Cosmology: Basics and Current Trends

Nicolas Berg and Andreas de Vries

Version: November 26, 2025

Contents

Chapter 1

General relativity

1.1 The logical foundation of general relativity

Like any physical theory, general relativity theory as a logical construct has precisely defined axioms that are assumed and from which all other laws and properties can be deduced. The essential axiom of general relativity theory is the equivalence principle. Einstein elegantly formulated it in such a way that it "automatically" implied two laws that he wanted to physically require from his theory of gravity: the correspondence principle and the equality of inertial and gravitational mass. The "correspondence principle" refers to the relationship between gravitational theory and special relativity theory and states that locally, i.e., on a small scale or in areas of weak gravity, the laws of special relativity theory apply approximately. In this way, general relativity theory contains special relativity as a limiting case, because with vanishing masses and uniform motions of bodies they merge into one another [1, p. 266].

The equality of heavy and inertial mass, on the other hand, states that the weight of a body in a given gravitational field is always proportional to its inertial mass: all bodies fall at the same speed in the same gravitational field [1, pp. 37, 269ff], for example, in the Earth's gravitational field at the Earth's surface with a uniform acceleration $g = 9.81$ m/s^2. Accordingly, one can decrease one's weight by accelerating in the opposite direction, and vice versa. Today, when television images and video clips of freely floating astronauts are familiar to everyone, this fact hardly surprises anyone. Regarding these two principles, Einstein unified them and formulated the now so-called Einstein equivalence principle:

Axiom 1.1.1 (Einstein equivalence principle). *[35, S. 88] No local experiments can distinguish a non-rotating system falling freely in a gravitational field (a "local inertial frame") from a uniformly moving system in gravity-free space.*

Expressed somewhat sloppier, the equivalence principle states: *No local experiments can distinguish between acceleration and gravity.* However, both formulations of the equivalence principle are somewhat vague, as it is not entirely clear what is meant by "local experiments." It should initially be understood as a heuristic principle, i.e., a pragmatic principle based on empirical evidence, which is concisely specified by the mathematical requirement formulated in section 1.1.1.

1.1.1 Gravitation curves space and time

Formally, the equivalence principle leads to general relativity in the sense of the fundamental equality of all inertial systems, i.e., all freely falling reference frames. Physically, it provides an interpretation of gravity. According to the equivalence principle, an astronaut in a rocket cannot distinguish whether the rocket is still on the ground in the gravitational field of the Earth or is already in a vacuum and being pushed upward with the acceleration due to gravity g.

Let us imagine two observers, Alice and Bob, in a gravity-free space, with Bob in an elevator moving at a constant acceleration (Figure 1.1). For Alice, being at rest, a ray of

Figure 1.1. Alice and Bob in an elevator accelerated by g in a gravity-free space. On the left is Alice's view, on the right is Bob's view, who perceives a gravitational field g and Alice accelerated by g.

light through the accelerated elevator describes a straight line while for the accelerated Bob it describes a curved path. Since according to the equivalence principle Bob cannot distinguish whether he is accelerated at g or is at rest in a gravitational field g, we must conclude that a gravitational field deflects light rays. Assuming that light rays travel along the shortest possible path between two points in spacetime, a *geodesic*, Einstein drew the radical conclusion that the curvature of the path of a reference frame depends only on the geometry of space and time: Space and time form a four-dimensional geometric unit called *"spacetime"*. It is curved, but at each point it locally looks like flat Minkowski spacetime (the world of gravity-free special relativity with a causality structure of past and future), similar to how the Earth's surface looks locally like a plane at each of its points. [1, p. 273].

Definition 1.1.2. The mathematical model of a gravitational field according to general relativity is a *spacetime*, i.e. a four-dimensional semi-Riemannian manifold $(\mathscr{M}, \boldsymbol{g})$, whose metric tensor $\boldsymbol{g}$ has the same signature as the Minkowski metric $\eta_{ij} = \operatorname{diag}(1, -1, -1, -1)$. The metric $\boldsymbol{g}$ determines the causality relations of spacetime; locally, there is always a future and a past light cone. Moreover, in local coordinates we usually write "g_{ij}" instead of "$\boldsymbol{g}$," meaning the (4×4) matrix with components in the coordinate system.

The equivalence principle then states that all physical laws are determined by the following two requirements [35, §2.I.3].

1. In addition to the metric and its derivatives, the equations of the laws may only contain quantities that already appear in the special relativistic formulation. The metric and its derivatives are derived from the equations of the laws. The equations of the laws are derived from the metric and its derivatives.

2. They must be covariant and must reduce to the special relativistic form if they are stated in a coordinate frame where we have $g_{ij}(x_0) = \eta_{ij}$ and $g_{ij,k}(x_0) = 0$ for $x_0 \in \mathscr{M}$. Here the comma denotes the partial derivative, i.e., $g_{ij,k} = \partial g_{ij}/\partial x_k$.

The covariant form of a physical law is largely determined by these two requirements. However, there are fundamental ambiguities in equations with higher-order derivatives, since partial derivatives commute with each other, but covariant derivatives do not. [35, §2.I.4.4].

1.2 The Einstein field equations

In 1859, astronomer Urbain Le Verrier discovered Mercury's perihelion rotation of 45 arc seconds per century. It remained an unsolved mystery of celestial mechanics for 56 years. What was the problem? Contrary to Newton's classical theory of gravity, Mercury does

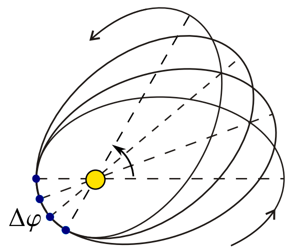


Figure 1.2. Perihelion precession of Mercury. $\Delta\varphi = 43''$ per century [31, p. 113]

not move in a closed ellipse, but in a rosette orbit, so that the perihelion, i.e. the point in the orbit closest to the sun, slowly advances with each orbit.

At a meeting of the *Königlich-Preußische Akademie der Wissenschaften* in Berlin on November 18, Einstein derived the perihelion rotation of Mercury as an approximate solution from his approach to a new theory of gravitation. [11, Eq. (13)], cf. [30, pp. 362–366] To achieve this result he had to use an approximation of second order of the curvature terms of spacetime. He obtained the value of 43″ per century, which nearer to Le Verrier's observation than any theoretical explanaton so far.[1] The following week, he presented the general formulation of the field equations [10, Equations (2a) and (6)] for the first time:[2] General Relativity was born at Thursday, November 25, 1915 [13, p. 419]. Today the Einstein field equations are written as

$$\underbrace{R_{ij} - \tfrac{1}{2} g_{ij}\, R - \Lambda g_{ij}}_{\text{geometry}} \;=\; \underbrace{\kappa T_{ij}}_{\text{physics}} \tag{1.1}$$

with the *Einstein constant* $\kappa = 8\pi G/c^4$ [9, Equ. (89)] and the *cosmological constant* $\Lambda \in \mathbb{R}$, representing physically a homogeneous negative effective energy density $\rho_{\text{eff}} = -2\Lambda/c^2\kappa$ of the vacuum. [23, §111], [31, pp. 69, 96], [30, p. 350], [34, p. 86], [35, pp. 149, 151] If $\Lambda \neq 0$, a vacuum spacetime has to be curved necessarily. [3] However, observations suggest that $\Lambda/c^2\kappa$ cannot surmount essentially the mean mass density of the universe, $\rho = 10^{-30}$ g/cm^3. [31, p. 143] Thus a non-vanishing Λ plays an important role only for cosmological models.

The quantities g_{ij}, R_{ij} and T_{ij} in (1.1) are second-order tensor fields, which can be represented for 4-dimensional spacetime in given coordinates at each spacetime point by

[1] Modern measurements [31, p. 113] correct Le Verriers value to $43.11'' \pm .45$, confirming Einstein's prediction.

[2] Einstein originally formulated the field equations without the Λ term and with a different the designation G_{im} for the Ricci tensor R_{im}, as it is now usually denoted, and with $\Lambda = 0$ in a different, but equivalent, form than in (1.1) below. Note moreover that both the Λ term and the T_{ij} term change sign, depending on the definition of the curvature tensors.

[3] At first sight a non-vanishing cosmological constant therefore contradicts the correspondence principle, that a spacetime without gravitation must be the Minkowski spacetime. However, formally the term $g_{ij}\,\Lambda$ can be brought to the right hand side of (1.1), thus being part of the physical energy-momentum distribution of the spacetime.

symmetric (4×4) matrices,

$$g_{ij} = \begin{pmatrix} g_{11} & g_{12} & g_{13} & g_{14} \\ g_{21} & g_{22} & g_{23} & g_{24} \\ g_{31} & g_{31} & g_{33} & g_{34} \\ g_{41} & g_{42} & g_{43} & g_{44} \end{pmatrix} \tag{1.2}$$

where $g_{ij} = g_{ji}$, i.e., there are only 10 independent components. Both R and the components of the tensor $R_{..}$ are functions of the tensor $g_{..}$ and its first and second partial derivatives with respect to x_1, x_2, x_3, and x_4:

$$R_{ij} = R_{ij}\left(g_{..}, \frac{\partial g_{..}}{\partial x_k}, \frac{\partial^2 g_{..}}{\partial x_l \partial x_m}\right), \qquad R = R\left(g_{..}, \frac{\partial g_{..}}{\partial x_k}, \frac{\partial^2 g_{..}}{\partial x_l \partial x_m}\right).$$

Especially, they depend linearly on the second derivatives of $g_{..}$, but nonlinear in the first derivatives. The central quantity on the left side of the equation (1.1) is therefore g_{ij}, the *metric tensor*. It determines the local distances between two points in spacetime, i.e., the "metric" of spacetime. The tensor is thus a purely geometric quantity and also indirectly determines the curvature of spacetime.

On the left-hand side of the field equations, there are therefore only terms that describe the *geometry* of spacetime, i.e., the curvatures of space and time. The tensor T_{ij} on the right-hand side, the "energy-momentum tensor," describes the physics, i.e., the distribution of energy and matter in spacetime.

The field equations are a system of 10 nonlinear second-order partial differential equations for the 10 components

$$g_{ij} = g_{ij}(x_0, x_1, x_2, x_3).$$

It is not easy to find exact solutions. Usually, one assumes certain physically reasonable symmetries of the spacetime, restricting the degrees of freedom for the g_{ij}. Often one also defines the physics of a gravitational system to be considered, i.e., the components T_{ij}.

To summarize, the basic idea of the field equations can be summarized as follows: *Matter and energy curve space and time, while on the other hand, the curvature of spacetime determines the motion and distribution of matter and light. Due to this recursion, the field equations* must *be nonlinear.* The degree of curvature depends on the size of the

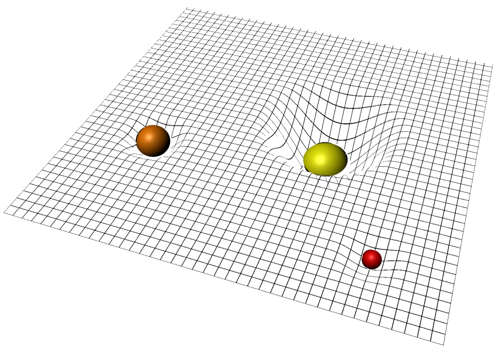


Figure 1.3. Mass and energy curve space-time, and conversely, the curvature of space-time determines the distribution of mass and energy. Graphics modified from `https://www.esa.int/spaceinimages/Images/2015/09/Spacetime_curvature`

masses and energies involved: the larger the mass, the greater the curvature (Figure 1.3).

1.2.1 The first confirmation by observation

On the initiative of British astronomer Arthur Eddington, who, despite the isolation of German scientists during the World War, had learned about Einstein's new theory from Dutch physicists, two expeditions succeeded in observing the total solar eclipse on May 29, 1919 in northern Brazil and on the west coast of Africa. They returned with several photographs of the stars in the sun's vicinity, but due to bad weather, only a few of them could be evaluated. The results of the measurements were published just under six months later, on November 6, 1919 [13, p. 444f, 492ff]. It was a triumph for Einstein's theory. The deflection of 1.75 arc seconds at the edge of the sun, which Einstein had predicted, was confirmed [1, p. 309].

The news spread around the world like wildfire, and Einstein became an international pop star practically overnight. The enormous public response was due, on the one hand, to the fascination that a completely new theory, radically contrary to everyday thinking, had been brilliantly confirmed. On the other hand, the news certainly also struck a chord

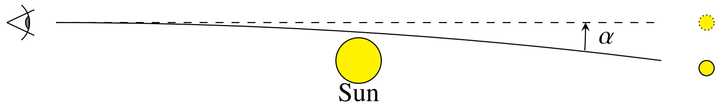


Figure 1.4. Light deflection by the sun. During a solar eclipse, stars appearing close to the sun appear to shift away from it. This effect can only be observed during total solar eclipses. The sun's light is deflected by the moon's gravity, causing stars to appear to shift away from the sun.

with the global public's desire for peace after the terrible and hitherto unprecedented industrialized violence of the world war that had just ended, because, after all, the theory of a German scientist, and Swiss citizen, had been confirmed by a British team of scientists: a signal for the international and peaceful impact of science.

1.2.2 Gravitational redshift

Let us consider two experimenters in a spacecraft that maintains a constant acceleration g. The distance between the two observers in the direction of g is h, cf. Fig. 1.5a). At time

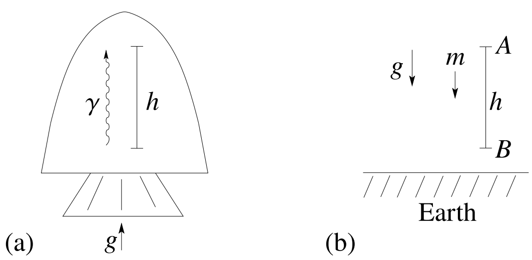


Figure 1.5. (a) Redshift by acceleration. (b) Redshift by gravitation.

$t = 0$, the lower observer sends a photon γ in the direction of the upper observer. Let us assume that the spacecraft is at rest at time $t = 0$ relative to an initial system. Let $v = gt$ be the velocity of the two observers after time t. Under the assumption $v \ll c$, we can neglect terms of the order v/c in the following approximation. Then the photon reaches the upper observer in the time $t \approx h/c$. Now the upper observer has reached the velocity

$v = gt \approx gh/c$. Due to the Doppler effect, he observes the photon with a redshift

$$z := \frac{\Delta\lambda}{\lambda} \approx \frac{v}{c} \approx \frac{gh}{c^2}. \tag{1.3}$$

According to the equivalence principle, the same redshift can also be expected for two observers in a homogeneous gravitational field. Since for a weak homogeneous gravitational field (such as on the Earth's surface) $\Delta\Phi = gh$ applies to Newton's gravitational potential Φ, we can also write Newton's approximation for the redshift

$$z = \frac{\Delta\Phi}{c^2}. \tag{1.4}$$

This formula could be confirmed up to 1% by terrestrial experiments basing on the Mößbauer effect.

The redshift (1.3), however, also follows from the conservation of energy. Let us assume two points A and B in a homogeneous gravitational field g at a distance h (Figure 1.5b). A mass m falls from A to B with an initial velocity of 0. At point B, according to Newton's theory, it has a kinetic energy $E_{\text{kin}} = mgh$. Now let us imagine that the entire energy of the falling body, i.e., rest energy plus kinetic energy, is annihilated at point B into a photon that moves back to point A in the gravitational field. If the photon had no interaction with the gravitational field, If the photon did not interact with the gravitational field, we could convert it back into a mass m there and would gain the energy $\Delta E = mgh$ in this circular process. In order to preserve the law of conservation of energy, the photon must therefore be shifted to the red. Therefore we obtain the following identities for the photon energy

$$E_{\text{below}} = E_{\text{above}} + mgh = mc^2 + mgh = mc^2\left(1 + \frac{gh}{c^2}\right) = E_{\text{above}}\left(1 + \frac{gh}{c^2}\right).$$

This means the the change of wave length

$$1 + z = \frac{\lambda_{\text{above}}}{\lambda_{\text{below}}} = \frac{\hbar\omega_{\text{below}}}{\hbar\omega_{\text{above}}} = \frac{E_{\text{below}}}{E_{\text{above}}} = 1 + \frac{gh}{c^2}, \tag{1.5}$$

coincides with (1.3).

1.3 Exact vacuum solutions of the field equations

Historically, the first exact solutions of the Einstein field equations with vanishing cosmological constant were found by Karl Schwarzschild in 1916 and by Alexander Friedmann in 1922. Schwarzschild's solution describes the gravitational field in a vacuum, enabling eventually the representation black holes, whereas Friedmann's solution represents a universe filled with a matter fluid of given density and pressure. In this section we shortly consider the first one as a prominent vacuum solution. We return to Friedmann's solution below in chapter 2.2.

The simplest class of exact solutions to Einstein's field equations (1.1), with $\Lambda = 0$, are the *vacuum solutions*. They solve the field equations with a vanishing energy-momentum tensor, i.e.,

$$T_{ij} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{1.6}$$

The vacuum solutions thus represent gravitational fields in a vacuum. Since the field equations with (1.6) are relatively easy to solve, the first exact solutions to be discovered were two vacuum solutions: the Minkowski space-time (by construction according to the equivalence principle, of course) and the Schwarzschild solution.

