} \mathrm{d}s^2 = & -\mathrm{d}t^2 + (r^2+a^2)\sin^2\theta\mathrm{d}\varphi^2 + \tfrac{2Mr\left(\mathrm{d}t - a\sin^2\theta\mathrm{d}\varphi\right)^2}{r^2+a^2\cos^2\theta} \\ & + (r^2 + a^2\cos^2\theta)\left(\mathrm{d}\theta^2 + \tfrac{\mathrm{d}r^2}{r^2-2Mr+a^2}\right)\ . \end{aligned} \tag{13.21}$$

This solution was found by Kerr in 1963. To prove that this is indeed a solution of Einstein's equations requires patience but is not difficult. For a derivation using more elementary principles more powerful techniques and machinery of mathematical physics are needed. The free parameter a in this solution can be identified with angular momentum.

c) The Newman et al solution

For sake of completeness we also mention that rotating planets can also be electrically charged. The solution for that case was found by Newman et al in 1965. The metric is:

$$\mathrm{d}s^2 = -\frac{\Delta}{Y}(\mathrm{d}t - a\sin^2\theta \mathrm{d}\varphi)^2 + \frac{\sin^2\theta}{Y}(a\mathrm{d}t - (r^2+a^2)\mathrm{d}\varphi)^2 + \frac{Y}{\Delta}\mathrm{d}r^2 + Y\mathrm{d}\theta^2 \,, \quad (13.22)$$

where

$$Y = r^2 + a^2\cos^2\theta \,, \quad (13.23)$$

$$\Delta = r^2 - 2Mr + Q^2/4\pi + a^2 \,. \quad (13.24)$$

The vector potential is

$$A_0 = -\frac{Qr}{4\pi Y} \,; \qquad A_3 = \frac{Qra\sin^2\theta}{4\pi Y} \,. \quad (13.25)$$

Exercise: show that when $Q = 0$ Eqs. (13.21) and (13.22) coincide.

Exercise: find the non-rotating magnetic monopole solution by postulating a radial magnetic field.

Exercise for the advanced student: describe geodesics in the Kerr solution.

14. The Robertson-Walker metric.

General relativity plays an important role in cosmology. The simplest theory is that at a certain moment "$t = 0$", the universe started off from a singularity, after which it began to expand. We assume maximal symmetry by taking as our metric

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + a^2(t)\mathrm{d}\omega^2 \,. \quad (14.1)$$

Here $\mathrm{d}\omega^2$ stands short for some fully isotropic 3-dimensional space, and $a(t)$ describes the (increasing) distance between two neighboring galaxies in space. Although we do embrace here the Copernican principle that all points in space look the same, we abandon the idea that there should be invariance with respect to time translations and also Lorentz invariance for this metric – the galaxies contain clocks that were set to zero at $t = 0$ and each provides for a local inertial frame.

First, we concentrate on the three-dimensional space described by $\mathrm{d}\omega^2$. Here, we take polar coordinates ϱ, θ, φ:

$$\mathrm{d}\omega^2 = B(\varrho)\mathrm{d}\varrho^2 + \varrho^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\varphi^2) \,, \quad (14.2)$$

then in this three dimensional space the Ricci tensor is (by using the same techniques as in chapter 11)

$$R_{11} = B'(\varrho)/\varrho B(\varrho) \ , \tag{14.3}$$

$$R_{22} = 1 - \frac{1}{B} + \frac{\varrho B'}{2B^2} \ . \tag{14.4}$$

In an isotropic (3-dimensional) space, one must have

$$R_{ij} = \lambda g_{ij} \ , \tag{14.5}$$

for some constant λ, and therefore

$$B'/B = \lambda B \varrho \ , \tag{14.6}$$

$$1 - \frac{1}{B} + \frac{\varrho B'}{2B^2} = \lambda \varrho^2 \ . \tag{14.7}$$

Together they give

$$\begin{aligned} 1 - \frac{1}{B} &= \tfrac{1}{2}\lambda\varrho^2 \ ; \\ B &= \frac{1}{1 - \frac{1}{2}\lambda\varrho^2} \ , \end{aligned} \tag{14.8}$$

which indeed also obeys (14.6) and (14.7) separately.

> *Exercise:* show that with $\varrho = \sqrt{\frac{2}{\lambda}}\sin\psi$, this gives the metric of the 3-sphere, in terms of its three angular coordinates ψ, θ, φ.

