Physics Before and After Einstein
M. Mamone Capria (Ed.)
IOS Press, 2005

Chapter 5

General Relativity: Gravitation as Geometry and the Machian Programme

Marco Mamone Capria

> To put it in another way, if only a relative meaning can be attached to the concept of velocity, ought we nevertheless to persevere in treating acceleration as an absolute concept?
>
> *Albert Einstein, 1934*

The path that led to the birth of general relativity is one of the most fascinating sequences of events in the whole history of theoretical physics. It was a tortuous path, to be sure, not without a fair amount of wandering and misunderstanding. The outcome was a theory with a curious destiny. On one hand it elicited the admiration of generations of theoretical physicists and provided the ground for several observational programs in astronomy and astrophysics up to the present day. On the other hand, the theory was judged inadequate by its very creator, who worked to the end of his life to devise a radical improvement of it, which eventually he did not achieve to his own satisfaction.

The very name of the theory contains a degree of ambiguity: 'general relativity' can be read as a shortening of either 'general theory of relativity – meaning a generalization of the previous *theory* of relativity – or 'theory of general relativity – meaning something more specific, i.e. a generalization of the *principle* of relativity. The latter is what Einstein had in mind. In fact his primary aim was to generalize the principle of relativity which he had introduced in his 1905 theory, the validity of which was restricted to inertial reference frames. Whether he succeeded or not in this endeavor and – more importantly – whether he *could* possibly succeed is just one of the main issues of a foundational debate on the theory which began when the theory was still far from completion, and which has not yet come to an end. This is not surprising, given the depth of the questions which found a provisional accommodation in the Einsteinian synthesis: the nature of space, time, motion, matter, energy.

Was there a need to 'generalize' the original theory of relativity? According to some of the most eminent physicists of the time, there was not. Max Planck himself, whose approval had been crucial in earning recognition for the special theory, warned Einstein against wasting time in further speculations on that thorny topic. A different question is: why did Einstein insist on searching for a theory which had to supersede special relativity?

Historians and textbook authors have often emphasized, rather too hastily, that special relativity was 'incompatible with gravitation', and that this is the very reason Einstein was looking for a more general theory. This view is not correct, however, the clearest evidence to the contrary being that in his major paper on relativity, in 1906, Henri

Poincaré had already offered not just one, but several alternatives to translate Newton's attraction law into a Lorentz-invariant expression, that is, into a special-relativistic law. He did not think that his proposals lacked physical reasonableness, but he did not feel competent either, as a theoretician, to decide between them; for this he soberly referred to future astronomical observations (the great conventionalist was in fact more inclined to submit to experiment than historiographical folklore portrays him). Poincaré was not alone in this attempt: in the following decades many other gravitational theories, even after the rise of general relativity, were advanced.[1]

What one can surmise, in the face of this preliminary evidence, is that theories of gravitation which were compatible with special relativity might have been in some sense *not enough*. As is so often the case in the history of science, the only safe way to know why a certain theory has been created is to pursue as far as possible the real path followed by its creators; and in this case, much more than for special relativity, the theory was the creation of one man (though by no means unaided).[2]

1. Absolute Space, Relative Motion and Relativity

The 1905 article introducing the special theory of relativity is remarkable for some of the things it contains, but it is rarely considered remarkable, as it should be, for one thing it does *not* contain: gravitation. In fact it should be *prima facie* surprising that a young physicist upturning the very foundations of physical science simply does not pay any attention to gravitation – there is no announcement of work in progress, not even an incidental remark. Even more surprising is that he can neglect such a crucial topic and still get away with it, that is, without his paper being rejected by such a prestigious journal as *Annalen der Physik*. The basic explanation for this curious circumstance lies in the fact that at the end of the nineteenth century mechanics and gravitation had lost their primacy, and a new view of the physical world, based on electromagnetism, had taken their place. It was widely thought that gravitation would have been explained in terms of electromagnetism rather than the reverse. So the received wisdom of the epoch considered it acceptable to redefine space and time in order to best accommodate electrodynamics. Nevertheless, sooner or later, a viable new theory of gravitation just *had* to be produced.