1.3.1 Minkowski spacetime of special relativity

The simplest solution to the field equation is the *Minkowski spacetime*. Its metric tensor in orthogonal coordinates (t, x, y, z) is constant

$$g_{ij} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{1.7}$$

Since all derivatives of g_{ij} disappear, space and time are not curved. In particular, light rays move in straight lines here, but there is also no gravity. Minkowski space-time is the model of special relativity. Of course, Einstein knew this trivial solution to his field equations; he had deliberately constructed his system of equations in such a way that Minkowski spacetime would solve them, cf. the correspondence principle above. However, for a non-vanishing cosmological constant this principle is not satisfied. [35, p. 151]

1.3.2 Black holes

Einstein himself was unable to determine a non-trivial exact solution to his field equations. But only a few months after their publication, the astronomer Karl Schwarzschild succeeded in finding the first exact solution to the complicated differential equations for the case of a vacuum outside a given mass. The solution found by Schwarzschild and now named after him ultimately describes a black hole, but this was not understood as a possible physical reality until the 1960s. The metric tensor that solves the field equations has a simple diagonal form

$$g_{ij} = \begin{pmatrix} g_{tt} & 0 & 0 & 0 \\ 0 & g_{rr} & 0 & 0 \\ 0 & 0 & g_{\vartheta\vartheta} & 0 \\ 0 & 0 & 0 & g_{\varphi\varphi} \end{pmatrix} \tag{1.8}$$

and is usually described by the coordinates $(r, \vartheta, \varphi, ct)$, where t is time and (r, ϑ, φ) are the spatial spherical coordinates. The three spatial components of g_{ij} have a sign different from g_{tt}, and the magnitude of the g_{rr} component is the reciprocal of the magnitude of g_{tt},

$$g_{rr} = -\frac{1}{g_{tt}}. \tag{1.9}$$

In the 1960s, Kerr discovered an exact solution that generalized the Schwarzschild solution and contains non-vanishing components in the secondary diagonal. The Kerr solution represents a rotating gravitational source in a vacuum. [2, Chap. 6]

1.4 General relativity in everyday life

Most of the most important applications of general relativity are, naturally, in astronomy, especially astrophysics and cosmology. After all, considerable gravitational fields or space and time scales are required to observe or utilize relativistic effects. In everyday life, one would not expect Einstein's hundred-year-old field equations to have any impact or even significance at all. After all, it is not exactly common to have a black hole in one's handbag or a small inflationary universe in one's closet. And warp drive, which would allow one to travel to the nearest wormhole around the corner, is not yet functional. It is therefore all the more surprising that there are indeed very practical This makes it all the more astonishing that there are indeed very practical applications and explanations of the general theory of relativity that affect our everyday lives.

1.4.1 Olbers' paradox

For a long time, it was an unexplained phenomenon that the sky is dark at night. Olbers' paradox states the following:

Theorem 1.4.1 (Olbers' paradox (1823) [26]). *If the universe is infinite and uniform in space and time and filled with stars like our sun, then the sky is at least as bright as the average luminosity L of the stars everywhere.*

Proof. [36] Each spherical shell with radius r and thickness $\mathrm{d}r$ around the Earth contains $4\pi\rho r^2\,\mathrm{d}r$ stars, where ρ denotes the average star density. The apparent luminosity of a star on Earth is L/r^2. Therefore, each spherical shell radiates with an apparent luminosity of $4\pi\rho L\,\mathrm{d}r$. This integrand is constant and, in particular, independent of r. □

But why is it dark at night? At least one of the three premises of the sentence must be false so that it does not contradict our observations. Either the universe is finite in space or time (or both), or it is not uniformly filled with sun-like stars. Einstein's field equations provide, for the first time in the history of science, physically testable cosmological models that solve Olbers' paradox by allowing for a temporally finite universe.

1.4.2 GPS

GPS (Global Positioning System, officially NAVSTAR GPS) is a global satellite-based navigation system for determining position. To do this, the satellites constantly transmit their current position and the exact time using coded radio signals. GPS receivers can then calculate their own position and speed from the signal transit times. In theory, signals from three satellites with simultaneous radio contact are sufficient for this. However, since GPS receivers do not necessarily use an accurate clock to measure the transit time, the signal from a fourth satellite is required to determine the exact time in the receiver. GPS was launched in February 1978 and is based on 31 satellites.

A physical system rotating in a circular orbit around a static spherically symmetric gravitational field of mass M can be described in general relativity by the Schwarzschild metric [7, 8] with $\mathrm{d}r = \mathrm{d}\varphi = 0$,

$$\mathrm{d}s^2 = c^2\,\mathrm{d}\tau^2 = \left(1 - \frac{2GM}{c^2 r}\right) c^2\,\mathrm{d}t^2 - r^2\,\mathrm{d}\vartheta^2. \tag{1.10}$$

Here, τ denotes the proper time of the rotating system. (t denotes the coordinate time of the entire gravitational field and corresponds to the proper time of a stationary observer at infinity, $r \to \infty$.) It follows that

$$\left(\frac{\mathrm{d}\tau}{\mathrm{d}t}\right)^2 = \left(1 - \frac{2GM}{c^2 r}\right) - \frac{r^2}{c^2}\,\frac{\mathrm{d}\vartheta}{\mathrm{d}t} = 1 + \frac{2\Phi(r)}{c^2} - \frac{v^2}{c^2} \tag{1.11}$$

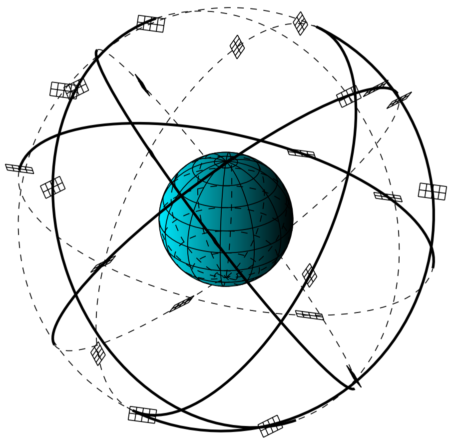


Figure 1.6. Satellite-based navigation systems such as GPS (Source: `https://tug.org/TUGboat/tb33-1/tb103wolcott.pdf`)

with the gravitational potential $\Phi(r)$ [35, p. VI.2]. and the orbital speed $v(r)$ given by

$$\Phi(r) = -\frac{MG}{r}, \qquad v(r) = r\,\frac{\mathrm{d}\vartheta}{\mathrm{d}t} \tag{1.12}$$

This results in the proper times τ_E and τ_S of a receiver on Earth with Earth radius r_E and a satellite orbiting the Earth in a circular orbit with radius r_S [27]

$$\left(\frac{\mathrm{d}\tau_S}{\mathrm{d}\tau_E}\right)^2 = \frac{1 + 2\Phi(r_S)/c^2 - v_S^2/c^2}{1 + 2\Phi(r_E)/c^2 - v_E^2/c^2} \tag{1.13}$$

Since the gravitational field of the Earth is quite small on the Earth's surface, i.e., $\Phi(r_E)/c^2 \ll 1$, and since the velocities v_E and V_S are much smaller than the speed of light, the approximation can be applied for equation (1.13), neglecting terms higher than $O(c^{-2})$, [27, 32].

$$\frac{\mathrm{d}\tau_S}{\mathrm{d}\tau_E} = \frac{\sqrt{1 + 2\Phi(r_S)/c^2 - v_S^2/c^2}}{\sqrt{1 + 2\Phi(r_E)/c^2 - v_E^2/c^2}} = 1 + \frac{MG}{c^2}\left(\frac{1}{r_E} - \frac{1}{r_S}\right) + \frac{v_E^2 - v_S^2}{2\,c^2} + O(c^{-3}) \tag{1.14}$$

With the difference $\Delta\Phi = \Phi(r_E) - \Phi(r_S)$ between the two gravitational potentials and the difference Δv^2 between the squares of the velocities,

$$\Delta\Phi = MG\left(\frac{1}{r_E} - \frac{1}{r_S}\right), \qquad \Delta v^2 = v_S^2 - v_E^2 \tag{1.15}$$

we thus have

$$\frac{\mathrm{d}\tau_S}{\mathrm{d}\tau_E} \approx 1 + \frac{\Delta\Phi}{c^2} - \frac{\Delta v^2}{2\,c^2}. \tag{1.16}$$

See [33, Gl. (14.12)]. This means that the time displayed by the atomic clocks on the GPS satellites is subject to special relativistic *and* gravitational time dilation. According to general relativity, the rate at which a clock runs depends on its location in the gravitational field, i.e., on $\Delta\Phi$, and according to special relativity, it depends on its velocity v_S. The lower gravitational potential in the satellite's orbit causes time to pass more quickly than on Earth, while the orbital motion of the satellites relative to a stationary observer on Earth slows it down. The two time dilations therefore have opposite effects. In other words, for an observer on the Earth's surface, time passes slower than for an observer on the satellite, who is not moving relative to him, by a factor of $1 + \frac{\Delta\Phi}{c^2}$.

For the values $M = 5.9810^{24}$ kg, $r_E = 6.3810^6$ m and $r_S = 26.5610^6$ m, equation (1.15) yields $\Delta\Phi/c^2 = 5.28 \cdot 10^{-10}$ and $\Delta v^2 = 8.5 \cdot 10^{-11}$ m/s. In the GPS satellite orbit, the gravitational effect therefore predominates by more than six times: time therefore passes faster on the satellites. Equation (1.16) yields the relative satellite velocity $v_S = 3900$ m/s and the relative rate difference

$$\frac{\mathrm{d}\tau_S}{\mathrm{d}\tau_E} = 1 + 4{,}44 \cdot 10^{-10} \tag{1.17}$$

the proper times on Earth and in the satellite. This ratio is significantly greater than the relative accuracy of cesium atomic clocks, which is of the order of 10^{-13}. To compensate for this time dilation, the oscillation frequency of the satellite clocks is detuned to $\mu_S =$ 10.229999995453 MHz, so that they always run synchronously with the frequency $\mu_E =$ 10.23 MHz of the terrestrial receiver clocks. (The sample actually yields $\mu_E/\mu_S - 1 = 4.44 \cdot 10^{-10}$.)

If time dilation were not compensated for, the difference in the rate of the clocks per year would add up to a time difference of of approximately 0,014 seconds. Accordingly, time dilation would lead to a deviation

$$\frac{\mathrm{d}s}{\mathrm{d}\tau_E} - c = c\left(\frac{\mathrm{d}\tau_S}{\mathrm{d}\tau_E} - 1\right) = 0{,}133 \text{ m/s} \tag{1.18}$$

Per second, the car would deviate 13 cm from its calculated course, or 8 meters per minute or 11.5 km per day. However, this error would not occur even without the corrective adjustment of the satellite clocks in the real GPS system, because at least four satellites are always contacted to determine time and location, so that only the satellite time is actually used.

At an altitude of about 3,000 km, special relativistic and gravitational time dilation cancel each other out.

Chapter 2

Cosmological models

In this chapter we will shortly review the standard models of cosmology. In essence they follow from the classes of geometry that are determined by the cosmological principle. In turn, the geometry and its symmetries prescribe the physical form and distribution of matter and energy in the universe, due to the Einstein field equations (1.1).

The standard models of cosmology are essentially the three geometries of the Friedmann-Lemaître-Robertson-Walker metrics, each of which has variants with or without a cosmological constant, i.e., $\Lambda \neq 0$ or $\Lambda = 0$, respectively. The all rely on the cosmological principle, which in essence fixes the possible geometries and therefore prescribe the overall distribution of energy and matter in the cosmos.

The standard models of the universe all go back to Alexander Friedmann. They and more modern modifications such as the "inflationary universe" in its initial phase have repeatedly been confirmed by physical observations. The first important confirmation was the discovery of the expansion of the universe by the Belgian priest and astrophysicist Georges Lemaître in 1927 [24] and the US astronomer Edwin Hubble in 1929 [17]. Further confirming observations were the discovery of cosmic background radiation by Penzias and Wilson in 1965 and the discovery of the Higgs boson in 2012, cf. [22, §20.1], [37, §13.3]

2.1 The cosmological principle

The cosmological principle summarizes two basic assumptions of cosmology that underlie its models of the universe as a whole.

Axiom 2.1.1 (Cosmological principle). *Viewed on a sufficiently large scale, the properties of the universe are the same for all observers. Especially the universe is homogeneous and isotropic. Homogeneity means that the universe appears the same to all observers regardless to their locations in space, isotropy means that the universe always appears the same to observers regardless of their direction of observation in space.*

In general, the two concepts homogeneity and isotropy are independent. For instance, a structure of equidistant parallel lines is homogeneous in the plane, but not isotropic, and concentric lines are istropic around the center but not homogeneous.

Remark 2.1.2. From the point of view of differential geometry, isotropy at each point means local rotational symmetry everywhere and implies the space to have a constant curvature. Therefore, isotropy at each point implies homogeneity, but a homogeneous universe could be anisotropic somewhere. [22, p. 448], [31, p. 130] □

The cosmological principle is based on the assumption that the uniformity of the universe observed from Earth cannot be explained by a special position, but rather by any position in the universe. This assumption cannot be proved. Its validity has to be assumed *a priori*, i.e., it is an axiom.

The cosmological principle does not apply to small distances. For example, the density of matter in the solar system is significantly greater than in interstellar space. Furthermore, individual galaxies are not evenly distributed, but form groups, clusters, superclusters, and filaments. On an even larger scale, however, the cosmological principle has been repeatedly confirmed by increasingly accurate measurements.

2.2 Friedmann-Lemaître-Robertson-Walker universes

Translated into the language of Riemannian geometry, the cosmological principle implies that the three-dimensional space has maximum symmetry, i.e., has constant curvature which, if variable at all, can only depend on time. There are three possible geometries of space satisfying this criterion, depending on a parameter k. [34, p. 222] It will turn out that k geometrically determines the spatial curvature of the universe, and physically $-k$ represents a part of its energy density. [31, p. 155]

Remark 2.2.1. One remarkable property of standard FLRW cosmology, as it will be described in the sequel, is that the energy density of the universe depends on its *global* geometry, represented by a discrete curvature parameter k. Of course, *locally* this follows immediately from the Einstein field equations (1.1), since geometry determines physics, and physics determines geometry. But *globally* it is not clear at first sight, why there should be only three cases. In fact, this is a consequence of the strong symmetry requirements due to the cosmological principle 2.1.1 and Remark 2.1.2. □

For $k \in \{-1, 0, 1\}$ let $\mathscr{M}_k$ denote one of the three manifolds

$$\mathscr{M}_k = \begin{cases} \mathbb{R}^4 & \text{for } k = -1, 0, \\ \mathbb{R} \times S^3 & \text{for } k = 1. \end{cases} \tag{2.1}$$

Define then the coordinates $(t, \chi, \theta, \varphi) \in \mathbb{R} \times I_k \times (0, \pi) \times (0, 2\pi)$ of $\mathscr{M}_k$ with the intervals

$$I_k = \begin{cases} (0, \infty) & \text{for } k = -1, 0, \\ (0, \pi) & \text{for } k = 1, \end{cases} \tag{2.2}$$

Especially for $k = -1, 0$, let $(\chi, \theta, \varphi) \in (0, \infty) \times (0, \pi) \times (0, 2\pi)$ denote the polar coordinates of $\mathbb{R}^3$ (with the radial coordinate χ). For $k = 1$, on the other hand, let $(\chi, \theta, \varphi) \in (0, \pi) \times (0, \pi) \times (0, 2\pi)$ denote the polar coordinates of S^3 (i.e., χ is an angular coordinate). Moreover, let

$$R \in C^\infty(\mathbb{R}, (0, \infty)), \qquad t \mapsto R(t), \tag{2.3}$$

be a function depending on t, and let $f_k \colon I_k \to J_k$ denote the functions

$$f_k(\chi) = \begin{cases} \sinh \chi & \text{for } k = -1, \\ \chi & \text{for } k = 0, \\ \sin \chi & \text{for } k = 1. \end{cases} \tag{2.4}$$

The parameter t is called *cosmic time*, and R the *scale factor* of the universe. Defining the covariant components of the metric tensor by

$$(g_{ij}) = \operatorname{diag}(1, -R^2, -R^2 f_k^2, -R^2 f_k^2 \sin^2 \theta),$$

the line element reads

$$ds^2 = dt^2 - R^2(t)\left(d\chi^2 - f_k^2(\chi)\,d\theta^2 - f_k^2(\chi)\,d\varphi^2\right). \tag{2.5}$$

cf. [23, (112, 2)]. $\sqrt{g} = R^3(t)\, f_k^2(\chi)\sin\theta$. For $k = 1$, the spatial part of(2.5), i.e., $R^2(t)\,(d\chi^2 + \sin^2\chi\, d\theta^2 + \sin^2\chi\sin^2\theta\, d\varphi^2)$, is exactly the three-dimensional line element of a hypersphere with radius $R(t)$ embedded in Euclidian space $\mathbb{R}^4$. Therefore, the scale factor R is sometimes also called the *world radius*. [31, pp. 149] Especially, for $k = 1$ the universe has a finite volume. In general, the volume of a 3-space with coordinates (x_1, x_2, x_3) and a metric with determinant g is defined as $V = \iiint \sqrt{g}\, dx_1\, dx_2\, dx_3$, cf. (A.13). For the line element (2.5) we therefore have

$$V = \int_0^{\pi}\int_0^{\pi}\int_0^{2\pi} \sqrt{g}\, d\varphi\, d\theta\, dr = R^3(t)\int_0^{\pi}\sin^2\chi\, d\chi \cdot \int_0^{\pi}\sin\theta\, d\theta \cdot \int_0^{2\pi} d\varphi = 2\pi^2 R^3(t). \tag{2.6}$$