Often one chooses a new coordinate u:

$$\varrho \stackrel{\text{def}}{=} \frac{\sqrt{2k/\lambda}\, u}{1 + (k/4)u^2} \ . \tag{14.9}$$

One observes that

$$\mathrm{d}\varrho = \sqrt{\frac{2k}{\lambda}} \, \frac{1 - \frac{1}{4}ku^2}{(1 + \frac{1}{4}ku^2)^2}\, \mathrm{d}u \qquad \text{and} \qquad B = \left(\frac{1 + \frac{1}{4}ku^2}{1 - \frac{1}{4}ku^2}\right)^2 , \tag{14.10}$$

so that

$$\mathrm{d}\omega^2 = \frac{2k}{\lambda} \cdot \frac{\mathrm{d}u^2 + u^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2)}{(1 + (k/4)u^2)^2} \ . \tag{14.11}$$

The parameter k is arbitrary except for its sign, which must be the same as the sign of λ. The factor in front of Eq. (14.11) may be absorbed in $a(t)$. Therefore we write for (14.1):

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + a^2(t) \frac{\mathrm{d}\vec{x}^2}{(1 + \frac{1}{4}k\,\vec{x}^2)^2} \ . \tag{14.12}$$

If $k = 1$ the spacelike piece is a sphere, if $k = 0$ it is flat, if $k = -1$ the curvature is negative and space is unbounded (in spite of the fact that then $|\vec{x}|$ *is* bounded, which is an artifact of our coordinate choice).

After some elementary calculations,

$$R_0^0 = \frac{3\ddot{a}}{a} . \tag{14.13}$$

$$R_1^1 = R_2^2 = R_3^3 = \frac{\ddot{a}}{a} + \frac{2}{a^2}(\dot{a}^2 + k) , \tag{14.14}$$

$$R = R_\mu^\mu = \frac{6}{a^2}(a\,\ddot{a} + \dot{a}^2 + k) . \tag{14.15}$$

The tensor $G_{\mu\nu}$ becomes (taking for simplicity $\vec{x} = 0$):

$$G_{00} = \frac{3}{a^2}(\dot{a}^2 + k) = 8\pi G_N\,\varrho + \Lambda , \tag{14.16}$$

$$G_{11} = G_{22} = G_{33} = -2a\,\ddot{a} - \dot{a}^2 - k = a^2(8\pi G_N\,p - \Lambda) . \tag{14.17}$$

Here, $\varrho = T_{44} = T_{00}/g_{00}$ is the energy density and p is the pressure: $T_{ij} = -p\,g_{ij}$.

Now what we have to do is to make certain assumptions about matter in the universe, and its equations of state, *i.e.* the relation between the energy density ϱ and the pressure p. The simplest case is to assume that there is no pressure (a “dust-filled universe”).

In this case, the energy density ϱ is just the matter density, which is inversely proportional to the volume:

$$\varrho = \frac{\varrho_0}{a^3} , \qquad \text{(dust)} \tag{14.18}$$

Then Eq. (14.16), also called *Friedmann's equation*, takes the form:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G_N}{3}\frac{\varrho_0}{a^3} - \frac{k}{a^2} + \frac{\Lambda}{3} . \tag{14.19}$$

We see that, as a increases, first the matter term dominates, then the space-curvature term (with k), and finally the cosmological constant dominates. By differentiating the Friedmann equation, we find that it agrees with Eq. (14.17).

Mathematically, the Friedmann equation can be compared with the equation for a one-dimensional particle (with total energy zero) moving in the potential

$$V(a) = -\frac{4\pi G_N}{3}\frac{\varrho_0}{a} + \frac{k}{2} - \frac{\Lambda}{6}a^2 . \tag{14.20}$$

Starting with small a, we have a rapid expansion. The expansion continues forever if $\Lambda > 0$ and $k = -1$. If $\Lambda < 0$, the expansion always comes to a halt, at which point the universe begins to shrink ($a^{\max}$ in Fig. 6.)