The reasons for Einstein's pursuit were mainly theoretical in nature, although also based on empirical indications. He was deeply dissatisfied with the absoluteness of the *space–time* structure in special relativity, which did not seem to him, and was not in fact, a big philosophical improvement with respect to classical physics with its absolute *space* and *time*. In his 1921 lectures at the Institute for Advanced Study at Princeton he was quite explicit on this:

> Just as it was consistent from the Newtonian standpoint to make both the statements, *tempus est absolutum, spatium est absolutum*, so from the standpoint of the special theory of relativity we must say, *continuum spatii et temporis est absolutum*. In this latter statement *absolutum* means not only 'physically real', but also 'independent in its physical properties, having a physical effect, but not itself influenced by physical conditions' [38, p. 55].

[1] Chapter 7 in [76] bears the title "Incompatibility of gravity and special relativity" and describes in which sense one might say that gravitational theories based on Minkowski space–time are not satisfactory; see [106] for a survey of the many alternatives which have been put forward up to 1965.

[2] Whenever convenient, in this chapter and in the following I use the translations of Einstein's writings contained in [49] and [50].

It has been disputed many times to what extent the views of philosopher–scientist Ernst Mach were important in the building of the special theory.[3] What is beyond dispute is Mach's influence in the birth of the general theory. Mach's criticism of the concepts of absolute space and absolute motion in his *Science of Mechanics* [72], a book first published in 1886 which Einstein read carefully and discussed with his two friends of the "Olympia Academy", was the starting point of Einstein's line of thought.

Particularly decisive was Mach's reply to the "bucket argument", put forward by Newton in order to demonstrate that absolute acceleration (that is, acceleration with respect to the absolute space), as opposed to absolute uniform motion, can be physically recognized. Newton had remarked that the surface of the water contained in a bucket becomes curved inwards if and only if the water rotates with respect to an inertial system – it is not enough if there is just a relative rotation between the water and the bucket. Mach replied that what Newton considered 'absolute' rotation was in fact rotation with respect to matter, though admittedly not the matter of the bucket itself, but the *huge* amount of astronomical matter which exerted a centrifugal force on the water. Mach's powerful answer implied that the whole material universe and the relative motions of its different parts, rather than 'absolute' space and motions, are responsible for the forces which in classical mechanics are called 'inertial' (or, even more dismissively, 'apparent').

It is clear that a theory in which motion is truly relative should not single out *a priori* a class of privileged systems, such as the inertial systems of Newtonian physics or special relativity. The inertial systems are all relatively at rest or moving with constant relative velocity, and any system having a constant velocity with respect to any of them is itself inertial. Now suppose – as is in fact the case – that the laws of physics according to special relativity (for example) can be expressed by certain formulas *which are the same within that class*; this would provide a sufficient condition to decide whether a given system S is accelerated with respect to one (and therefore to all) of the inertial systems. To do that, one would have to check experimentally whether or not certain phenomena, as described in S, satisfy those formulas. If they don't, then S is accelerating with respect to the inertial frames, and so it certainly *moves* with respect to them.

This concept can be aptly illustrated by Galileo's famous ship experiment, vividly described in his *Dialogue Concerning the Two Chief World Systems – Ptolemaic and Copernican*, of 1632. If we are below deck, can we decide whether the ship is moving with respect to the shore without having to look (or otherwise communicating with) outside? Can we settle this question by just performing experiments inside? The answer, as Galileo (and Giordano Bruno before him) claimed, is negative if the ship is moving with constant velocity – which of course normally requires that the sea or the river is calm. On the other hand, during a tempest (or if for whatever reason the ship is gaining or losing speed), passengers below deck can usually establish quite easily whether the ship is moving: they may point, for instance, to an hanging rope or a lamp which 'no one has touched' and which has begun to oscillate, and this is sufficient proof that the ship is moving. There is no need to 'look outside' to be certain of this fact.