Remark 2.2.2. In cosmology it is more familiar to consider another coordinate than χ, namely the parameter $r \in J_k$, with $J_k = I_k$ for $k = 0, -1$ and $J_1 = (0, 1)$ for $k = 1$, given by

$$\chi \mapsto r = f_k(\chi), \qquad \text{i.e.,} \qquad d\chi = \frac{dr}{\sqrt{1 - kr^2}}. \tag{2.7}$$

where $\chi = f_k^{-1}(r) \in (0, \frac{\pi}{2})$. Thus the metric (2.5) transforms to

$$ds^2 = c^2\, dt^2 - R^2(t)\left(\frac{dr^2}{1 - kr^2} + r^2\, d\theta^2 + r^2\sin^2\theta\, d\varphi^2\right), \tag{2.8}$$

and the determinant of the metric is $\sqrt{g} = R^3(t)\frac{r^2}{\sqrt{1-r^2}}$. This is the line element of the *Friedmann-Lemaître-Robertson-Walker (FLRW) models*. It is a metric of a space-time that is obtained from the cosmological principle of space-like isotropy at each fixed cosmic time $t = \text{const.}$ [16], [31]

For $k = 1$, i.e., a three-dimensional hypersphere $S^3(R)$ with world radius R, the metric gets singular at $r = 1$. However, this is not a physical singularity, but a pure coordinate singularity. An analogy in two dimensions may illustrate this, see Fig. 2.1. We have

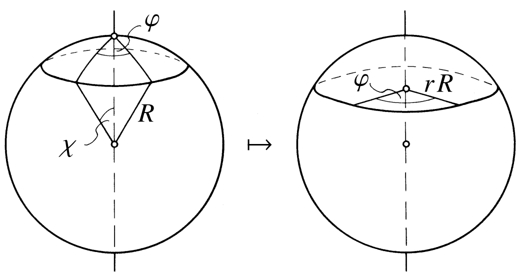


Figure 2.1. The coordinate change $(\chi, \varphi) \mapsto (r, \varphi)$ on the sphere $S^2(R)$ embedded in three-space $\mathbb{R}^3$, as an analog to the coordinate change (2.7) on $S^3(R)$. Graphics modified from [31, p. 151]

$r = \sin\chi$, but there is no singularity at the equator $\chi = \frac{\pi}{2}$. Moreover, this analogy illustrates the role of the cosmological scale factor R. A sphere S^2 is a two-dimensional surface embedded in a 3-dimensional space. Its radius lives in the third dimension, it is not part of the surface. However, the value of this radius affects distances measured on the two-dimensional surface. Similarly, the cosmological scale factor is not a distance in our 3-dimensional space, but its value affects the measurement of distances. □

2.3 Solving the field equations by FLRW geometries

So far, we only derived the *geometry* of a universe satisfying the cosmological principle. To solve the Einstein field equations (1.1), however, we have to find suitable distributions of energy and matter in the spacetime, determined by T_{ij}. In other words, a physical model of the universe is derived, if (2.8) is inserted in the Einstein field equations. By the strong spatial symmetries implied by the cosmological principle they only can determine the evolution of the universe in time, i.e., its dynamics. It turns out that, given the FLRW geometry, the energy-momentum tensor must represent an ideal fluid with pressure p and energy density ρ, i.e., must be of the form

$$T_{ij} = pg_{ij} + (\rho + p/c^2)u_i u_j \tag{2.9}$$

with c the speed of light and u_i the four-velocity of the fluid with respect to the rest system of the FLRW metric (2.8). Moreover, p and ρ depend only on the cosmic time but not on their spatial positions, i.e., $p = p(t)$ and $\rho = \rho(t)$. [34, pp. 83, 233–234] In the rest system of the fluid, we have

$$T_{ij} = \begin{pmatrix} c^2\rho & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{pmatrix} \tag{2.10}$$

Inserting this tensor into the Einstein field equations yields the two ordinary differential equations for $R(t)$,

$$\frac{\dot{R}^2}{R^2} + \frac{2\ddot{R}}{R} + \frac{c^2 k}{R^2} = -\kappa p + c^2\Lambda \tag{2.11}$$

and

$$\frac{3\dot{R}^2}{R^2} + \frac{3c^2 k}{R^2} = \kappa c^4 \rho + c^2\Lambda \tag{2.12}$$

[31, p. 154], [34, p. 234] Here $\dot{R}$ denotes the derivative with respect to the cosmic time, i.e., $\dot{R} = \mathrm{d}R/\mathrm{d}t$. For $p = 0$, these equations have been derived first by Friedmann in 1922. [14, (4) & (5)] Subtracting $\frac{1}{6}$ times (2.12) from $\frac{1}{2}$ times (2.11) gives

$$\frac{\ddot{R}}{R} = \frac{c^2\Lambda}{3} - \frac{\kappa c^4}{6}\left(\rho + 3p/c^2\right). \tag{2.13}$$

cf. Unsöld and Baschek [38, p. 485]. Hence the system of the two equations (2.11) and (2.12) is equivalent to the system of (2.13) and (2.12). Nowadays, (2.12) and (2.13) are called *Friedmann equations*. Note that equation (2.13) does not contain the geometric parameter k, in contrast to Friedmann's original equation (2.11).

Equation (2.13) states that both the energy density and the pressure cause the expansion rate of the universe $\dot{R}$ to decrease. This is a consequence of gravitation, with pressure playing a similar role to that of energy density. The cosmological constant, on the other hand, causes an acceleration in the expansion of the universe. Under the assumption that the total mass of the universe is constant, i.e., $\rho R =$ const, Equation (2.11) on the other hand expresses the energy conservation of the universe. [31, p. 155]

Another important identity can be deduced from the Friedmann equations, assuming that all the terms R, $\dot{R}$, and $(c^2\rho + p)$ do nat vanish: Deriving first (2.12) times R^2 with respect to t, rearranging the resulting equation for $\ddot{R}/R$, and inserting this into (2.13), we obtain

$$\frac{\dot{\rho}}{\rho + p/c^2} = -\frac{3\dot{R}}{R}. \tag{2.14}$$

cf. [34, (26, 9)]. Physically this equation expresses the first law of thermodynamics, assuming the expansion of the universe is an adiabatic process, i.e., a process without a change of entropy.

2.4 FLRW models with $\Lambda = 0$ and $p = 0$: Exact solutions

The Friedmann equation (2.13) expresses the dynamics of the FLRW models. That is, the behavior in time follows from the energy density ρ and the pressure p. In a matter-dominated universe, as we observe it today, we have $p \approx 0 < \rho$. A cosmos like this is also called "incoherent" or "dust-like". If we set $p = 0$ and let moreover the cosmological constant vanish, (2.14) reduces to $\dot{\rho}/\rho = -3\dot{R}/R$ and can be immediately integrated by

$$\rho R^3 = \mathscr{A} = \text{const.} \tag{2.15}$$

The integration constant $\mathscr{A}$ can be interpreted as a multiple of the total mass $\mathfrak{M}$ of the universe. With (2.6), for instance, we have $\mathfrak{M} = 2\pi^2 \mathscr{A}$. By (2.15) we have $\rho = \mathscr{A}/R^3$, i.e., equation (2.12) simplifies to $(\dot{R})^2 = \kappa c^2 \mathscr{A}/R - k$, which by the transformation $t \mapsto \tau = \pm\frac{1}{c}\int \frac{\mathrm{d}t}{R(t)}$, i.e., $\mathrm{d}\tau = \pm\frac{c}{R}\,\mathrm{d}t$ or $\frac{\mathrm{d}}{\mathrm{d}t} = \pm\frac{c}{R}\frac{\mathrm{d}}{\mathrm{d}\tau}$, attains the form

$$(R')^2 = \kappa\mathscr{A}R/3 - kR^2 \tag{2.16}$$

with $A_0 = \kappa\mathscr{A}/3$, and $'$ denoting the derivative with respect to τ. [34, p. 237] Separation of variables[1] then yields the solutions

$$R(\tau) = \begin{cases} \frac{1}{2}\,A_0\,(1 - \cos\tau) & \text{for } k = 1, \\ \frac{1}{4}\,A_0\,\tau^2 & \text{for } k = 0, \\ \frac{1}{2}\,A_0\,(\cosh\tau - 1) & \text{for } k = -1, \end{cases} \tag{2.17}$$

with $\tau \in I_k$. The cosmic time then satisfies $t = \frac{1}{c}\int_0^\tau R\,\mathrm{d}\tau$, i.e., $t \in J_k$ and

$$t = \pm \begin{cases} \frac{1}{2c}\,A_0(\tau - \sin\tau) & \text{for } k = 1, \\ \frac{1}{12c}\,A_0\tau^3 & \text{for } k = 0, \\ \frac{1}{2c}\,A_0(\sinh\tau - \tau) & \text{for } k = -1. \end{cases}$$

Here $t = 0$ denotes the "beginning of the world." Since R (except for a positive factor corresponding to the radius of the universe) describes the expansion of the universe, there are three different development scenarios depending on the parameter k, Fig. 2.2. The first scenario $k = +1$ is the so-called "big crunch," in which the universe collapses into a singularity after reaching its maximum expansion. This scenario is based on the assumption that the universe is finite and has a finite volume.

Also for a radiation-dominated Friedmann universe elementary solutions can be obtained for $\Lambda = 0$, cf. [34, §26], [31, §5, especially pp. 156f and pp. 160ff], [25, p. 347ff], [23, §§111–113].

[1] In more detail, separation of variables (2.16) yields $\frac{\mathrm{d}R}{\sqrt{A_0 R - kR^2}} = \pm\mathrm{d}\tau$, or $\pm\tau = \int \frac{\mathrm{d}R}{\sqrt{A_0 R - kR^2}}$. For $k = -1$ this is $\pm\tau = \int \frac{\mathrm{d}R}{\sqrt{A_0 R + R^2}} = \operatorname{arcosh}(1 + 2R/A_0)$, for $k = 0$ we have $\pm\tau = \int \frac{\mathrm{d}R}{\sqrt{A_0 R}} = 2\sqrt{R/A_0}$, and for $k = 1$, finally, $\pm\tau = \int \frac{\mathrm{d}R}{\sqrt{A_0 R - R^2}} = \arccos(1 - 2R/A_0)$. The identities for the integral can be shown elementarily by derivation with respect to R. Rearranging the equations for R yields (2.17), independently from the sign of τ.

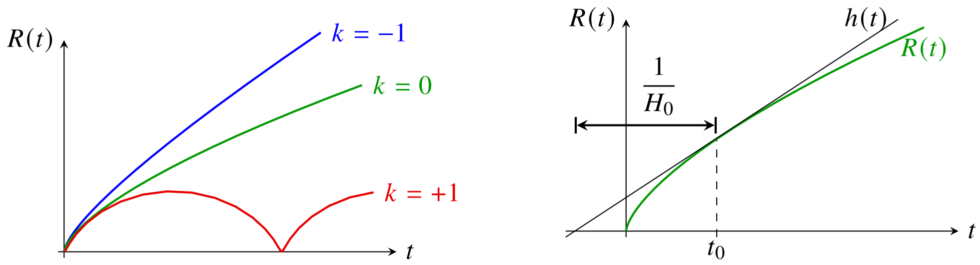


Figure 2.2. Expansions of the three FLRW cosmologies with $\Lambda = 0$. The reciprocal of the Hubble constant H_0 represents an upper limit for the age of the universe, as can be seen from the tangent equation $h(t) = \dot{R}(t_0)(t - t_0) + R(t_0)$.

If we form the quotient of the derivative $\dot{R}$ of the world radius and R, we obtain the relative rate of change of the universe

$$H(t) = \frac{\dot{R}(t)}{R(t)}, \tag{2.18}$$

the *Hubble function* H. For the time t_0 = "now," $H_0 := H(t_0)$ is called the *Hubble constant*. Its reciprocal $1/H_0$ indicates the maximum age of the universe for a non-accelerating expanding universe [31], because the equation of the tangent of the graph of $R(t)$ at the point t_0 is $h(t) = \dot{R}(t_0)(t - t_0) + R(t_0)$, and its zero point is given by

$$t_0 - t = \frac{R(t_0)}{\dot{R}(t_0)} = \frac{1}{H_0}, \tag{2.19}$$

see Figure 2.2. For a nonvanishing cosmological constant, $\Lambda > 0$, we have qualitatively the same results for $k = 0, +1$. For $k = -1$, however, the vacuum energy density prohibits the big crunch occurring for Λ greater than a critical value Λ_c, i.e., $\Lambda > \Lambda_c \geqq 0$. In this case the universe instead expands forever. For the illustration of the six FLRW universes see Figure 2.4.

2.5 FLRW models with $\Lambda > 0$

A positive cosmological constant Λ represents a vacuum energy density. In general, for $\Lambda > 0$ there are no elementary solutions of the Friedmann equations (2.12) and (2.13), as determined in Section 2.4. A remarkable exception is the Einstein cosmos which will we consider shortly. It turns out, however, that the qualitative behaviors of the FLRW models with $\Lambda > 0$ can be classified by the three elementary solutions (2.17).

2.5.1 The Einstein cosmos

The first exact cosmological solution of the field equations (1.1) was accomplished by Einstein in 1917 [12], two years after their publication. According to the state of knowledge at that time, he assumed the universe to be static. This implies that all derivatives with respect to time in (2.12) and (2.13) vanish, i.e.,

$$\frac{3k}{R^2} = \Lambda + \kappa c^2 \rho, \qquad \Lambda = \frac{\kappa}{2}\left(c^2 \rho + 3p\right). \tag{2.20}$$

Inserting the second equation into the first one we obtain $\frac{3k}{R^2} = \frac{\kappa}{2}\left(c^2\rho + 3p\right) + \kappa c^2\rho$, or

$$k = \frac{\kappa R^2}{2}\left(c^2\rho + p\right) > 0. \tag{2.21}$$

Since k can only be -1, 0, or $+1$, it must be $k = 1$. This, in turn, determines the constant radius R to be

$$R^2 = \frac{2}{\kappa(c^2\rho + p)}. \tag{2.22}$$

By $\mathfrak{M} = V\rho$ and (2.6) the total mass of the universe is

$$\mathfrak{M} = 2\pi^2 R^3 \rho = \frac{\sqrt{32}\,\pi^2\rho}{\sqrt{\kappa^3(c^2\rho + p)^3}}. \tag{2.23}$$

In other words:

Theorem 2.5.1. *A static isotropic universe must be a 3-sphere with a constant radius given by (2.22) and with a total mass given by (2.23). It is called the Einstein cosmos.*

Example 2.5.2. According to astronomical observations, the current energy density of the universe is[2] $\rho = 9.9 \cdot 10^{-27}$ kg/m^3 and $p = 0$. With $\kappa = 2.07665 \cdot 10^{-43}$ s^{2}kg^{-1}m^{-3} equation (2.22) gives

$$R = \frac{1}{\sqrt{\Lambda}} = \sqrt{\frac{2}{\kappa c^2\rho}} = 1.04 \cdot 10^{26} \text{ m} = 11 \cdot 10^9 \text{ ly}, \tag{2.24}$$

and (2.23)

$$\mathfrak{M} = \frac{\sqrt{32}\,\pi^2}{\sqrt{\kappa^3 c^6 \rho}} = \frac{\pi^2}{c^3}\sqrt{\frac{32}{\kappa^3\rho}} = 2.2 \cdot 10^{53} \text{ kg} = 1.1 \cdot 10^{23}\,\mathfrak{M}_\odot. \tag{2.25}$$

A light ray propagates on a geodesic line of the 3-sphere, i.e., a great circle. The length of the longest possible path of a photon therefore is the circumference of the sphere, $\ell = 2\pi R$. For the Einstein cosmos we therefore have $\ell = 6.54 \cdot 10^{26}$ m $= 69 \cdot 10^9$ ly. Interestingly, the diameter of the observable universe, which is a 2-sphere with the Earth in the center, is estimated to be 93×10^9 ly. If we lived in the Einstein cosmos, it would last "only" about another $69 - 93/2 = 69 - 46.5 = 22.5$ billion lightyears so that we can see light from the Earth having been emitted then 69 billion years ago. □

However, the Einstein cosmos is unstable in the sense that any slight change in either the value of the cosmological constant or the matter density will result in a universe that either expands and accelerates forever, or collapses to a singularity.

2.5.2 The de Sitter universes

In 1917, shortly after Einstein published his *Kosmologische Betrachtungen*, de Sitter introduced vacuum models of the universe with a non-vanishing cosmological constant. [4–6] In fact, for $\rho = p = 0$ the Friedmann equations (2.12) and (2.13) read

$$\frac{3\dot{R}^2}{R^2} + \frac{3c^2k}{R^2} = c^2\Lambda, \qquad \ddot{R} = \frac{c^2\Lambda}{3}\,R. \tag{2.26}$$

[2] https://map.gsfc.nasa.gov/universe/uni_matter.html

In fact, multiplying the first equation by R^2 and deriving it, gives exactly the second equation, i.e., the second equation is redundant here. However, to solve the first one it is convenient to start the general solutions of the second one and inserting them into the first one. Thereby they are filtered by the curvature parameter k such that we obtain the possible elementary solutions [34, p. 235]

$$R(t) = \begin{cases} \dfrac{1}{A}\cosh Act & \text{for } k = +1, \\ R_0\, \mathrm{e}^{Act} & \text{for } k = 0, \\ \dfrac{1}{A}\sinh Act & \text{for } k = -1, \end{cases} \qquad \text{with } A = \sqrt{\Lambda/3},\ R_0 \in \mathbb{R}^+. \tag{2.27}$$

In any case, the scale factor thus grows exponentially in time. Remarkably, de Sitter universes with a spherical geometry, $k = 1$, or a flat geometry, $k = 0$, do not initiate with a bing bang, but with a positive scale factor $R(0) = 1/\sqrt{\Lambda/3}$ or R_0, respectively. Moreover, for a negative cosmological constant we have another solution,

$$R(t) = \frac{1}{A}\sin Act \quad \text{or} \quad R(t) = \frac{1}{A}\cos Act \qquad \text{with } A = \sqrt{-\Lambda/3},\ k = -1. \tag{2.28}$$

Therefore the only solution to representing an isotropic vacuum universe with a negative cosmological constant has a hyperbolic geometry, may start either with a big bang, $R(0) = 0$, or with a finite scale factor $R(0) = 1/\sqrt{-\Lambda/3} > 0$, and eventually terminates with a big crunch in any case. However, a negative cosmological constant expressing negative

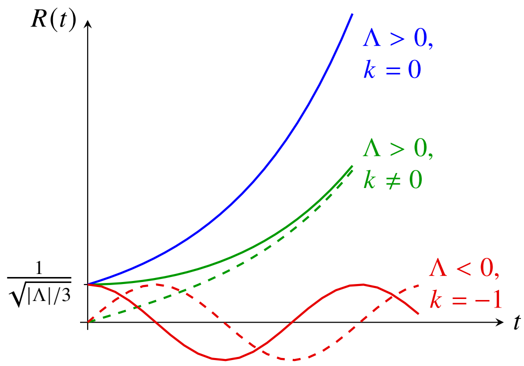


Figure 2.3. Scale factor $R(t)$ of the de Sitter universes according to equations (2.27) and (2.28). The dashed lines show the two cases, both with $k = -1$, starting with a big bang.

vacuum energy seems unprobable according to the astronomical observations.