It is instructive to consider the solutions to the equations (14.16) and (14.17) when there are other relations between the pressure and the density. For instance in a radiation-filled universe, we have $p = \varrho/3$, and since we may assume that the radiation is thermal,

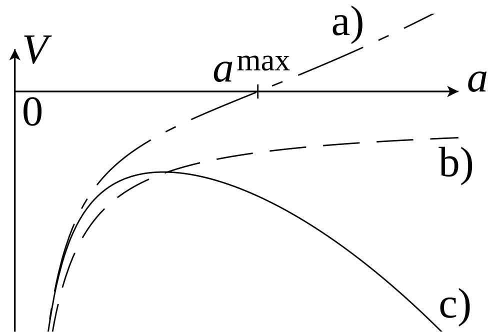

Figure 6: The potential (14.20) for the cases a) $k = 0, \Lambda < 0$, b) $k = -1, \Lambda = 0$ and c) $k = 0, \Lambda > 0$. In the case (a), there is a turning point at $a = a^{\max}$.

and the number of photons is conserved, we may conclude that $\varrho = \varrho_0/a^4$ instead of Eq. (14.18). Indeed, this agrees with Eqs. (14.16) and (14.17).

In the case $\Lambda = 0$, the solutions to the Friedmann Equation (14.19) are well-known mathematical curves. We have

$$a\dot{a}^2 + ka = \frac{8\pi G_N\ \varrho_0}{3} \equiv D; \tag{14.21}$$

$$\dot{a}^2 \ = \ D/a - k\,, \tag{14.22}$$

and from (14.17):

$$\ddot{a} \ = \ -D/2a^2\,. \tag{14.23}$$

Write Eq. (14.22) as

$$\frac{\mathrm{d}t}{\mathrm{d}a} \ = \ \sqrt{\frac{a}{D-ka}}\,, \tag{14.24}$$

then we try

$$a \ = \ \frac{D}{k}\sin^2\varphi\ , \tag{14.25}$$

$$\frac{\mathrm{d}t}{\mathrm{d}\varphi} \ = \ \frac{\mathrm{d}a}{\mathrm{d}\varphi}\frac{\mathrm{d}t}{\mathrm{d}a} \ = \ \frac{2D}{k\sqrt{k}}\sin\varphi\,\cos\varphi\cdot\frac{\sin\varphi}{\cos\varphi}\ , \tag{14.26}$$

$$t(\varphi) \ = \ \frac{D}{k\sqrt{k}}(\varphi - \tfrac{1}{2}\sin 2\varphi)\ , \tag{14.27}$$

$$a(\varphi) \ = \ \frac{D}{2k}(1 - cos2\varphi)\ . \tag{14.28}$$

These are the equations for a cycloid. Since $D > 0$, $t > 0$ and $a > 0$ we demand

$$\begin{aligned} k > 0 \quad &\rightarrow \quad \varphi \text{ real };\\ k < 0 \quad &\rightarrow \quad \varphi \text{ imaginary };\\ k = 0 \quad &\rightarrow \quad \varphi \text{ infinitesimal }. \end{aligned} \tag{14.29}$$

See Fig. 7.

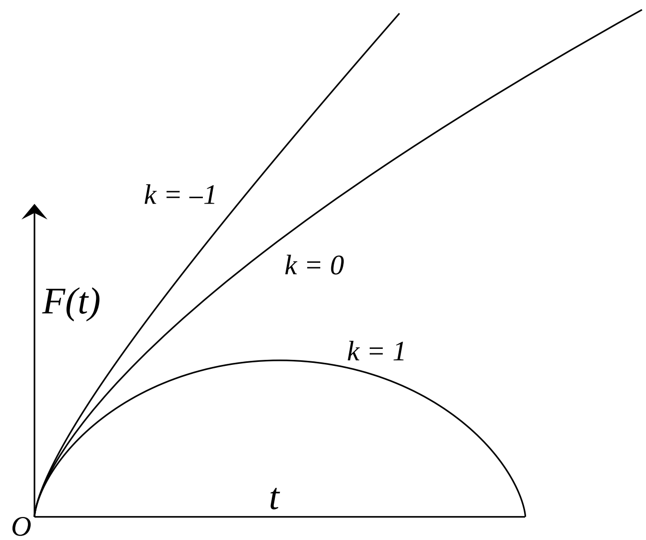

Figure 7: The Robertson-Walker universe with $\Lambda = 0$, for $k = 1$, $k = 0$, and $k = -1$.

All solutions start with a “big bang” at $t = 0$. Only the cycloid in the $k = 1$ case also shows a “big crunch” in the end. If $k \leq 0$ not only space but also time are unbounded.

Other cases, such as $p = -\varrho/3$ and $p = -\varrho$ are good exercises.