Thus, if we wish to rule out such a possibility, we need to formulate laws which maintain the same form not only in inertial systems, but also in accelerated ones. But is this a reasonable programme? As we have just shown, by testing the validity of certain

[3] One of the reasons for this perplexity is that Mach did not accept special relativity, of which he wrote in 1913 that he did not want to be considered a "forerunner" and which in his opinion was "growing more and more dogmatical".

laws in a system we *can* infer whether we are moving or not with respect to the inertial systems. The only possible answer to this objection from a Machian point of view is to concede that one can indeed have laws which are valid only for selected systems, and that therefore can be used to detect one's state of motion with respect to them, but that this fact does not rule out the possibility of finding other, more fundamental laws which are just the same in all systems.

Einstein endorsed this viewpoint, and embodied it in his *principle of general covariance* (which he also called "Principle of Relativity" in 1918 [33]), and that in the outline of general relativity he published in 1916 is stated as follows:

> *The general laws of nature are to be expressed by equations which hold good for all systems of coordinates, that is, are covariant with respect to any substitutions whatever (generally covariant)* [31, p. 117].

We shall discuss in due course what this formulation has to do with a 'principle of relativity'. However, before arriving at the confidence expressed in this statement, Einstein had to travel a long way. He started studying elementary accelerated reference frames: a freely falling chest and a rotating disk. In fact the simplicity of these examples was to a large extent deceptive.

2. The Principle of Equivalence

The concept of an absolute space–time continuum is unavoidable insofar as one wishes to maintain the ordinary law of inertia. But there are two reasons, according to Einstein [38], to abandon this law, at least as usually construed.

The first reason, for which Einstein refers to Mach's criticism of classical mechanics, is that such an absolute entity would enjoy the property, which runs "contrary to the mode of thinking in science", of acting without being acted upon by anything. One might reformulate this objection by saying that in classical physics the action and reaction principle does not hold in noninertial systems (cf. quotation in §1).

The other reason is the numerical equality, experimentally established with "very high accuracy" by Roland von Eötvös, between *inertial mass* and *gravitational mass*, which still remain conceptually separate in classical physics.

The latter point had arisen early in Einstein's thought; it was first introduced in a survey of special relativity published in 1907 [22] and then presented in a paper published in 1911, which contains the basic physical insights of general relativity [23].

The equality of the two kinds of mass (after suitable choice of the units) follows from Newton's force law and Galileo's law of free fall which, put in a general form, states that all bodies in a homogeneous gravitational field have the same constant acceleration. Using a torsion balance Eötvös had checked in 1896 a simple consequence of this equality, namely, that in a reference frame attached to the surface of a rotating homogenous sphere (modelling the Earth) the total (i.e. gravitational plus inertial) force acting on a body must have one and the same direction, independent of the mass and constitution of the body. In other terms, any difference between the two kinds of mass would result in different directions for the 'plumb lines' determined by using differently constituted bodies. Eötvös's experiments gave a negative outcome: no difference could be detected (cf. Chap. 7). The rôle of his experiments with respect to general relativity was compared by Einstein to that played by Michelson's experiments for special relativity [24, Sect. 1].

Assuming that the two kinds of mass are equal means that, for gravitational phenomena, a uniformly accelerated reference system, far from all bodies, cannot be distinguished from one at rest in a homogeneous gravitational field. This was illustrated by the ideal experiment of "a spacious chest resembling a room with an observer inside who is equipped with apparatus" [40, p. 68]: such an observer cannot distinguish, by making mechanical experiments inside this container, whether it is moving with uniform acceleration or it is at rest but in a constant gravitational field (like the gravity field near the Earth's surface). If this is true, then a freely falling chest in a constant gravitational field *will behave as an inertial reference frame*, at least as far as mechanical phenomena are concerned. This idea is what Einstein called, in an unpublished manuscript of 1920, "the happiest thought of my life". From the Kyoto address we learn that 'Einstein's apple' fell down "all of a sudden" when he was "sitting in a chair in the patent office at Bern" and the following remark occurred to him: "If a person falls freely, he should not feel his weight himself" [1, p. 15].