Theorem 2.5.3. *If the energy density and the pressure are negligibly small compared to the cosmological constant, i.e., $c^2\rho$, $p \ll \Lambda$, all solutions of the Friedmann equations (2.12) and (2.13) are approximately given by (2.27). Especially, the scaling factor R of the universe then grows exponentially with respect to the cosmological time t.*

The exponential expansion of the scale factor means that the physical distance between any two non-accelerating observers will eventually be growing faster than the speed of light. At this point they will no longer be able to communicate with each other.

Nowadays, “the” de Sitter universe usually is referred to the flat geometry case $k = 0$ with $\Lambda > 0$ in (2.27).

2.5.3 Radiation-dominated universe

2.5.4 Matter-dominated universe

2.5.5 The ΛCDM model

The present-day density ρ_0, which according to the current state of knowledge suggest to give zero curvature k. Substituting these conditions to the Friedmann equation gives

$$\rho_{\text{crit}} = \frac{3H_0^2}{8\pi G}. \tag{2.29}$$

If the cosmological constant were actually zero, the critical density would also mark the dividing line between eventual recollapse of the universe to a Big Crunch, or unlimited expansion.

The present-day density parameter Ω_x for various species is defined as the dimensionless ratio

$$\Omega_x \equiv \frac{\rho_x(t_0)}{\rho_{\text{crit}}} = \frac{8\pi G \rho_x(t_0)}{3H_0^2}, \tag{2.30}$$

where the subscript x is one of b for baryons, c for cold dark matter, rad for radiation (photons plus relativistic neutrinos), and Λ for dark energy. By construction we therefore have

$$\Omega_{\text{b}} + \Omega_{\text{c}} + \Omega_{\text{rad}} + \Omega_\Lambda = 1. \tag{2.31}$$

2.5.6 The Gödel universe

2.5.7 Astronomical evidence

FLRW models of the universe and more modern modifications such as the “inflationary universe” in its initial phase have repeatedly been confirmed by physical observations. The first important confirmation was the discovery of the expansion of the universe by the

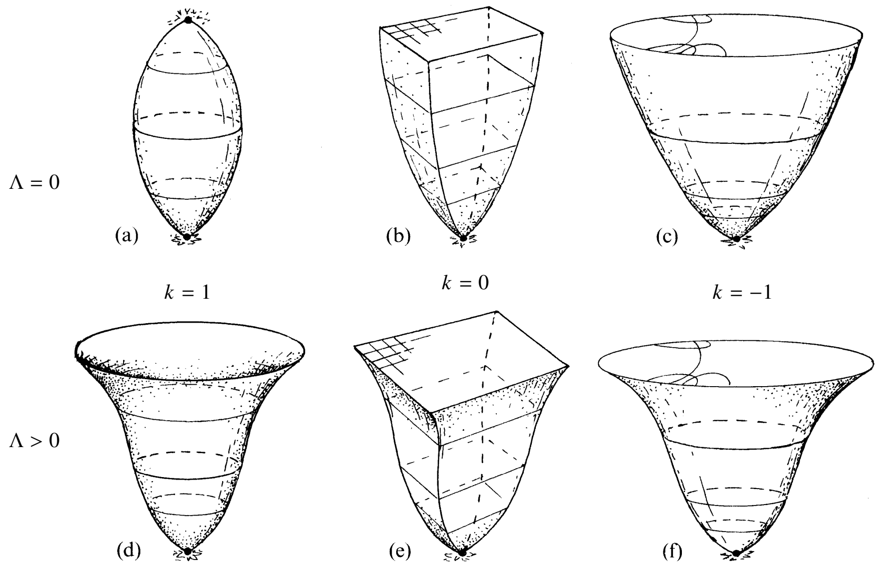


Figure 2.4. The FLRW cosmological models. Time is depicted upwards and each model starts with a big bang. Graphics from [28, p. 719]

Belgian priest and astrophysicist Georges Lemaître in 1927 [24] and the US astronomer Edwin Hubble in 1929 [17]. Further confirming observations were the discovery of cosmic background radiation by Penzias and Wilson in 1965 and the discovery of the Higgs boson in 2012, cf. [22, §20.1], [37, §13.3]

Chapter 3

Conformal geometry

3.1 The complex plane

3.2 Holomorphic mappings

3.3 Conformal mappings

3.4 Singularities and infinities

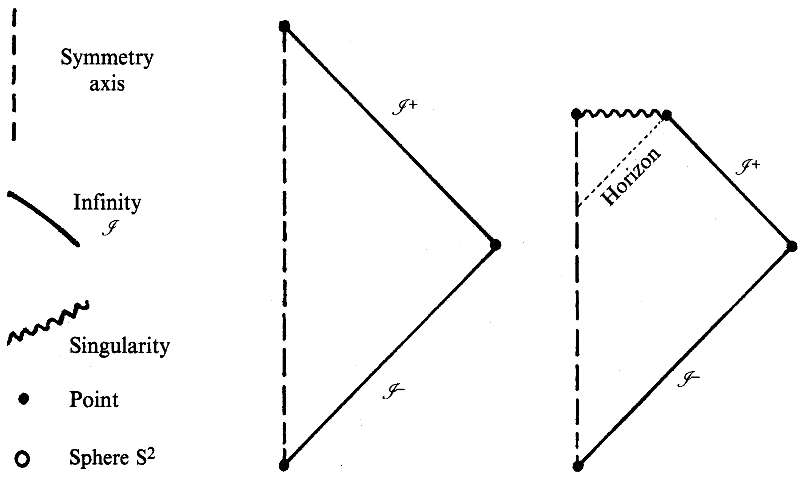


Figure 3.1. Legend for reading Penrose diagrams.

3.5 Conformal null infinity

In this section we consider null cones in General Relativity more precisely. Loosely speaking, the null cones in a space-time represent the sets of all photon trajectories, including closed photon orbits. A useful technique to study regions of infinity in time or space is the conformal rescaling of the space-time leading to a compact manifold the finite boundary of which is the set of the spacetime points at infinity. In this way the analytical behavior null cones at infinity can be considered precisely.

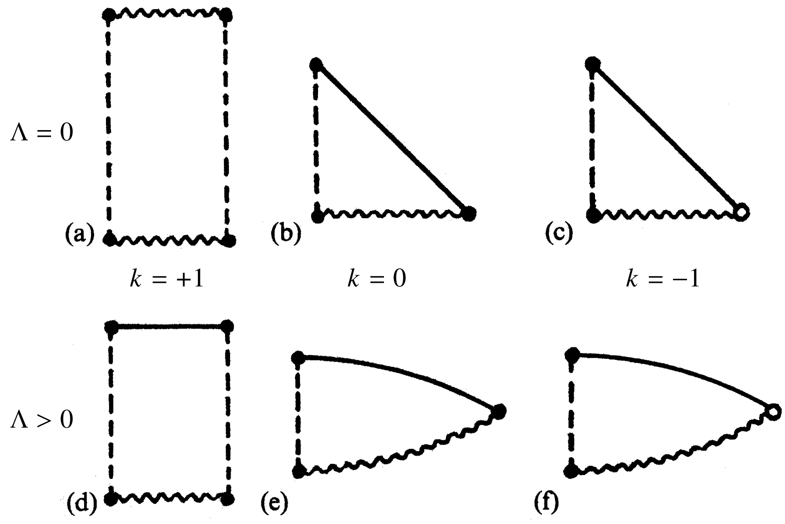


Figure 3.2. Penrose diagrams of the FLWR universes. Cf. Figure 2.4

One of the most useful areas in general relativity has turned out to be the study of asymptotic questions, important examples of which are the definitions of the total energy-momentum contained in an asymptotically flat space-time and of gravitational radiation. For this the technique of conformal rescalings of the space-time $\mathscr{M}$ can be applied, replacing the original physical metric $\mathrm{d}s$ by a new ("unphysical") metric $\mathrm{d}\hat{s}$, which is conformally related to it,

$$\mathrm{d}\hat{s} = \Omega\,\mathrm{d}s, \tag{3.1}$$

Ω being a suitably smooth, everywhere positive function defined on $\mathscr{M}$. The metric tensor g_{ij} and its inverse g^{ij} are accordingly rescaled by

$$g_{ij} \mapsto \hat{g}_{ij} = \Omega^2 g_{ij}, \quad g^{ij} \mapsto \hat{g}^{ij} = \Omega^{-2} g^{ij}. \tag{3.2}$$

Provided that the asymptotic structure of $\mathscr{M}$ is suitable and that Ω is chosen appropriately, it is possible to adjoin to $\mathscr{M}$ a certain boundary surface, denoted by $\mathscr{I}$ (pronounced 'scri' — a contraction of 'script I'), in such a way that the "unphysical" metric $\hat{g}_{ij}$ extends non-degenerately and with some degree of smoothness to these new points. The function Ω may also be extended appropriately but becomes zero on $\mathscr{I}$. This implies that the physical metric would have to be infinite on $\mathscr{I}$, so it cannot be so extended. Thus, from the point of view of the physical metric, the points on $\mathscr{I}$ are infinitely distant from their neighbors. Physically they represent 'points at infinity'.

The adjoining of $\mathscr{I}$ to a space-time $\mathscr{M}$ provides us with a smooth manifold with boundary, denoted by $\bar{\mathscr{M}}$, i.e.

$$\mathscr{I} = \partial\mathscr{M}, \qquad \mathscr{M} = \operatorname{int}\bar{\mathscr{M}} \tag{3.3}$$

(∂ = boundary, int = interior). The advantage is that the powerful *local* techniques of differential geometry can now be employed on $\bar{\mathscr{M}}$ with implications for the asymptotics of $\mathscr{M}$. Indeed, the very definition of asymptotic flatness in general relativity can be given in a convenient and coordinate-free way. Conformal methods are particularly appropriate in general relativity because many of the important concepts are conformally invariant. Among these are the massless free-field equations, the Weyl conformal tensor, null geodesics, null hypersurfaces, and relativistic causality. The technique is similar to that used in complex analysis, where the 'point of infinity' is adjoint to the complex plane to obtain the Riemann sphere, and in projective geometry.

3.5.1 Minkowski space

Let us begin by examinig the construction of conformal infinity for Minkowski space $\mathbb{M}$. It is topologically equivalent to the four-dimensional space, $\mathbb{M} \cong \mathbb{R}^4$. The physical metric, in spherical polar coordinates, is

$$\mathrm{d}s^2 = c^2\,\mathrm{d}t^2 - \mathrm{d}r^2 - r^2(\,\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2). \tag{3.4}$$

For convenience we introduce a retarded time parameter $u = ct - r$ and an advanced time parameter $v = ct + r$, to obtain

$$\mathrm{d}s^2 = \mathrm{d}u\,\mathrm{d}v - \frac{1}{4}\,(v-u)^2(\,\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2). \tag{3.5}$$

There is much freedom in the choice of a conformal factor Ω. However, for 'asymptotically flat' space-times it turns out from general considerations (in the context of the 'peeling theorem') that the choice of Ω must be such that along any ray it approaches zero, both in the past and in the future, like the reciprocal of an affine parameter λ on the ray, i.e. $\Omega\lambda \to$ constant as $\lambda \to \pm\infty$ along the ray. Each u = constant hypersurface is a future light cone, generated by the light rays (null straight lines) for which θ and φ remain constant. The coordinate v serves as an affine parameter into the future on each of these radial rays. Similarly, the coordinate u serves as an affine parameter into the past on these rays. Thus we shall require $\Omega v \to$ constant as $v \to \infty$ on $u, \theta, \varphi =$ constant, and $\Omega u \to$ constant as $u \to -\infty$ on $v, \theta, \varphi =$ constant. If we wish to keep Ω smooth over the finite parts of space-time, then the choice

$$\Omega = \frac{2}{\sqrt{(1+u^2)(1+v^2)}} \tag{3.6}$$

can be made (the factor 2 being for later convenience), and thus, by (3.1),

$$\mathrm{d}\hat{s}^2 = \frac{4\,\mathrm{d}u\;\mathrm{d}v}{(1+u^2)(1+v^2)} - \frac{(v-u)^2(\,\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2)}{(1+u^2)(1+v^2)} \tag{3.7}$$

Many other choices of Ω are possible, of course, but this one is especially convenient.

In order that the 'points at infinity' may be assigned finite coordinates, we replace (u, v) by (p, q), where

$$u = \tan p, \qquad v = \tan q. \tag{3.8}$$

Then

$$\mathrm{d}\hat{s}^2 = 4\;\mathrm{d}p\,\mathrm{d}q - \sin^2(q-p)\,(\,\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2). \tag{3.9}$$

The range of the variables p, q is as indicated in Figure 3.3, in which each point represents a 2-sphere of radius $\sin(q-p)$. The vertical line $q - p = 0$ represents the spatial origin ($r = 0$) and is just a coordinate singularity: the space-time is non-singular on this line. The sloping lines $p = -\frac{1}{2}\pi$, for $-\frac{1}{2}\pi < q < \frac{1}{2}\pi$, and $q = \frac{1}{2}\pi$, for $-\frac{1}{2}\pi < p < \frac{1}{2}\pi$, represent null infinity, denoted by $\mathscr{I}^-$ and $\mathscr{I}^+$, respectively, for Minkowski space (corresponding to $u = -\infty$ and to $v = \infty$). There are indicated three exceptional points representing 2-spheres with vanishing radius $\sin(q-p)$, the points $\mathrm{i}^\pm$ given by $p = q = \pm\pi/2$, and the point i^0 by $q = -p = \pi/2$. They are considered not to belong to $\mathscr{I}^-$ or to $\mathscr{I}^+$.

Physically, we interpret i^- as representing past temporal infinity, $\mathscr{I}^-$ as past null infinity, i^0 as spatial infinity, $\mathscr{I}^+$ as future null infinity, and i^+ as future temporal infinity. The reason for this terminology becomes clear when we examine the behavior of straight lines in Minkowski space (with metric $\mathrm{d}s$). A timelike straight line acquires a past end-point i^- and a future end-point i^+; a null straight line acquires a past end-point on $\mathscr{I}^-$

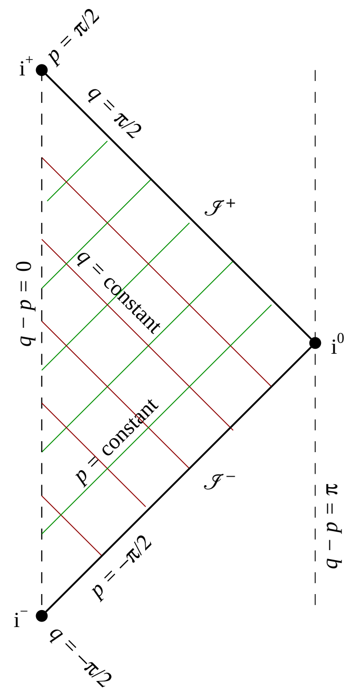


Figure 3.3. Penrose diagram of Minkowski space-time $\mathbb{M}$. The region of (p, q)-space which corresponds to $\mathbb{M}$ and the future (past) null infinity $\mathscr{I}^+$ ($\mathscr{I}^-$), defined by $q = \frac{\pi}{2}$ ($p = -\frac{\pi}{2}$). The line $q - p = 0$ is an axis of spherical symmetry (and so also $q - p = \pi$). Each point in the diagram defines a 2-sphere of radius $\sin(q - p)$. Each line $p = \text{constant}$ and $q = \text{constant}$ represents a light ray.

and a future end-point on $\mathscr{I}^+$; a spacelike straight line becomes a closed curve through i^0. Since rays remain rays after conformal rescalings, the null straight lines become rays with respect to the $d\hat{s}$ metric, whereas the timelike or spacelike straight lines are not, in general, geodesics with respect to $d\hat{s}$.