15. Gravitational radiation.

Fast moving objects form a time dependent source of the gravitational field, and causality arguments (information in the gravitational fields should not travel faster than light) then suggest that gravitational effects spread like waves in all directions from the source. Far from the source the metric $g_{\mu\nu}$ will stay close to that of flat space-time. To calculate this effect one can adopt a linearized approximation. In contrast to what we did in previous chapters it is now convenient to choose units such that

$$16\pi G_N \;=\; 1\,. \tag{15.1}$$

The linearized Einstein equations were already treated in chapter 7, and in chapter 9 we see that, after gauge fixing, wave equations can be derived (in the absence of matter, Eq. (9.17) can be set to zero). It is instructive to recast these equations in Euler-Lagrange form. The Lagrangian for a linear equation however is itself quadratic. So we have to expand the Einstein-Hilbert action to second order in the perturbations $h_{\mu\nu}$ in the metric:

$$g_{\mu\nu} \;=\; \eta_{\mu\nu} + h_{\mu\nu}\,, \tag{15.2}$$

and after some calculations we find that the terms quadratic in $h_{\mu\nu}$ can be written as:

$$\sqrt{-g}(R + \mathcal{L}^{\mathrm{matter}}) \;=$$

$$\begin{aligned} &\tfrac{1}{8}(\partial_\sigma h_{\alpha\alpha})^2 - \tfrac{1}{4}(\partial_\sigma h_{\alpha\beta})(\partial_\sigma h_{\alpha\beta}) - \tfrac{1}{2}T_{\mu\nu}h_{\mu\nu} \\ &+\tfrac{1}{2}{A_\sigma}^2 + \text{ total derivative } + \text{ higher orders in } h\,, \end{aligned} \tag{15.3}$$

where

$$A_\sigma \;=\; \partial_\mu h_{\mu\sigma} - \tfrac{1}{2}\partial_\sigma h_{\mu\mu}\,, \tag{15.4}$$

and $T_{\mu\nu}$ is the energy momentum tensor of matter when present. Indices are summed over with the flat metric $\eta_{\mu\nu}$, Eq. (7.2).

The Lagrangian is invariant under the linearized gauge transformation (compare (8.16) and (8.17))

$$h_{\mu\nu} \;\rightarrow\; h_{\mu\nu} + \partial_\mu u_\nu + \partial_\nu u_\mu\,, \tag{15.5}$$

which transforms the quantity A_σ into

$$A_\sigma \;\rightarrow\; A_\sigma + \partial^2 u_\sigma\,. \tag{15.6}$$

One possibility to fix the gauge is to choose

$$A_\sigma \;=\; 0 \tag{15.7}$$

(the linearized De Donder gauge). For calculations this is a convenient gauge. But for a better understanding of the real physical degrees of freedom in a radiating gravitational field it is instructive first to look at the "radiation gauge" (which is analogous to the electromagnetic case $\partial_i A_i = 0$):

$$\partial_i h_{ij} \;=\; 0\;; \qquad \partial_i h_{i4} \;=\; 0\,, \tag{15.8}$$

where we stick to the earlier agreement that indices from the middle of the alphabet, $i, j, \dots$, in a summation run from 1 to 3. So we do *not* impose (15.7).

First go to "momentum representation":

$$h(\vec{x},t) \;=\; (2\pi)^{-3/2}\int \mathrm{d}^3\vec{k}\,\hat{h}(\vec{k},t)\,e^{i\vec{k}\cdot\vec{x}}\;; \tag{15.9}$$

$$\partial_i \;\rightarrow\; ik_i\,. \tag{15.10}$$

We will henceforth omit the hat($\hat{}$) since confusion is hardly possible. The advantage of the momentum representation is that the different values of $\vec{k}$ will decouple, so we can concentrate on just one $\vec{k}$ vector, and choose coordinates such that it is in the z direction: $k_1 = k_2 = 0,\ k_3 = k$. We now decide to let indices from the beginning of the alphabet run from 1 to 2. Then one has in the radiation gauge (15.8):

$$h_{3a} \;=\; h_{33} \;=\; h_{30} \;=\; 0\,. \tag{15.11}$$

Furthermore

$$\begin{aligned} A_a &= -\dot{h}_{0a}\;, \\ A_3 &= -\tfrac{1}{2}ik(h_{aa} - h_{00})\;, \\ A_0 &= \tfrac{1}{2}(-\dot{h}_{00} - \dot{h}_{aa})\;. \end{aligned} \tag{15.12}$$