However, when we consider *curved* lines in Minkowski space, the question of which end-points they acquire is more complicated. For example, the helix curve given by $\varphi = ct$, $r = 1$, $\theta = \pi/2$, is a null curve, since $du = dv = d\varphi = c\,dt$ and $dr = d\theta = 0$ in (3.7), but has a past end-point at i^- and a future end-point at i^+, since $p = \arctan(ct - 1)$, $q = \arctan(ct + 1)$ and $p, q \longrightarrow \pm\infty$ as $t \longrightarrow \pm\infty$. At both points the completed curve is not smooth. On the other hand, the timelike curve $r = c\sqrt{1 + t^2}$, $\theta = \pi/2$, $\varphi = 0$, satisfying $dr^2 = \frac{c^2 t^2\, dt^2}{1+t^2}$, i.e., $ds^2 = \frac{c^2}{t^2+1}\, dt^2 > 0$ in (3.4), smoothly acquires a past-end point on $\mathscr{I}^-$ and a future end-point on $\mathscr{I}^+$, since $p = \arctan(ct - r) \longrightarrow 0$ and $q = \arctan(ct + r) \longrightarrow \pi/2$ as $t \longrightarrow \infty$, and $p \longrightarrow -\pi/2$ and $q \longrightarrow 0$ as $t \longrightarrow -\infty$. The only thing one can prove is that two *causal* curves have the same past-end point on $\mathscr{I}^-$ iff they share the same future, and that they have the same future-end point iff they have the same past. In particular, a past-endless causal curve acquires a past end-point at i^- or at $\mathscr{I}^-$ according as its future is or is not the whole Minkowski space (since, e.g., the future of the time axis $t = 0$ is the whole of $\mathbb{M}$). In exactly the same way, a future-endless causal curve reaches i^+ or a point of $\mathscr{I}^+$ according as its past set is or is not the whole Minkowski space [29, §9.1].

The metric (3.9) is perfectly regular on these regions at infinity, $\mathscr{I}^\pm$, and the points $i^\pm$, i^0. Indeed, the space-time and its metric $d\hat{s}$ can be extended beyond these regions in a non-singular fashion. The vertical line $q - p = \pi$ is again a coordinate singularity, precisely of the same type as that of $q - p = 0$. The entire vertical strip $0 \leqq q - p < \pi$ may be used to define a space-time $\mathscr{E} \cong \mathbb{R} \times S^3$, the *Einstein static universe*. To see this, we choose new coordinates

$$\tau = p + q, \quad \chi = q - p \tag{3.10}$$

and obtain

$$ds^2 = d\tau^2 - \left[d\chi^2 + \sin^2\chi \,(d\theta^2 + \sin^2\theta \, d\varphi)\right]. \tag{3.11}$$

The part in the square bracket represents the metric of a unit 3-sphere S^3. The portion of $\mathscr{E}$ which is conformal to the original Minkowski space may be described as that lying between the light cones of two points i^- and i^+. This portion wraps around the Einstein space to meet at the 'back' in the single point i^0. The situation is illustrated in Fig. 3.4, under suppression of two dimensions. Minkowski 2-space is conformal to the interior of

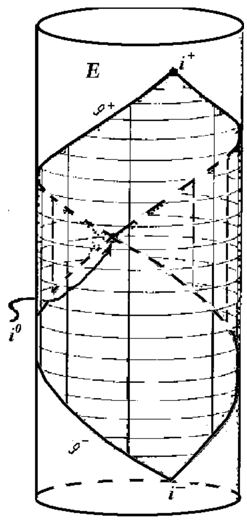


Figure 3.4. The region of the Einstein cylinder $\mathscr{E}$ which corresponds to Minkowski space $\mathbb{M}$. Figure taken from [29].

a square, represented as tilted at 45°. This square wraps around the cylinder which is the two-dimensional version of the Einstein universe.

3.5.2 Infinity in Schwarzschild space-time

We examine the conformal infinity of the Schwarzschild solution. The familiar form of the metric is

$$ds^2 = \left(1 - \frac{2M}{r}\right) c^2\, dt^2 - \left(1 - \frac{2M}{r}\right)^{-1} dr^2 - r^2(d\theta^2 + \sin^2\theta \, d\varphi^2). \tag{3.12}$$

Rather than attempting to obtain $\mathscr{I}^+$ and $\mathscr{I}^-$ simultaneously, as was done for Minkowski space, it is more appropriate to introduce a retarded time coordinate

$$u = ct - r - 2M \ln(r - 2M) \tag{3.13}$$

and an advanced time coordinate

$$v = ct + r + 2M \ln(r - 2M) \tag{3.14}$$

separately. In the first case the metric form becomes

$$ds^2 = \left(1 - \frac{2M}{r}\right) du^2 + 2\, du\, dr - r^2(d\theta^2 + \sin^2\theta \, d\varphi^2), \tag{3.15}$$

and in the second one,

$$ds^2 = \left(1 - \frac{2M}{r}\right) dv^2 - 2\, dv\, dr - r^2 (d\theta^2 + \sin^2\theta\, d\varphi^2). \tag{3.16}$$

In each case we can choose $\Omega = r^{-1} = w$, say. Then the "unphysical" metric is

$$d\hat{s}^2 = \Omega^2\, ds^2 = \left(w^2 - 2Mw^3\right) du^2 - 2\, du\, dw - d\theta^2 + \sin^2\theta\, d\varphi^2 \tag{3.17}$$

in the first case, and

$$d\hat{s}^2 = \left(w^2 - 2Mw^3\right) dv^2 + 2\, dv\, dw - d\theta^2 + \sin^2\theta\, d\varphi^2 \tag{3.18}$$

in the second one. The metrics (3.17) and (3.18) are manifestly regular and analytic on their respective hypersurfaces $w = 0$. (Clearly the determinants are non-zero at $w = 0$.) The physical space-time is given when $w > 0$ in (3.17) and we can extend the manifold to include the boundary hypersurface $\mathscr{I}^+$, given when $w = 0$. Similarly, in (3.18) the physical space-time corresponds to $w > 0$ and can be extended to include $\mathscr{I}^-$, given when $w = 0$. In fact, we could extend the space-time across $w = 0$ to negative values of w, but this will not be done here. Only the boundary $\mathscr{I} = \mathscr{I}^- \cup \mathscr{I}^+$ will be adjoined to the space-time.

There is a difficulty if we try to identify $\mathscr{I}^-$ with $\mathscr{I}^+$. If we do extend the region of definition of (3.17) to include negative values of w, and then make the replacement $w \mapsto -w$, we see that the metric has the form (3.18) with u in place of v, but with a negative mass $-M$ in place of M. Thus, the extension across $\mathscr{I}$ involves a reversal of the sign of the mass. In fact, the derivative at $\mathscr{I}$ of the conformal curvature contains the information of the mass. Therefore, if we attempt to identify $\mathscr{I}^+$ with $\mathscr{I}^-$, and want the *same* sign of the non-zero mass M to occur on the two sides, then there must be a discontinuity in the derivative of the curvature across $\mathscr{I}$, so that the metric must fail to be C^3 at $\mathscr{I}$.

Accepting, then, that is is not reasonable to identify $\mathscr{I}^+$ with $\mathscr{I}^-$, we are led to a picture closely resembling that obtained for Minkowski space. The only essential difference occurs with the points i^-, i^0, i^+. It turns out that whenever mass is present, the point i^0, and normally also $i^\pm$, must, if adjoined to the manifold, be singular for the conformal geometry. It is therefore reasonable not to attempt to include these points, in the general case, as part of the conformal infinity. (Even in Minkowski space the boundary surface at i^0, $i^\pm$ is not smooth.) The picture, then, is as indicated in Fig. 3.5. We have two disjoint boundary hypersurfaces $\mathscr{I}^-$ and $\mathscr{I}^+$ each of which is a "cylinder" with topology $S^2 \times \mathbb{R}$. It is clear from (3.17) and (3.18) that each of $\mathscr{I}^\pm$ is a *null* hypersurface, the induced metric at $w = 0$ being degenerate. These null hypersurfaces are generated by rays, given by θ, φ = constant, $w = 0$, whose tangents are *normals* to the hypersurfaces. These rays may be taken to be the "$\mathbb{R}$s" of the topological product $S^2 \times \mathbb{R}$.

The picture in Fig. 3.5 serves as a model for asymptotically flat spaces generally.

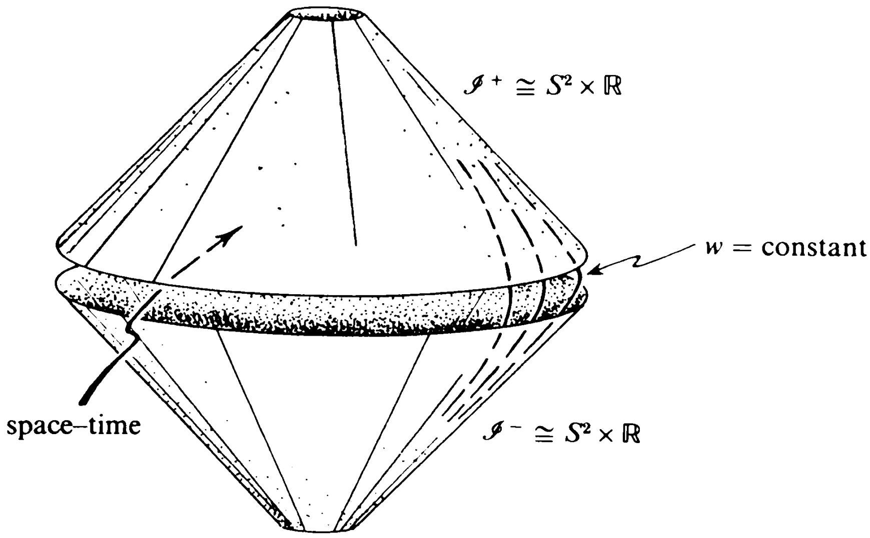


Figure 3.5. Null infinity for the Schwarzschild space-time. Note that $w = 0$ corresponds both to $\mathscr{I}^+$ and to $\mathscr{I}^-$. Figure taken from [29].

Appendix A

Differential geometry

A.1 Smooth manifolds

A *topological space* is a pair $(X, \mathcal{O})$ of a set X and a set $\mathcal{O}$ of subsets of X, defining the "open" subsets, such that the following three axioms are valid. [19, p. 7]

- *Axiom 1:* Any union of open sets is open.
- *Axiom 2:* The intersection of two open sets is open.
- *Axiom 3:* $\emptyset$ and X are open.

Then $\mathcal{O}$ is called the *topology* of X. Often, we shortly write X instead of $(X, \mathcal{O})$. A topological space X is called *Hausdorff* if any two different points of X have disjoint open neighborhoods. [19, p. 22]

Example A.1.1. The topological space $(X, \{\emptyset, X\})$ is not Hausdorff. [19, p. 23] □

A topological space satisfies the *second countability axiom* if its topology $\mathcal{O}$ has a countable basis. [19, p. 98]

Remark A.1.2. The second countability axiom enables to find a countable subcovering for every covering $\{U_\lambda\}_{\lambda\in\Lambda}$, especially for any family of open coverings $\{U_x\}_{x\in X}$. It is required for many inductive constructions and proofs. [19, p. 106] □

A mapping $f : X \to Y$ between two topological spaces X and Y is called *continuous* if the pre-image $f^{-1}(U')$ of an open set $U' \in Y$ is open in X. [19, p. 16]

Remark A.1.3. Especially, if $f : X \to Y$ is continuous and U' is an open neighborhood of $f(x)$ in Y, then $f^{-1}(U')$ is an open neighborhood of x. □

A bijective mapping $f : X \to Y$ is called a *homeomorphism* if both f and f^{-1} are continuous, i.e., if $U \subset X$ is open if and only if $f(U) \subset Y$ is open. [19, p. 17] We then often write shortly $f : X \xrightarrow{\cong} Y$ and call the spaces X and Y being *homeomorphic*.

Definition A.1.4. Let X be a topological space. Then an *n-dimensional chart*, or *coordinate system*, of X is a homeomorphism $h : U \xrightarrow{\cong} U'$ of an open subset $U \subset X$ to an open subset $U' \subset \mathbb{R}^n$. The values $h = (x^1, \ldots, x^n)$ around $p \in U \subset X$ are called *local coordinates*. The set U is called the *chart domain*. The inverse h^{-1} is called a *parametrization* of X. [18, p. 1], [20, p. 369]

Definition A.1.5. If (U, h) and (V, k) are two n-dimensional charts for X, the homeomorphism $k \circ (h^{-1} \mid h(U \cap V))$ from $h(U \cap V)$ onto $k(U \cap V)$ is called a *coordinate change*. If it is differentiable as a function from $\mathbb{R}^n$ to $\mathbb{R}^n$, it is called a *diffeomorphism* and the charts are said to *change differentiably*.

Definition A.1.6. A set $\mathfrak{A}$ of n-dimensional charts of a topological space X, whose chart domains cover the whole space X, is called an *n-dimensional atlas* of X. The atlas is *differentiable* if all coordinate changes are changing differentiably.

The set $\mathscr{D}(\mathfrak{A})$ of all charts (U, h) of X changing differentiably is then an n-dimensional maximal atlas of X. [18, p. 2]

Definition A.1.7. An *n-dimensional differential structure* of a topological space X is a maximal differentiable atlas.

Definition A.1.8. A *smooth manifold* of dimension n is a Hausdorff topological space $\mathscr{M}$ with a differentiable structure $\mathscr{D}$, satisfying the second countability axiom and the property that every point of $\mathscr{M}$ has a neighborhood diffeomorphic to $\mathbb{R}^n$.

Many approximations of complicated situations in physics are done by linearizations. In case of manifolds a kind of local linearization is given by the tangent space of some given point $p \in \mathscr{M}$ of the manifold.

Definition A.1.9. A *tangent vector* v to a smooth manifold $\mathscr{M}$ at a point $p \in \mathscr{M}$ is a linear function from the space of smooth functions defined on some neighborhood of $p \in \mathscr{M}$ which satisfies the Leibniz rule:

- *Linearity:* $v(\alpha f + \beta g) = \alpha v(f) + \beta v(g)$ for $\alpha, \beta \in \mathbb{R}$ and functions f, g on $\mathscr{M}$ differentiable at p;
- *Leibniz rule:* $v(fg) = f(p)\, v(g) + g(p)\, v(f)$.

Then $v(f)$ is the *directional derivative* of f along v. The space $T_p\mathscr{M}$ of tangent vectors to $\mathscr{M}$ at p together with addition and scalar multiplication defined by

$$(\alpha u + \beta v)(f) = \alpha u(f) + \beta v(f)$$

is a vector space of dimension n, the *tangent vector space*.

This definition can be formulated more precisely by identifying functions which coincide on a neighborhood of p: two function f and g on $\mathscr{M}$ differentiable at p have the same *germ* at p if there exists a neighborhood of p where they coincide. The equivalence class of differentiable functions at p whch have the same germ as a function f is called a *germ* of f. The disjoint equivalence classes are called the *germs of differentiable functions* at p. Germs form an algebra. A *tangent vector* then is derivation on the algebra of of germs of differentiable functions at p.

In a local chart (U, h) with coordinates $(x^1, \ldots, x^n)$ a tangent vector $v \in T_p\mathscr{M}$ can be expressed as

$$v^i \partial x^i := v^i \frac{\partial}{\partial x^i} \tag{A.1}$$

with the Einstein summation convention that double indices are summed over. Here the components $(v^1, \ldots, v^n)$ are defined as $v^i = v(x^i)$. [3, p. 119, 18, p. 32]

More intuitively a tangent vector can be described in terms of differentiable parametrized curves $\gamma : (-\varepsilon, \varepsilon) \to \mathscr{M}$ with $\gamma(0) = p$, see Figure A.1. [3, p. 120, 18, p. 29]

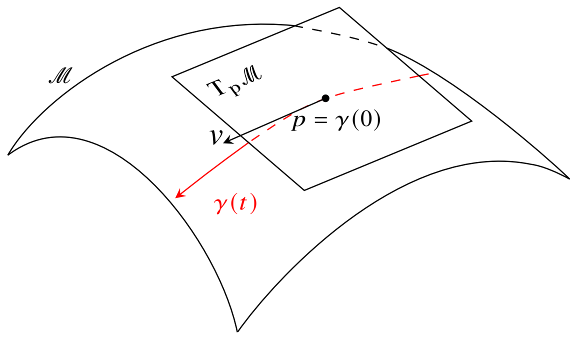


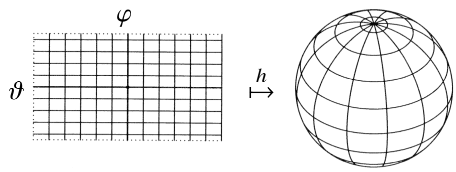
Figure A.1. The tangent space $T_p\mathscr{M}$ of $\mathscr{M}$ at $p \in \mathscr{M}$ can be viewed as the derivative of a curve $\gamma : (-\varepsilon, \varepsilon) \longrightarrow \mathscr{M}$ with $\gamma(0) = p$. [3, p. 120, 18, p. 29] Graphic modified from G. Mezzovilla under CC-BY 4.0

Example A.1.10. An illustrative example of a two-dimensional manifold is the surface $S^2(r)$ of a three-dimensional sphere of radius r in R^3,

$$S^2(r) = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = r^2\}. \tag{A.2}$$

Without its poles $\{z = \pm r\}$ it has the chart $h : S^2(r) \setminus \{z = \pm r\} \longrightarrow (0, \pi) \times (0, 2\pi)$,

$$\begin{pmatrix} \vartheta \\ \varphi \end{pmatrix} \overset{h}{\mapsto} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r\ \sin\vartheta\cos\varphi \\ r\ \sin\vartheta\sin\varphi \\ r\ \cos\vartheta \end{pmatrix}. \tag{A.3}$$



The local coordinates $(\vartheta, \varphi) \in (0, \pi) \times [0, 2\pi)$ are called *spherical coordinates*. □

Remark A.1.11. If $\mathscr{M}$ is a differentiable manifold of class Ck, $T\mathscr{M}$ is a differentiable manifold of class C^{k-1}, and it is called a *differentiable tangent bundle* of class C^{k-1}. However there is no canonical isomorphism between the fiber at a point and the typical fiber, and hence no canonical isomorphism between fibers at different points, unless the fiber bundle is given an additional structure. Such a structure may be for instance parallel displacement as considered in case of semi-Riemannian manifolds. □

A.2 Smooth fiber bundles

In mathematics, and particularly topology, a fiber bundle is a space that is locally a product space, but globally may have a different topological structure. In fact, bundle have been introduced to generalize topological products. The need to generalize topological products can be seen already by a simple example: a cylinder obtained by glueing a strip of paper is the Cartesian product $S^1 \times (-1, 1)$ of a 1-sphere S^1 (a circle) and a line segment $(-1, 1)$, see Figure A.2. But a Möbius strip is obtained by twisting a strip of paper and then glueing it. Therefore it cannot be described globally as a product space: For $U \subsetneq S^1$ the topological product $U \times [0, 2\pi)$ decribes a segment of the Möbius strip. But we need some mechanism to say that twisting occurs somewhere.

Definition A.2.1. A *fiber bundle* is a 6-tuple $(E, B, F, \pi, \mathscr{F}, G)$ consisting of:

- Manifolds E, B, and F, called the *total space*, the *base*, and the *typical fiber*, respectively.

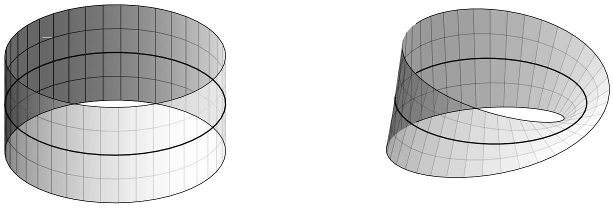


Figure A.2. A cylinder $S^1 \times (-1, 1)$ and a Möbius strip. For each open arc $U \subsetneq S^1$ of the circle, $U \times (-1, 1)$ is a local chart of both the cylinder and the Möbius strip.

- A surjective mapping $\pi : E \longrightarrow B$, called the *projection*. For any point $p \in B$ the preimage F_p is called the *fiber* at x.
- A family $\mathscr{F} = \{(U_i, \varphi_i)\}$ of open sets $U_i \subset B$ covering B and homeomorphisms $\varphi_i : \pi^{-1}(U_i) \to U_i \times F$ such that the following diagram commutes:

$$\begin{tikzcd} \pi^{-1}(U_i) \arrow[r, "\varphi"] \arrow[d, "\pi"'] & U_i \times F \arrow[dl, "\mathrm{pr}_1"] \\ U_i & \end{tikzcd}$$

  The sets $\{(U_i, \varphi_i)\}$ are called *local trivializations* of the bundle. As a consequence, for each $x \in \pi^{-1}(U_i) \subset E$ the homeomorphisms can be written as $\varphi(x) = (\pi(x), \hat{\varphi}_i(x))$ by unique homeomorphisms $\hat{\varphi}_i : F_p \longrightarrow F$ where F_p with $p = \pi(x) \subset B$ is the fiber over p. To simplify notation, let for $p \in U_i$ denote $\hat{\varphi}_{i,p} := \hat{\varphi}_i|_{F_p}$.
- A topological group G of homeomorphisms $F \longrightarrow F$, called the *structure group*, such that for any $p \in U_i \cap U_j \subset B$ the homeomorphism $\hat{\varphi}_{i,p} \circ \hat{\varphi}_{j,p}^{-1} : F \longrightarrow F$ is an element of G. The mappings $f_{ij} : U_i \cap U_j \longrightarrow G$ induced by $p \mapsto f_{ij}(p) = \hat{\varphi}_{i,p} \circ \hat{\varphi}_{j,p}^{-1}$ are called *transition functions*.

Often, a fiber bundle is shortly written as the short sequence $F \longrightarrow E \xrightarrow{\pi} B$. If F and G are isomorphic, $F \cong G$, and G acts on F by left translation, the bundle is called a *principal fiber bundle*.

In other words, the similarity between a space E and a product space $B \times F$ is defined using the projection $\pi : E \longrightarrow B$, that in small regions of E behaves just like a projection from corresponding regions of $B \times F$ to B.

Example A.2.2. *(Möbius strip)* According to [3, pp. 126–127] we can represent the Möbius strip as a smooth fiber bundle according to the following profile.

Profile	
Bundle space:	$E =$ Möbius strip
Base space:	$B = S^1 \cong \{(\cos t, \sin t) \mid 0 \leqq t < 2\pi\}$
Fiber:	$F = (-1, 1)$
Covering of B:	$U_1 = \{t \mid -\pi < t < \frac{\pi}{2}\}$, $U_2 = \{t \mid 0 < t < \frac{3\pi}{2}\}$
Transition function:	$f_{12}\big|_{(-\pi, -\frac{\pi}{2})}(t) = -\mathbf{1}_F$, $f_{12}\big|_{(0, \frac{\pi}{2})}(t) = +\mathbf{1}_F$
Structure group:	$G = \{-\mathbf{1}_F, +\mathbf{1}_F\} \cong \mathbb{Z}_2$

Here U_1 and U_2 are two local charts for the sphere S^1 overlapping in two disjoint regions,

$$U_1 \cap U_2 = (-\pi, -\tfrac{\pi}{2}) \cup (0, \tfrac{\pi}{2})$$

The first region then is the set $(-\pi, -\frac{\pi}{2}) \times F$ for which every point on the fiber $\pi^{-1}(t)$ is reflected with respect to the origin, whereas for the second regions $(0, \frac{\pi}{2}) \times F$ every point on the fiber is identified with itself. Thus the structure group is $G = \{-\mathbf{1}_F, +\mathbf{1}_F\} \cong \mathbb{Z}_2$. □

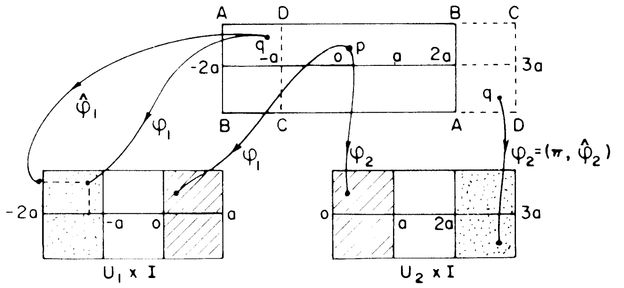


Figure A.3. Identifying scheme to construct a Möbius strip as a fiber bundle. Here $a = \frac{\pi}{2}$, $I = (-1, 1)$, and $\{U_1, U_2\}$ covers S^1. Points with he same label A, B, C, D are identified. Modified from [3, p. 126]

Example A.2.3. *(Tangent bundle)* [3, pp. 127–128] Let $\mathscr{M}$ be a manifold of dimension n. Then let $T\mathscr{M}$ be the space of pairs (p, v_p) of all points $p \in \mathscr{M}$ and all $v_p \in T_p\mathscr{M}$. Then the *tangent bundle* as given in the following profile is a fiber bundle.

Profile of a tangent bundle	
Bundle space:	$E = T\mathscr{M} = \bigcup_{p \in \mathscr{M}} \{(p, v_p) \mid v_p \in T_p\mathscr{M})\}$
Base space:	$B = \mathscr{M}$
Fiber:	$F = \mathbb{R}^n$, $F_p = \pi^{-1}(p) = T_p\mathscr{M}$
Covering of *B*:	$\{U_i \mid \{(U_i, \psi_i)\}$ is an atlas of $\mathscr{M}\}$
Transition function:	$f_{ij} = \psi'_{i,p} \circ \psi'^{-1}_{j,p}$ where $\psi'_{i,p}(v_p)$ is the representative of $v_p \in T_p\mathscr{M}$ in the chart (U_i, ψ_i)
Structure group:	$G = GL(n, \mathbb{R})$

For each fixed $p \in \mathscr{M}$, we have $\{p\} \times T_p\mathscr{M} \subset T\mathscr{M}$. The projection $\pi : T\mathscr{M} \longrightarrow \mathscr{M}$ simply maps $(p, v_p) \mapsto p$, and the fiber at p is $F_p = \pi^{-1}(p) = T_p\mathscr{M} \cong \mathbb{R}^n$, i.e., $F = \mathbb{R}^n$. The covering of $\mathscr{M}$ is given by the chart domains U_i of an arbitrary atlas of $\mathscr{M}$, and their charts ψ_i induce the homeomorphisms φ_i as the pair of the projection π and the mapping $(\pi, \psi' \circ \pi_2)$ where $\pi_2(p, v_p) = v_p$ and $\psi'_i(v_p)$ is the representative of v_p in the chart (U_i, ψ_i), namely

$$\varphi_i = (\pi, \psi' \circ \pi_2) : \pi^{-1}(U_i) \longrightarrow U_i \times \mathbb{R}^n, \qquad (p, v_p) \mapsto (p, \psi'_i(v_p)). \tag{A.4}$$

The fiber coordinates on $T_p\mathscr{M}$ are given by the mappings

$$(\psi_i, \mathrm{id}_{\mathbb{R}^n}) \circ (\psi' \circ \pi_2) : \pi^{-1}(U_i) \to \mathbb{R}^n \times \mathbb{R}^n, . \tag{A.5}$$

The coordinates of a point $y = (p, v_p) \in \pi^{-1}(U_i) \subset T\mathscr{M}$ are thus

$$(x^1, \ldots, x^n, v_p^1, \ldots, v_p^n) \tag{A.6}$$

where $(x^1, \ldots, x^n)$ denote the coordinates of p in $\mathscr{M}$. A change of fiber coordinates on $T\mathscr{M}$ is therefore entirely determined by a change of coordinates on $\mathscr{M}$. Thus the structure group is $GL(n,\mathbb{R})$ of automorphisms of $\mathbb{R}^n$ whose matrix representation is given as the set of $n \times n$ matrices with non-vanishing determinant. The tangent bundle is a $2n$-dimensional real manifold. [15, p. 584] $\square$

Definition A.2.4. A *cross-section* of the fiber bundle $F \to E \xrightarrow{\pi} B$ is a mapping $f : B \to E$ such that $f \circ \pi$ is the identity on B. A *vector field* v is a cross-section on a tangent bundle $T\mathscr{M}$. In other words, a vector field associates to each point $p \in \mathscr{M}$ a tangent vector $v_p \in T_p\mathscr{M}$ by the mapping $v : \mathscr{M} \to T\mathscr{M}$, $p \mapsto (p, v_p)$, often abbreviated $p \mapsto v_p$. A vector field v on a smooth manifold $\mathscr{M}$ is called *smooth* or *differentiable* if the mapping $v : \mathscr{M} \to T\mathscr{M}$ is smooth. [3, p. 132]

Remark A.2.5. A vector field can be regarded as a derivation on the algebra of $C^k(\mathscr{M})$ of functions ofclass C^k on $\mathscr{M}$:

$$v : C^k(\mathscr{M}) \to C^{k-1}(\mathscr{M}) \qquad \text{by} \qquad v(f) = vf. \tag{A.7}$$

Hence $v_p(f) = (vf)(p)$, or in local coordinates (x^i) of $p \in \mathscr{M}$: $(vf)(x^i) = v_p \partial_{x^i}$. $\square$

A.3 Differential forms on smooth manifolds

Definition A.3.1. [21, p. 219] Let $\mathscr{M}$ be a smooth manifold of dimension n. Then a *differential k-form*, or shortly a *k-form* on $\mathscr{M}$ is a mapping that assigns to each $p \in \mathscr{M}$ an alternating multilinear mapping $\omega_p \in \Lambda^k(T_p\mathscr{M})$ [cf. 20, p. 217]. Its dimension is $\binom{n}{k}$. It is called *differentiable* if it is differentiable in local coordinates. The vector space of all differentiable (C^∞) k-forms on $\mathscr{M}$ is denoted by $\Omega^k\mathscr{M}$. By definition functions $f : \mathscr{M} \to \mathbb{R}$ are 0-forms, and thus $\Omega^0\mathscr{M}$ is the space of differentiable functions on $\mathscr{M}$.

A differentiable 1-form $\omega : \Omega^1\mathscr{M}$ is called a *Pfaffian form*. A special kind of Pfaffian forms are the “exact Pfaffian forms” which are differentials of differentiable functions.

Definition A.3.2. [18, p. 60] Let f be a differentiable function $f : \mathscr{M} \to \mathbb{R}$ from a smooth manifold $\mathscr{M}$ to a real number. Then the differentiable 1-form $\mathrm{d}f \in \Omega^1\mathscr{M}$ given by $p \to \Lambda^1 T_p\mathscr{M}$ is called the *differential* of f.

In local coordinates $h = (x^1, \dots, x^n)$ around $p \in U \subset \mathscr{M}$ the differentials $\mathrm{d}x^1, \dots,$ $\mathrm{d}x^n$ form a basis of the space $T_p^*\mathscr{M}$, in fact the dual basis of $\partial_i := \frac{\partial}{\partial x^i} \in T_p\mathscr{M}$. Therefore a general 1-form $\omega \in \Omega^1\mathscr{M}$ on the local chart (U, h) can be expressed as

$$\omega = \sum_{i=1}^{n} \omega_i \, \mathrm{d}x^i \tag{A.8}$$

where $\omega_i : U \to \mathbb{R}$ are the component functions $\omega_i := \omega(\partial_i)$. Especially the differential of a smooth function $f : \mathscr{M} \to \mathbb{R}$ is given by the Leibniz rule

$$\mathrm{d}f = \sum_{i=1}^{n} \partial_i f \, \mathrm{d}x^i \tag{A.9}$$

cf. [18, p. 61].

Definition A.3.3. [3, p. 196, 18, p. 135] Let V be a real vector space. smooth manifold. Then the *exterior product*, also called *wedge product* or *Graßmann product*,

$$\wedge : \Omega^r\mathscr{M} \times \Omega^s\mathscr{M} \to \Omega^{r+s}\mathscr{M}, \qquad (\omega, \eta) \mapsto \omega \wedge \eta \tag{A.10}$$

of differential forms on $\mathscr{M}$ pointwise by $(\omega \wedge \eta) := \omega_p \wedge \eta_p$ for each $p \in \mathscr{M}$.

Definition A.3.4. [18, p. 139] Let $\mathscr{M}$ be a smooth manifold. Then the *exterior product*, also called *wedge product* or *Graßmann product*,

$$\wedge : \Omega^r\mathscr{M} \times \Omega^s\mathscr{M} \to \Omega^{r+s}\mathscr{M}, \qquad (\omega, \eta) \mapsto \omega \wedge \eta \tag{A.11}$$

of differential forms on $\mathscr{M}$ is defined pointwise by $(\omega \wedge \eta) := \omega_p \wedge \eta_p$ for each $p \in \mathscr{M}$. Here

$$\begin{aligned}(\omega_p \wedge \eta_p)&(v_1, \dots, v_{r+s}) \\ &:= \frac{1}{r!s!} \sum_\pi \operatorname{sign}\pi \cdot \omega_p(v_{\pi(1)}, \dots, v_{\pi(r)}) \cdot \eta_p(v_{\pi(r+1)}, \dots, v_{\pi(r+s)})\end{aligned}$$

where $v_i \in T_p\mathscr{M}$ and π is a permutation of $(1, 2, \dots, r+s)$.

Remark A.3.5. The antisymmetry of the wedge product implies $\mathrm{d}x^i \wedge \mathrm{d}x^j = -\mathrm{d}x^j \wedge \mathrm{d}x^i$ as well as $\mathrm{d}x^i \wedge \mathrm{d}x^i = 0$. [39, p. 307] □

Example A.3.6. *(Classical differential forms in 3-dimensional vector analysis)* [18, p. 169] Let $\mathscr{M} \in \mathbb{R}^3$ be open. Then the vector-valued 1-form $\mathrm{d}\boldsymbol{s}$ and the vector-valued 2-form $\mathrm{d}\boldsymbol{S}$,

$$\mathrm{d}\boldsymbol{s} = \begin{pmatrix} \mathrm{d}x^1 \\ \mathrm{d}x^2 \\ \mathrm{d}x^3 \end{pmatrix}, \qquad \mathrm{d}\boldsymbol{S} = \begin{pmatrix} \mathrm{d}x^2 \wedge \mathrm{d}x^3 \\ \mathrm{d}x^3 \wedge \mathrm{d}x^1 \\ \mathrm{d}x^1 \wedge \mathrm{d}x^2 \end{pmatrix} \tag{A.12}$$

are called the *vector-valued line element* $\mathrm{d}\boldsymbol{s} \in \Omega^1\mathscr{M}$ and the *vector-valued surface element* $\mathrm{d}\boldsymbol{S} \in \Omega^2\mathscr{M}$; the real-valued 3-form

$$\mathrm{d}V = \mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \mathrm{d}x^3 \tag{A.13}$$

is called the *volume element* of $\mathscr{M}$. Note that $\mathrm{d}x^i \in \Omega^1\mathscr{M}$, $\mathrm{d}x^i \wedge \mathrm{d}x^j \in \Omega^2\mathscr{M}$, and $\mathrm{d}V \in \Omega^3\mathscr{M}$. Note moreover that the “line element” $\mathrm{d}s = \sqrt{\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2}$, which is applied to compute the arc length of a curve in $\mathbb{R}^3$, is *not* a 1-form. [18, p. 68] □

Example A.3.7. *(Variable transformations)* [39, pp. 307–310] For a coordinate transformation $x = x(u, v)$, $y = y(u, v)$ applied to a 2-form $\omega = a\,\mathrm{d}x \wedge \mathrm{d}y$ with $a = a(x, y)$ we have

$$\omega = (\partial_u x\, \partial_v y - \partial_v x\, \partial_u y)\, a\, \mathrm{d}u \wedge \mathrm{d}v, \tag{A.14}$$

since $\mathrm{d}x = \partial_u x\, \mathrm{d}u + \partial_v x\, \mathrm{d}v$, $\mathrm{d}y = \partial_u y\, \mathrm{d}u + \partial_v y\, \mathrm{d}v$, and thus $\mathrm{d}\omega = (\partial_u x\, \mathrm{d}u + \partial_v x\, \mathrm{d}v) \wedge (\partial_u y\, \mathrm{d}u + \partial_v y\, \mathrm{d}v)$. In terms of the Jacobian determinant

$$\frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \partial_u x & \partial_v x \\ \partial_u y & \partial_v y \end{vmatrix} = \partial_u x\, \partial_v y - \partial_v x\, \partial_u y$$