Let us split off the trace of h_{ab} :

$$h_{ab} = \tilde{h}_{ab} + \tfrac{1}{2}\delta_{ab}h \ , \tag{15.13}$$

with

$$h = h_{aa} \ ; \qquad \tilde{h}_{aa} = 0 \, . \tag{15.14}$$

Then we find that

$$\begin{aligned} \mathcal{L} &= \mathcal{L}_1 + \mathcal{L}_2 + \mathcal{L}_3 \ , \\ \mathcal{L}_1 &= \tfrac{1}{4}(\dot{\tilde{h}}_{ab})^2 - \tfrac{1}{4}k^2\tilde{h}^2_{ab} - \tfrac{1}{2}\tilde{T}_{ab}\tilde{h}_{ab} \ , & (15.15) \\ \mathcal{L}_2 &= \tfrac{1}{2}k^2h^2_{0a} + h_{0a}T_{0a} \ , & (15.16) \\ \mathcal{L}_3 &= -\tfrac{1}{8}\dot{h}^2 + \tfrac{1}{8}k^2h^2 - \tfrac{1}{2}k^2hh_{00} - \tfrac{1}{2}h_{00}T_{00} - \tfrac{1}{4}hT_{aa} \ . & (15.17) \end{aligned}$$

Here we used the abbreviated notation:

$$\begin{aligned} h^2 &= \int \mathrm{d}^3\vec{k}\, h(\vec{k},t)h(-\vec{k},t) \ , \\ k^2h^2 &= \int \mathrm{d}^3\vec{k}\, k^2h(\vec{k},t)h(-\vec{k},t) \ ,. & (15.18) \end{aligned}$$

The Lagrangian $\mathcal{L}_1$ has the usual form of a harmonic oscillator. Since $\tilde{h}_{ab} = \tilde{h}_{ba}$ and $\tilde{h}_{aa} = 0$, there are only two degrees of freedom (forming a spin 2 representation of the rotation group around the $\vec{k}$ axis: "gravitons" are particles with spin 2). $\mathcal{L}_2$ has no kinetic term. It generates the following Euler-Lagrange equation:

$$h_{0a} = -\frac{1}{k^2}T_{oa} \, . \tag{15.19}$$

We can substitute this back into $\mathcal{L}_2$:

$$\mathcal{L}_2 = -\frac{1}{2k^2}T^2_{0a} \, . \tag{15.20}$$

Since there are no further kinetic terms this Lagrangian produces directly a term in the Hamiltonian:

$$\begin{aligned} H_2 &= -\int \mathcal{L}_2\mathrm{d}^3\vec{k} = \int \frac{1}{2k^2}T^2_{0a}\mathrm{d}^3\vec{k} = \int \Big(\frac{\delta_{ij} - k_ik_j/k^2}{2k^2}\Big)T_{0i}(\vec{k})T_{0j}(-\vec{k})\mathrm{d}^3\vec{k} = \\ &= \tfrac{1}{2}\int T_{0i}(\vec{x})\left[\Delta(\vec{x}-\vec{y})\delta_{ij} - E_{ij}(\vec{x}-\vec{y})\right]T_{oj}(\vec{y})\mathrm{d}^3\vec{x}\mathrm{d}^3\vec{y} \ ; & (15.21) \\ &\text{with} \quad \partial^2\Delta(\vec{x}-\vec{y}) = -\delta^3(\vec{x}-\vec{y}) \quad \text{and} \quad \Delta = \frac{1}{4\pi|\vec{x}-\vec{y}|} \ , \end{aligned}$$

whereas E_{ij} is obtained by solving the equations

$$\partial^2E_{ij}(\vec{x}-\vec{y}) = \partial_i\partial_j\Delta(\vec{x}-\vec{y}) \quad \text{and} \quad (x_i - y_i)\, E_{ij}(\vec{x}-\vec{y}) = 0 \, , \tag{15.22}$$

so that

$$E_{ij} = \frac{\delta_{ij}}{8\pi|\vec{x}-\vec{y}|} - \frac{(\vec{x}-\vec{y})_i(\vec{x}-\vec{y})_j}{8\pi|\vec{x}-\vec{y}|^3}\,. \tag{15.23}$$

Thus, $\mathcal{L}_2$ produces effects which are usually only very tiny relativistic corrections to the instantaneous interactions between the Poynting components of the stress-energy-momentum tensor.