Equation (A.14) can be rewritten as

$$\omega = \frac{\partial(x, y)}{\partial(u, v)}\, a\, \mathrm{d}u \wedge \mathrm{d}v. \tag{A.15}$$

Similarly, by the coordinate transformation $x = x(u, v, w)$, $y = y(u, v, w)$, $z = z(u, v, w)$ the three-dimensional 3-form $\omega = a\,\mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z$

$$\omega = \frac{\partial(x, y, z)}{\partial(u, v, w)}\, a\, \mathrm{d}u \wedge \mathrm{d}v \wedge \mathrm{d}z. \tag{A.16}$$

with the Jacobian

$$\frac{\partial(x,y,z)}{\partial(u,v,w)} = \begin{vmatrix} \partial_u x & \partial_v x & \partial_w x \\ \partial_u y & \partial_v y & \partial_w y \\ \partial_u z & \partial_v z & \partial_w z \end{vmatrix}.$$

Finally, the variable transformation $x = x(u, v)$, $y = y(u, v), z = z(u, v)$ applied to a 2-form

$$\omega = a\,\mathrm{d}y \wedge \mathrm{d}z + b\,\mathrm{d}z \wedge \mathrm{d}x + c\,\mathrm{d}x \wedge \mathrm{d}y$$

gives the expression

$$\omega = \left(a\,\frac{\partial(y,z)}{\partial(u,v)} + b\,\frac{\partial(z,x)}{\partial(u,v)} + c\,\frac{\partial(x,y)}{\partial(u,v)} \right) \mathrm{d}u \wedge \mathrm{d}v. \tag{A.17}$$

□

Remark A.3.8. By the exterior product $\Omega^* := \bigoplus_0^\infty \Omega^k$ becomes a contravariant functor from the category of manifolds and differentiable mappings into the category of real graduated anticommutative algebras with a unit element. [18, p. 139] □

Theorem A.3.9 (Cartan derivative). *[18, p. 140, 21, p. 246] If $\mathscr{M}$ is a smooth manifold of dimension n, there is only one possibility to introduce a sequence of linear mappings*

$$0 \longrightarrow \Omega^0\mathscr{M} \xrightarrow{\mathrm{d}} \Omega^1\mathscr{M} \xrightarrow{\mathrm{d}} \Omega^2\mathscr{M} \xrightarrow{\mathrm{d}} \cdots \xrightarrow{\mathrm{d}} \Omega^{n-1}\mathscr{M} \xrightarrow{\mathrm{d}} \Omega^n\mathscr{M} \longrightarrow 0 \tag{A.18}$$

such that the following three conditions are satisfied:

(1) Leibniz rule: For $f \in \Omega^0\mathscr{M}$, $\mathrm{d}f \in \Omega^1\mathscr{M}$ is the differential of f.

(2) Poincaré rule: $\mathrm{d}^2 = \mathrm{d} \circ \mathrm{d} = 0$

(3) Product rule: $\mathrm{d}(\omega \wedge \eta) = \mathrm{d}\omega \wedge \eta + (-1)^r \omega \wedge \mathrm{d}\eta$ *for $\omega \in \Omega^r\mathscr{M}$.*

Then $\mathrm{d}\omega$ *is called the* exterior *or* Cartan derivative *of the differential form ω, and the whole sequence* $(A.18)$ *is called the* de Rham complex *of $\mathscr{M}$. We have automatically the property:*

(4) Naturalness: For any differentiable mapping $f : \mathscr{M} \to N$ between smooth manifolds $\mathscr{M}$ and N and all differential forms ω on N we have $\mathrm{d}(f^*\omega) = f^*(\mathrm{d}\omega)$.

Remark A.3.10. For a differential form $\omega = \sum \omega_{j_1 \dots j_r} \mathrm{d}x^{j_1} \wedge \ldots \wedge \mathrm{d}x^{j_r} \in \Omega^r\mathscr{M}$ we have the Cartan rule

$$\mathrm{d}\omega = \sum \mathrm{d}\omega_{j_1 \dots j_r} \wedge \mathrm{d}x^{j_1} \wedge \ldots \wedge \mathrm{d}x^{j_r}. \tag{A.19}$$

[39, p. 307] □

Example A.3.11. [39, p. 307] For $\omega = a\,\mathrm{d}x + b\,\mathrm{d}y$ we have

$$\begin{aligned} \mathrm{d}\omega &= \mathrm{d}a \wedge \mathrm{d}x + \mathrm{d}b \wedge \mathrm{d}y = (\partial_x a\,\mathrm{d}x + \partial_y a\,\mathrm{d}y) \wedge \mathrm{d}x + (\partial_x b\,\mathrm{d}x + \partial_y b\,\mathrm{d}y) \wedge \mathrm{d}y \\ &= (\partial_x b - \partial_y a)\,\mathrm{d}x \wedge \mathrm{d}y, \end{aligned} \tag{A.20}$$

since $\mathrm{d}x \wedge \mathrm{d}x = \mathrm{d}y \wedge \mathrm{d}y = 0$ and $\mathrm{d}y \wedge \mathrm{d}x = -\,\mathrm{d}x \wedge \mathrm{d}y$. □

Example A.3.12. [18, pp. 254–262] Let be $\mathscr{M} \subset \mathbb{R}^3$. Then $\Omega^3\mathscr{M} \cong \Omega^0\mathscr{M} = C^\infty(\mathscr{M}, \mathbb{R}) =: \mathscr{F}(\mathscr{M})$ of real-value differential functions on $\mathscr{M}$ and $\Omega^1\mathscr{M} \cong \Omega^2\mathscr{M} \cong C^\infty(\mathscr{M}, \mathbb{R}^3) =: \mathscr{V}$ of differental vector fields on $\mathscr{M}$. Then the classical differential operators from vector analysis are given by

$$\begin{aligned}
\operatorname{grad} : \mathscr{F}(\mathscr{M}) \to \mathscr{V}(\mathscr{M}), \qquad & f \mapsto \begin{pmatrix} \partial_x f \\ \partial_y f \\ \partial_z f \end{pmatrix}, \\
\operatorname{curl} : \mathscr{V}(\mathscr{M}) \to \mathscr{V}(\mathscr{M}), \qquad & \begin{pmatrix} u \\ v \\ w \end{pmatrix} \mapsto \begin{pmatrix} \partial_y w - \partial_z v \\ \partial_z u - \partial_x w \\ \partial_x y - \partial_y u \end{pmatrix}, \\
\operatorname{div} : \mathscr{V}(\mathscr{M}) \to \mathscr{F}(\mathscr{M}), \qquad & \begin{pmatrix} u \\ v \\ w \end{pmatrix} \mapsto \partial_x u + \partial_y v + \partial_z w,
\end{aligned} \tag{A.21}$$

With the *nabla operator* $\nabla \in \Omega^1\mathscr{M}$,

$$\nabla := \begin{pmatrix} \partial_x \\ \partial_y \\ \partial_z \end{pmatrix}, \tag{A.22}$$

the operators (A.21) can be compactly written as

$$\operatorname{grad} f = \nabla f, \qquad \operatorname{curl} \boldsymbol{v} = \nabla \times \boldsymbol{v} \qquad \operatorname{div} \boldsymbol{v} = \nabla \cdot \boldsymbol{v}. \tag{A.23}$$

□

Definition A.3.13. [18, p. 222, 20, p. 266] Let $\mathscr{M}$ be a smooth n-dimensional manifold whose tangent spaces $T_p\mathscr{M}$ for $p \in \mathscr{M}$ are oriented and have a nondegenerate symmetric bilinear form $\langle \cdot, \cdot \rangle$. (In more general contexts such as semi-Riemannian manifolds and Minkowski space, the bilinear form may not be positive.) Then the *Hodge star operator*

$$\star : \Omega^k\mathscr{M} \xrightarrow{\simeq} \Omega^{n-k}\mathscr{M} \tag{A.24}$$

for differential forms is defined pointwise by $(\star\omega)_p = \star(\omega_p)$ for each $p \in \mathscr{M}$ where $\omega_p \in \Lambda^k(T_p\mathscr{M})$ is a multilinear mapping and $\star : \Lambda^k(T_p\mathscr{M}) \to \Lambda^{n-k}(T_p\mathscr{M})$ is defined by mapping the oriented orthonormal basis $\{e_1, \dots, e_k\} \in \Lambda^k(T_p\mathscr{M})$ to the oriented orthonormal basis $\{e_1, \dots, e_{n-k}\} \in \Lambda^{n-k}(T_p\mathscr{M})$. (Note that $\dim \Lambda^k(T_p\mathscr{M}) = \dim \Lambda^{n-k}(T_p\mathscr{M}) = \binom{n}{k}$.)

Example A.3.14. Let be $\{e_x, e_y, e_z\}$ be the standard basis of the Euclidean space $\mathbb{R}^3$, with the usual orientation, and let be $\{\mathrm{d}x, \mathrm{d}y, \mathrm{d}z\}$ be the dual basis; note that they can be considered as differential 1-forms on $\mathbb{R}^3$. The Hodge star operator $\star : \Omega^1\mathbb{R}^3 \to \Omega^2\mathbb{R}^3$ then implies

$$\star\,\mathrm{d}x = \mathrm{d}y \wedge \mathrm{d}z, \qquad \star\,\mathrm{d}y = \mathrm{d}z \wedge \mathrm{d}x, \qquad \star\,\mathrm{d}z = \mathrm{d}x \wedge \mathrm{d}y. \tag{A.25}$$

□

The Hodge star operator transforms the Cartan derivative d to a *codifferential* $\delta : \Omega^k\mathscr{M} \to \Omega^{k-1}\mathscr{M}$ defined by $\delta = (-1)^k \star^{-1} \mathrm{d}\star$, i.e.,

$$\begin{array}{ccc}
\Omega^k\mathscr{M} & \xrightarrow{\mathrm{d}} & \Omega^{k+1}\mathscr{M} \\
\star\downarrow & & \star\downarrow \\
\Omega^{n-k}\mathscr{M} & \xrightarrow{(-1)^k\delta} & \Omega^{n-k-1}\mathscr{M}
\end{array} \tag{A.26}$$

Therefore each $\Omega^k\mathscr{M}$ is flanked from two Cartan derivatives and two codifferentials:

$$\Omega^{k-1}\mathscr{M} \underset{\delta}{\overset{\mathrm{d}}{\rightleftarrows}} \Omega^k\mathscr{M} \underset{\delta}{\overset{\mathrm{d}}{\rightleftarrows}} \Omega^{k+1}\mathscr{M} \tag{A.27}$$

This defines the *Laplace-Beltrami operator* $\Delta := \mathrm{d}\delta + \delta\mathrm{d} : \Omega^k\mathscr{M} \to \Omega^k\mathscr{M}$. Differential forms satisfying $\Delta\omega = 0$ are called *harmonic*.

The Hodge star operator transforms the de Rham complex (A.18) into a dual complex of decreasing differential form degrees:

$$\begin{tikzcd}
0 \arrow[r] & \Omega^0\mathscr{M} \arrow[r, "\mathrm{d}"] \arrow[d, "\star"'] & \Omega^1\mathscr{M} \arrow[r, "\mathrm{d}"] \arrow[d, "\star"'] & \cdots \arrow[r, "\mathrm{d}"] & \Omega^{n-1}\mathscr{M} \arrow[r, "\mathrm{d}"] \arrow[d, "\star"'] & \Omega^n\mathscr{M} \arrow[r] \arrow[d, "\star"'] & 0 \\
0 \arrow[r] & \Omega^n\mathscr{M} \arrow[r, "\delta"] & \Omega^{n-1}\mathscr{M} \arrow[r, "-\delta"] & \cdots \arrow[r, "(-1)^{n-2}\delta"] & \Omega^1\mathscr{M} \arrow[r, "(-1)^{n-1}\delta"] & \Omega^0\mathscr{M} \arrow[r] & 0
\end{tikzcd} \tag{A.28}$$

Example A.3.15. *(Maxwell equations)* [18, p. 238] Let be $\mathscr{M} \subset \mathbb{R}^3$. Then $\Omega^3\mathscr{M} \cong \Omega^0\mathscr{M} = C^\infty(\mathscr{M}, \mathbb{R}) =: \mathscr{F}(\mathscr{M})$ of real-value differential functions on $\mathscr{M}$ and $\Omega^1\mathscr{M} \cong \Omega^2\mathscr{M} \cong C^\infty(\mathscr{M}, \mathbb{R}^3) =: \mathscr{V}$ of differental vector fields on $\mathscr{M}$. Then with the classical differential operators from vector analysis (A.21) the de Rham complex and its dual (A.26) are now given by

$$0 \rightleftarrows \mathscr{F}(\mathscr{M}) \underset{\mathrm{div}}{\overset{\mathrm{grad}}{\rightleftarrows}} \mathscr{V}(\mathscr{M}) \underset{-\mathrm{curl}}{\overset{\mathrm{curl}}{\rightleftarrows}} \mathscr{V}(\mathscr{M}) \underset{\mathrm{grad}}{\overset{\mathrm{div}}{\rightleftarrows}} \mathscr{F}(\mathscr{M}) \rightleftarrows 0 \tag{A.29}$$

Then the Laplace-Beltrami operator for 0-forms and for 3-forms is given by

$$\Delta = \mathrm{div}\,\mathrm{grad} : \mathscr{F}(\mathscr{M}) \to \mathscr{F}(\mathscr{M}),$$

and for 1-forms and 2-forms by

$$\Delta = \mathrm{grad}\,\mathrm{div} - \mathrm{curl}\,\mathrm{curl} : \mathscr{V}(\mathscr{M}) \to \mathscr{V}(\mathscr{M}).$$

With the nabla operator (A.22) Maxwell's equations are given by

$$\begin{aligned}
\nabla \cdot \boldsymbol{E} &= \frac{\rho}{\varepsilon_0}, & \nabla \times \boldsymbol{E} &= -\dot{\boldsymbol{B}} \\
\nabla \cdot \boldsymbol{B} &= 0, & \nabla \times \boldsymbol{B} &= \mu_0 \left(\boldsymbol{J} + \varepsilon_0 \dot{\boldsymbol{E}}\right)
\end{aligned} \tag{A.30}$$

with $\boldsymbol{E}$ the electric field, $\boldsymbol{B}$ the magnetic field, ρ the electric charge density, and $\boldsymbol{J}$ the current density; here ε_0 is the electric constant and μ_0 the magnetic constant. Note that the speed of light c is then given by

$$c = \sqrt{\frac{1}{\varepsilon_0\mu_0}}. \tag{A.31}$$

Let the Faraday tensor $F \in \Omega^2(\mathbb{R} \times \mathscr{M})$ denote the 2-form in four-dimensional spacetime given in coordinates by $F = F_{ij}\,\mathrm{d}x^i \wedge \mathrm{d}x^j$ with

$$F_{ij} = \begin{pmatrix} 0 & E_x/c & E_y/c & E_z/c \\ -E_x/c & 0 & -B_z & B_y \\ -E_y/c & B_z & 0 & -B_x \\ -E_z/c & -B_y & B_x & 0 \end{pmatrix}, \tag{A.32}$$

where $\boldsymbol{E} = (E_x, E_y, E_z)$ and $\boldsymbol{B} = (B_x, B_y, B_z)$, and the four-current density $j \in \Omega^3(\mathbb{R}\times\mathscr{M})$ the 3-form $j = (c\rho, \boldsymbol{J})$. [18, p. 265, 23, p. 73] Then the Maxwell equations (A.30) can be rewritten as

$$\mathrm{d}F = 0 \qquad \text{and} \qquad \mathrm{d} \star F = j. \tag{A.33}$$

The continuity equation $\mathrm{d}j = \mathrm{d}^2 \star F = 0$ follows automatically by the Cartan calculus, in terms of $\boldsymbol{J}$ and ρ it reads $\operatorname{div}\boldsymbol{J} + c\dot{\rho} = 0$. The skew-symmetrical Faraday tensor can be derived as the differential of a four-potential $A_i = (\phi, -\boldsymbol{A})$ of a scalar potential $\phi \in \Omega^0(\mathbb{R}^4)$ and a vector potential $\boldsymbol{A} \in \Omega^1(\mathbb{R} \times \mathscr{M})$ related to the electric field $\boldsymbol{E}$ and the magnetic field $\boldsymbol{B}$ by

$$\boldsymbol{E} = -\frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t} - \operatorname{grad}\phi, \qquad \boldsymbol{B} = \operatorname{curl}\boldsymbol{A} \tag{A.34}$$

cf. [23, pp. 53, 57, 73] □

Example A.3.16. *(Cotangent bundle)* [3, p. 138] Let $\mathscr{M}$ be a smooth manifold of dimension n. Then let $T^*\mathscr{M}$ be the space of pairs (p, ω_p) of all points $p \in \mathscr{M}$ and all differential forms $v_p \in T_p\mathscr{M}$. Then the *cotangent bundle* as given in the following profile is a fiber bundle.