In $\mathcal{L}_3$ we find that h_{00} acts as a Lagrange multiplier. So the Euler-Lagrange equation it generates is simply:

$$h \;=\; -\frac{1}{k^2}T_{00}\,, \tag{15.24}$$

leading to

$$\mathcal{L}_3 \;=\; -\dot{T}_{00}^2/8k^4 + T_{00}^2/8k^2 + T_{00}T_{aa}/4k^2\,. \tag{15.25}$$

Now for the source we have in a good approximation

$$\partial_\mu T_{\mu\nu} \;=\; 0\;, \tag{15.26}$$

$$\text{so}\qquad ikT_{3\nu} \;=\; \dot{T}_{0\nu} \qquad\text{and}\qquad ikT_{30} \;=\; \dot{T}_{00}\;, \tag{15.27}$$

and therefore one can write

$$\mathcal{L}_3 \;=\; -T_{30}^2/8k^2 + T_{00}^2/8k^2 + T_{00}T_{aa}/4k^2\;; \tag{15.28}$$

$$H_3 \;=\; -\int \mathcal{L}_3\mathrm{d}^3\vec{k}\;. \tag{15.29}$$

Here the second term is the dominant one:

$$-\int \mathrm{d}^3\vec{k}T_{00}^2/8k^2 \;=\; -\int\frac{T_{00}(\vec{x})T_{00}(\vec{y})\mathrm{d}^3\vec{x}\mathrm{d}^3\vec{y}}{8\cdot 4\pi|\vec{x}-\vec{y}|} \;=\; -\frac{G_N}{2}\int\frac{\mathrm{d}^3\vec{x}\mathrm{d}^3\vec{y}}{|\vec{x}-\vec{y}|}T_{00}(\vec{x})T_{00}(\vec{y})\;, \tag{15.30}$$

where we re-inserted Newton's constant. This is the linearized gravitational potential for stationary mass distributions. The other terms have to be processed as in Eqs. (15.21)–(15.23).

We observe that in the radiation gauge, $\mathcal{L}_2$ and $\mathcal{L}_3$ generate contributions to the forces between the sources. It looks as if these forces are instantaneous, without time delay, but this is an artifact peculiar to this gauge choice. There is gravitational radiation, but it is all described by $\mathcal{L}_1$. We see that $\tilde{T}_{ab}$, the traceless, spacelike, transverse part of the energy momentum tensor acts as a source. Let us now consider a small, localized source; only in a small region V with dimensions much smaller than $1/k$. Then we can use:

$$\begin{aligned}\int T^{ij}\mathrm{d}^3\vec{x} \;&=\; \int T^{kj}(\partial_k x^i)\mathrm{d}^3\vec{x} \;=\; -\int x^i\partial_k T^{kj}\mathrm{d}^3\vec{x}\\ &=\; \partial_0\int x^iT^{0j}\mathrm{d}^3\vec{x} \;=\; \partial_0\int x^i(\partial_k x^j)T^{0k}\mathrm{d}^3\vec{x}\\ &=\; \tfrac{1}{2}\partial_0\int\partial_k(x^ix^j)T^{0k}\mathrm{d}^3\vec{x} \;=\; -\tfrac{1}{2}\int x^ix^j\partial_kT^{0k}\mathrm{d}^3\vec{x}\end{aligned}$$

$$= \tfrac{1}{2}\partial_0^2 \int x^i x^j T^{00} \mathrm{d}^3\vec{x}\ . \tag{15.31}$$

This means that, when integrated, the space-space components of the energy momentum tensor can be identified with the second time derivative of the *quadrupole moment* of the mass distribution T_{00}.

We would like to know how much energy is emitted by this radiation. To do this let us momentarily return to electrodynamics, or even simpler, a scalar field theory. Take a Lagrangian of the form

$$\mathcal{L} = \tfrac{1}{2}\dot{\varphi}^2 - \tfrac{1}{2}k^2\varphi^2 - \varphi J\ . \tag{15.32}$$