Profile	
Bundle space:	$E = T^*\mathscr{M} = \bigcup_{p\in\mathscr{M}}\{(p, \omega_p) \mid \omega_p \in T^*_p\mathscr{M})\}$
Base space:	$B = \mathscr{M}$
Fiber:	$F = \mathbb{R}^n$, $F_p = \pi^{-1}(p) = T^*_p\mathscr{M}$
Covering of B:	$\{U_i \mid \{(U_i, \psi_i)\}$ is an atlas of $\mathscr{M}\}$
Transition function:	$f_{ij} = \psi'_{i,p} \circ {\psi'}^{-1}_{j,p}$ where $\psi'_{i,p}(\omega_p)$ is the representative of $\omega_p \in T^*_p\mathscr{M}$ in the chart (U_i, ψ_i)
Structure group:	$G = GL(n, \mathbb{R})$

Analogously to the discussion in Example A.2.3, for each fixed $p \in \mathscr{M}$ we have $\{p\} \times T^*_p\mathscr{M} \subset T^*\mathscr{M}$, the projection $\pi : T^*\mathscr{M} \to \mathscr{M}$ simply maps $(p, \omega_p) \mapsto p$, and the fiber at p is $F_p = \pi^{-1}(p) = T^*_p\mathscr{M} \cong \mathbb{R}^n$, i.e., $F = \mathbb{R}^n$. The coordinates of a point $y = (p, \omega_p) \in \pi^{-1}(U_i) \subset T\mathscr{M}$ are

$$(x^1, \ldots, x^n, \mathrm{d}x^1, \ldots, \mathrm{d}x^n) \tag{A.35}$$

where $(x^1, \ldots, x^n)$ denote the coordinates of p in $\mathscr{M}$. A change of fiber coordinates on $T^*\mathscr{M}$ is therefore entirely determined by a change of coordinates on $\mathscr{M}$. Hence the structure group is $GL(n, \mathbb{R})$. The cotangent bundle is a $2n$-dimensional real manifold. [15, p. 585] □

Definition A.3.17. [3, p. 138] A *covariant vector field* is a cross section $T\mathscr{M} \to \mathscr{M}$ of a cotangent bundle. It is often called a *1-form*. The *reciprocal image of a covariant vector field* $\theta_{f(p)}$ under a differentiable mapping $f : \mathscr{M} \to \mathscr{N}$ between two differentiable manifolds is defined by the numerical equality

$$(f^*\theta)_p v_p = \theta_{f(p)}(f'v)_{f(p)} \tag{A.36}$$

where $f' : T\mathscr{M} \to T\mathscr{N}$ is the differential, defined pointwisely by $f'_p : T_p\mathscr{M} \to T_{f(p)}\mathscr{N}$.

$$\begin{tikzcd} T\mathscr{M} \arrow[r, "f'"] \arrow[d] & T\mathscr{N} \arrow[d] \\ \mathscr{M} \arrow[r, "f"] & \mathscr{N} \end{tikzcd} \qquad \begin{tikzcd} T_p\mathscr{M} \arrow[r, "f'_p"] \arrow[d] & T_{f(p)}\mathscr{N} \arrow[d] \\ \{p\} \arrow[r, "f"] & \{f(p)\} \end{tikzcd}$$

The *reciprocal image of a 1-form* θ under a differentiable mapping $f : \mathscr{M} \to \mathscr{N}$ is defined by the function equality

$$(f^*\theta)v = \theta(f'v) \circ f. \tag{A.37}$$

The mapping $f^*\theta$ is also called the *pullback* of θ by f. In fact, $f^* : T^*\mathscr{N} \to T^*\mathscr{M}$,

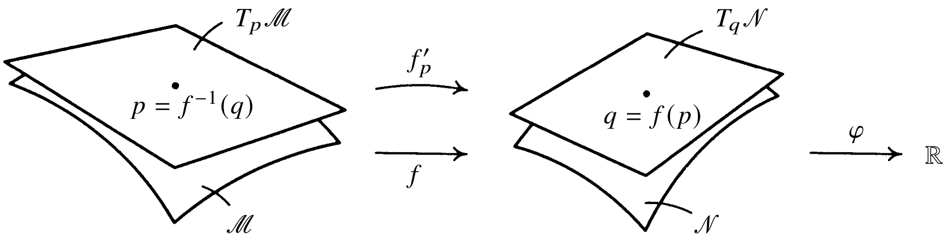
$$\begin{tikzcd} T^*\mathscr{M} \arrow[d] & T^*\mathscr{N} \arrow[l, "f^*"'] \arrow[d] \\ \mathscr{M} \arrow[r, "f"] & \mathscr{N} \end{tikzcd} \qquad \begin{tikzcd} T^*_p\mathscr{M} \arrow[d] & T^*_{f(p)}\mathscr{N} \arrow[l, "f^*_p"'] \arrow[d] \\ \{p\} \arrow[r, "f|_{\{p\}}"] & \{f(p)\} \end{tikzcd}$$

with $f^*_{f(p)} : T^*_{f(p)}\mathscr{N} \to T^*_p\mathscr{M}$ for each $p \in \mathscr{M}$.

Remark A.3.18. For differentiable functions $f : \mathscr{M} \to \mathscr{N}$ and $\varphi : \mathscr{N} \to \mathbb{R}$, i.e., $\varphi \circ f : \mathscr{M} \to \mathbb{R}$, the expression for the image of a vector field reads $(f'v)_{f(p)}(\varphi) = v_p(\varphi \circ f)$, which can be rewritten

$$(f'v)(\varphi) = v(\varphi \circ f) \circ f^{-1} \tag{A.38}$$

if $f : \mathscr{M} \to \mathscr{N}$ is invertible. [3, pp. 121, 134, 18, p. 39]



Expressing the image of a vector field therefore involves f^{-1}. The expression (A.37) of the reciprocal image of a covariant vector field, however, does not involve f^{-1}. In this respect a 1-form is more interesting than a vector field; $f^*\theta$ is always a differentiable 1-form if the mapping f and the 1-form θ are differentiable. [3, p. 138] □

A.4 Semi-Riemannian manifolds

So far we have considered smooth manifolds, i.e., manifolds having with a differential structure. In case of a curved n-dimensional semi-Riemannian or manifold, a "metric" is added, and the de Rham complex (A.18) generalizes to a vector-valued analogon

$$0 \longrightarrow \Omega^0(\mathscr{M}, T\mathscr{M}) \xrightarrow{\mathrm{d}^\nabla} \Omega^1(\mathscr{M}, T\mathscr{M}) \xrightarrow{\mathrm{d}^\nabla} \cdots \xrightarrow{\mathrm{d}^\nabla} \Omega^n(\mathscr{M}, T\mathscr{M}) \longrightarrow 0 \tag{A.39}$$

where now the forms are defined on the tangential bundle $T\mathscr{M}$ of $\mathscr{M}$ and the operator d^∇ depends on the curvature of the manifold and does not satisfy the Poicaré rule in general, i.e., $\mathrm{d}^\nabla \circ \mathrm{d}^\nabla \neq 0$. [21, p. 252] More precisely, the Cartan structure equations in semi-Riemannian manifold yield the relation

$$\mathrm{d}\theta^i = \theta^j \wedge \gamma^i_{kj}\theta^k \tag{A.40}$$

where $\{\theta^i\}$ form a basis of the cotangential bundle $T^*\mathscr{M}$ (???) and γ^i_{kj} are the connection coefficients given by the unique torsion-free linear connection ∇ through $\nabla e_i = \gamma^j_{kl}\theta^k \otimes e_j$ where $\{e_i\}$ and $\{\theta^i\}$ are dual frames. [3, pp. 301, 310]

Definition A.4.1. Let $\mathscr{M}$ be a semi-Riemannian oriented manifold of dimension n with the metric tensor g_{ij} given in the local coordinates $(x^1, \ldots, x^n)$. Then the *canonical volume form* $\omega_{\mathscr{M}} \in \Omega^n\mathscr{M}$ is defined as

$$\omega_{\mathscr{M}} = \sqrt{|g|}\, \mathrm{d}x^1 \wedge \ldots \wedge \mathrm{d}x^n \tag{A.41}$$

where g is the determinant of the metric tensor. Moreover, the *star operator* $\star : \Omega^k\mathscr{M} \to \Omega^{n-k}\mathscr{M}$ is given by

$$(*\zeta)_{\tau_{k+1}\ldots\tau n} = \operatorname{sgn}\tau \cdot \sqrt{|g|}\, \zeta^{\tau_1\ldots\tau_k} \tag{A.42}$$

in local coordinates preserving the orientation, where τ is a permutation of $\{1, \ldots, n\}$. [18, pp. 253–254]

The coderivative $\delta : \Omega^1\mathscr{M} \to \Omega^0\mathscr{M}$ is given in local coordinates by

$$\delta\alpha = \alpha^i\partial_i = \frac{1}{\sqrt{|g|}}\partial_i\left(\sqrt{|g|}\alpha^i\right). \tag{A.43}$$

The function $\delta\alpha$ the is called the *divergence* of the vector field $\alpha^i\partial_i$. Hence for functions or 0-forms the Laplace-Operator $\Delta : \Omega^0\mathscr{M} \to \Omega^0\mathscr{M}$ is defined by $\delta\mathrm{d}$, i.e.,

$$\Delta f = \frac{1}{\sqrt{|g|}}\partial_i\left(\sqrt{|g|}\partial^i f\right) = \frac{1}{\sqrt{|g|}}\sum_{i,j=1}^{n}\frac{\partial}{\partial x^i}\left(\sqrt{|g|}g^{ij}\frac{\partial}{\partial x^j}f\right). \tag{A.44}$$

Example A.4.2. Let $\mathscr{M} = S^2 \subset \mathbb{R}^3$ be the sphere with spherical coordinates $(\vartheta, \varphi) \in (0, \pi) \times [0, 2\pi]$ as in Example A.1.10. Then the metric tensor induced by the Euclidean space $\mathbb{R}^3$ in these coordinates is given by

$$(g_{ij}) = \begin{pmatrix} 1 & 0 \\ 0 & \sin^2\vartheta \end{pmatrix}, \qquad (g^{ij}) = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{\sin^2\vartheta} \end{pmatrix}. \tag{A.45}$$

Therefore [18, p. 256]

$$\Delta_{S^2} = \frac{1}{\sin\vartheta}\frac{\partial}{\partial\vartheta}\left(\sin\vartheta\frac{\partial}{\partial\vartheta}\right) + \frac{1}{\sin^2\vartheta}\frac{\partial^2}{\partial^2\varphi}. \tag{A.46}$$

□

Bibliography

[1] M. Born. *Die Relativitätstheorie Einsteins. Kommentiert und erweitert von Jürgen Ehlers und Markus Pössel*. Berlin Heidelberg: Springer, 2001. URL: https://books.google.com/books?id=xR4fBgAAQBAJ.

[2] S. Chandrasekhar. *The Mathematical Theory of Black Holes*. Oxford: Oxford University Press, 1983. URL: http://books.google.com/books?id=LBOVcrzFfhsC.

[3] Y. Choquet-Bruhat, C. de Witt-Morette, and M. Dillard-Bleick. *Analysis, Manifolds and Physics*. Amsterdam: North-Holland Publ. Co., 1982.

[4] W. de Sitter. "Einstein's theory of gravitation and its astronomical consequences. Third paper". In: *Month. Not. Roy. Astr. Soc.* 78 (Nov. 1917), pp. 3–28.

[5] W. de Sitter. "On the curvature of space". In: *Proc. Kon. Ned. Akad. Wet.* 20 (June 30, 1917), pp. 229–243.

[6] W. de Sitter. "On the relativity of inertia. Remarks concerning Einstein's latest hypothesis". In: *Proc. Kon. Ned. Acad. Wet.* 19.2 (Mar. 31, 1917), pp. 1217–1225.

[7] A. de Vries. *Über die Beschränktheit der Energienorm bei der Evolution der Dirac-, Weyl- und Maxwellfelder in gekrümmten Raumzeiten*. Bochum: Brockmeyer, 1994. URL: http://math-it.org/Publikationen/diss.pdf.

[8] A. de Vries. "The evolution of the Weyl and Maxwell fields in curved space-times". In: *Math. Nachr.* 179 (1996), pp. 27–45. DOI: 10.1002/mana.19961790103.

[9] A. Einstein. "Die formale Grundlage der allgemeinen Relativitätstheorie". In: *Sitzungsber. Königl. Preuß. Akad. Wiss., phys.-math. Klasse* II (1914), pp. 1030–1085. URL: http://adsabs.harvard.edu/abs/1914SPAW......1030E.

[10] A. Einstein. "Die Feldgleichungen der Gravitation". In: *Sitzungsber. Königl. Preuß. Akad. Wiss., phys.-math. Klasse* II (Nov. 25, 1915), pp. 844–847. URL: http://adsabs.harvard.edu/abs/1915SPAW.......844E.

[11] A. Einstein. "Erklärung der Perihelbewegung des Merkur aus der allgemeinen Relativitätstheorie". In: *Sitzungsber. Königl. Preuß. Akad. Wiss., phys.-math. Klasse* II (Nov. 18, 1915), pp. 831–839. URL: http://adsabs.harvard.edu/abs/1915SPAW.......831E.

[12] A. Einstein. "Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie". In: *Sitzungsber. Königl. Preuß. Akad. Wiss., phys.-math. Klasse* (Feb. 8, 1917), pp. 844–847. URL: https://ui.adsabs.harvard.edu/abs/1917SPAW.......142E.

[13] A. Fölsing. *Albert Einstein. Eine Biographie*. Frankfurt: Suhrkamp, 1993.

[14] A. Friedman. "Über die Krümmung des Raumes". In: *Z. Physik* 10 (1922), pp. 377–386. DOI: 10.1007/BF01332580.

[15] G. Grosche, E. Zeidler, and D. Ziegler, eds. *Teubner Taschenbuch der Mathematik. Teil 2*. Leipzig: B. G. Teubner, 1995.

[16] S. W. Hawking and G. F. R. Ellis. *The Large Scale Structure of Space-time*. Cambridge: Cambridge University Press, 1973.

[17] E. Hubble. “A relation between distance and radial velocity among extra-galactic nebulae”. In: *PNAS* 15.3 (1929), pp. 168–173. DOI: 10.1073/pnas.15.3.168.

[18] K. Jänich. *Vektoranalysis*. Berlin Heidelberg: Springer-Verlag, 1992.

[19] K. Jänich. *Topologie*. 6th ed. Berlin Heidelberg: Springer-Verlag, 1999.

[20] K. Jänich. *Mathematik 1. Geschrieben für Physiker*. Berlin Heidelberg: Springer-Verlag, 2001.

[21] K. Jänich. *Mathematik 2. Geschrieben für Physiker*. Berlin Heidelberg: Springer-Verlag, 2002.

[22] H. Karttunen, P. Kröger, H. Oja, M. Poutanen, and J. Donner. *Fundamental Astronomy*. 2nd ed. Berlin Heidelberg New York: Springer Verlag, 2000.

[23] L. D. Landau and E. M. Lifschitz. *Klassische Feldtheorie*. 12th ed. Frankfurt: Harri Deutsch, 1997.

[24] G. Lemaître. “Un Univers homogène de masse constante et de rayon croissant rendant compte de la vitesse radiale des nébuleuses extra-galactiques”. In: *Annales de la Société Scientifique de Bruxelles* A47 (1927), pp. 49–59. URL: http://adsabs.harvard.edu/abs/1927ASSB...47...49L.

[25] B. O’Neill. *Semi-Riemannian Geometry. With Applications to Relativity*. San Diego: Academic Press, 1983.

[26] H. W. M. Olbers. “Ueber die Durchsichtigkeit des Weltraumes”. In: *Astronomisches Jahrbuch* (1826). URL: https://archive.org/stream/wilhelmolberssei00olbeuoft#page/132/mode/2up.

[27] J.-F. Pascual-Sánchez. “Introducing relativity in global navigation satellite systems”. In: *Annalen der Physik* 16.4 (2007), pp. 258–273. ISSN: 1521-3889. DOI: 10.1002/andp.200610229.

[28] R. Penrose. *The Road to Reality*. New York: Vintage Books, 2004.

[29] R. Penrose and W. Rindler. *Spinors and Space-Time. Vol. 2: Spinor and Twistor Methods in Space-Time Geometry*. Cambridge: Cambridge University Press, 1986.

[30] F. Scheck. *Theoretische Physik 3. Klassische Feldtheorie*. 4th ed. Berlin: Springer Spektrum, 2017.

[31] R. U. Sexl and H. K. Urbantke. *Gravitation und Kosmologie. Eine Einführung in die Allgemeine Relativitätstheorie*. Mannheim Wien Zürich: BI Wissenschaftsverlag, 1987.

[32] S. F. Singer. “Application of an Artificial Satellite to the Measurement of the General Relativistic “Red Shift” ”. In: *Phys. Rev.* 104 (Oct. 1956), pp. 11–14. DOI: 10.1103/PhysRev.104.11.

[33] B. Sonne and R. Weiß. *Einsteins Theorien. Spezielle und Allgemeine Relativitätstheorie für Einsteiger und zur Wiederholung*. Berlin Heidelberg: Springer Spektrum, 2013.

[34] H. Stephani. *Allgemeine Relativitätstheorie*. Berlin: Deutscher Verlag der Wissenschaften, 1991.

[35] N. Straumann. *Allgemeine Relativitätstheorie und relativistische Astrophysik*. Berlin: Springer-Verlag, 1988.

[36] F. J. Tipler. “Johann Mädler’s Resolution of Olbers’ Paradox”. In: *Quat. J. Roy. Astr. Soc.* 29 (1988), pp. 313–325. URL: `http://adsabs.harvard.edu/abs/1988QJRAS..29..313T`.

[37] A. Unsöld and B. Baschek. *Der neue Kosmos. Einführung in die Astronomie und Astrophysik*. 6th ed. Berlin Heidelberg New York: Springer Verlag, 1999.

[38] A. Unsöld and B. Baschek. *Der neue Kosmos. Einführung in die Astronomie und Astrophysik*. 7th ed. Berlin Heidelberg New York: Springer Verlag, 2002.

[39] E. Zeidler, ed. *Teubner Taschenbuch der Mathematik. Teil 1*. Leipzig: B. G. Teubner, 1996.