Let J be periodic in time:

$$J(\vec{x},t) = J(\vec{x})e^{-i\omega t}\,, \tag{15.33}$$

then the solution of the field equation (see the lectures about classical electrodynamics) is at large r:

$$\varphi(\vec{x},t) = -\frac{e^{ikr}}{4\pi r}\int J(x')\mathrm{d}^3x'\ ; \qquad k = \omega\,, \tag{15.34}$$

where x' is the *retarded* position where one measures J. Since we took the support V of our source to be very small compared to $1/k$ the integral here is just a spacelike integral. The energy P emitted per unit of time is

$$\frac{\mathrm{d}E}{\mathrm{d}t} = P = 4\pi r^2\Big(\tfrac{1}{2}\dot{\varphi}^2 + \frac{k^2}{2}\varphi^2\Big) = \frac{k^2}{4\pi}\Big|\int J(x')\mathrm{d}^3x'\Big|^2 = \frac{1}{4\pi}\Big|\int \partial_0 J(\vec{x})\mathrm{d}^3\vec{x}\Big|^2\ . \tag{15.35}$$

Now this derivation was simple because we have been dealing with a scalar field. How does one handle the more complicated Lagrangian $\mathcal{L}_1$ of Eq. (15.15)?

The traceless tensor

$$\hat{T}_{ij} = T_{ij} - \tfrac{1}{3}\delta_{ij}T_{kk}\,, \tag{15.36}$$

has 5 mutually independent components. Let us now define inner products for these 5 components by

$$\hat{T}^{(1)}\cdot\hat{T}^{(2)} = \tfrac{1}{2}\hat{T}^{(1)}_{ij}\hat{T}^{(2)}_{ij}\,, \tag{15.37}$$

then (15.15) has the same form as (15.32), except that in every direction only 2 of the 5 components of $\hat{T}_{ij}$ act. If we integrate over all directions we find that all components of $\hat{T}_{ij}$ contribute equally (because of rotational invariance, but the total intensity is just $2/5$ of what it would have been if we had $\hat{T}$ in $\mathcal{L}_1$ instead of $\tilde{T}_{ab}$. Therefore, the energy emitted in total will be

$$\begin{aligned} P &= \frac{2k^2}{5\cdot 4\pi}\cdot\tfrac{1}{2}\Big(\int \hat{T}_{ij}(\vec{x})\mathrm{d}^3\vec{x}\Big)^2 \\ &= \frac{2}{20\pi}\cdot\tfrac{1}{2}\Big(\tfrac{1}{2}\partial_0^3\hat{t}_{ij}\Big)^2 \\ &= \frac{G_N}{5}(\partial_0^3\hat{t}_{ij})^2\ , \end{aligned} \tag{15.38}$$

with, according to (15.31),

$$\hat{t}_{ij} \;=\; \int (x^i x^j - \tfrac{1}{3}\vec{x}^2\delta_{ij})T_{00}\mathrm{d}^3\vec{x}\ . \tag{15.39}$$

For a bar with length L one has

$$\begin{aligned}\hat{t}_{11} &= \frac{1}{18}ML^2\,,\\ \hat{t}_{22} &= \hat{t}_{33} = -\frac{1}{36}ML^2\ .\end{aligned} \tag{15.40}$$

If it rotates with angular velocity Ω then $\hat{t}_{11}$, $\hat{t}_{12}$ and $\hat{t}_{22}$ each rotate with angular velocity 2Ω:

$$\begin{aligned}\hat{t}_{11} &= ML^2(\frac{1}{72}+\frac{1}{24}\cos 2\Omega t\Big)\,,\\ \hat{t}_{22} &= ML^2(\frac{1}{72}-\frac{1}{24}\cos 2\Omega t\Big)\,,\\ \hat{t}_{12} &= ML^2(\frac{1}{24}\sin 2\Omega t\Big)\ ,\\ \hat{t}_{33} &= -\frac{1}{36}ML^2\ .\end{aligned} \tag{15.41}$$

Eqs. (15.41) are derived by realizing that the $\hat{t}_{ij}$ are a (5 dimensional) representation of the rotation group. Only the rotating part contributes to the emitted energy per unit of time:

$$P \;=\; \frac{G_N}{5}(2\Omega)^6\Big(\frac{ML^2}{24}\Big)^2(2\cos^2 2\Omega t) + 2\sin^2 2\Omega t) \;=\; \frac{2G_N}{45c^5}M^2L^4\Omega^6\,, \tag{15.42}$$

where we re-inserted the light velocity c to balance the dimensionalities.

Eq. (15.38) for the emission of gravitational radiation remains valid as long as the movements are much slower than the speed of light and the linearized approximation is allowed. It also holds if the moving objects move just because they are in each other's gravitational fields (a binary pulsar for example), but this does not follow from the above derivation without any further discussion, because in our derivation it was assumed that $\partial_\mu T_{\mu\nu} = 0$.