Einstein proposed that the impossibility of distinguishing experimentally between these two types of coordinate systems should be raised to a fundamental principle holding *for all physical phenomena*, not just the mechanical ones.

In a theory where this *principle of equivalence* (as Einstein called it) holds, the inertial and gravitational masses are not just *empirically* equal: they are *identical* – the very same physical property as measured by different procedures. We recognize in this move a close methodological resemblance to the principle of special relativity. What experience suggested in 1905 was that the laws of electromagnetism do not differ appreciably across different inertial frames: and Einstein (and Poincaré before him) assumed *that these laws do not differ at all*. The importance of the asserted identity of the two masses was stressed by Einstein in unambiguous terms:

> The possibility of explaining the numerical equality of inertia and gravitation by the unity of their nature gives to the general theory of relativity, according to my conviction, such a superiority over the conceptions of classical mechanics, that all the difficulties encountered must be considered as small in comparison with this progress [38, p. 58].

The 'explanation' mentioned here is of the same kind as the 'explanation' of the Michelson–Morley and other optical ether-drift experiments by the constancy postulate plus the principle of relativity. One might say as well that if these postulates are valid, then *there is nothing to explain* in those experimental results, except for their deviating from a perfectly 'null' outcome.

3. Two Early Empirical Consequences

Already in his 1907 survey [22] Einstein derived two of the three effects that were to provide the main testing ground for general relativity during the next fifty years (cf. Chap. 7): the gravitational redshift and the deflection of light rays.

If system S is uniformly accelerated, its natural time coordinate τ is linked to a suitable inertial time t by the following approximate formula:

$$t = \tau\left(1 + \frac{gh}{c^2}\right),$$

where h is the height in the direction of the acceleration. Now the principle of equivalence enables us to infer that in an inertial system embedded in a uniform gravitational field the same formula should be applied, the term gh being interpreted as the gravitational potential ϕ:

$$t = \tau\left(1 + \frac{\Phi}{c^2}\right).$$

This result implies that a clock at rest in a place with a higher potential goes *faster* than a clock at the origin (which is the zero level in this case).

If in the last statement we traslate 'clock' into 'every physical process', we can apply it to the "generators of the spectral lines". Thus, assuming that the same argument can also be applied, to some extent, to inhomogeneous gravitational fields, Einstein concluded that the radiation coming from material on the surface of the Sun and other heavenly bodies must arrive to us redshifted (i.e. with a lower frequency) with respect to radiation produced by similar material on the Earth.

From the equivalence principle Einstein derived in 1907 a second consequence which was at variance with one of the postulates of the special theory. In a gravitational field the speed of light is *not* constant, but it depends on the place according to the law:

$$c' = c_0\left(1 + \frac{\Phi}{c^2}\right),$$

where c_0 is the speed of light at the origin. Formally, this is tantamount to interpreting the gravitational field as a medium with varying refraction index, which implies that the paths of the light-rays are in general not rectilinear.

In a paper of 1911 ("On the influence of gravitation on the propagation of light" [23]) Einstein came back to these results, and by using Huyghens's principle derived, for light passing close to an heavenly body, the following approximate formula for the total deflection angle:

$$\alpha \approx \frac{2GM}{c^2\Delta}, \tag{1}$$

where Δ is the distance between the ray and the centre of the body. In the case of the Sun he obtained the following estimate: $\alpha = 4 \cdot 10^{-6} = 0.83$ seconds of arc. He ended this paper by inviting astronomers to check "apart from any theory [...] whether it is possible with the equipment at present available to detect an influence of gravitational fields on the propagation of light". In fact Einstein had to correct this prediction a few years later, when the theory reached its final form.

4. Special Relativity as a Geometric Theory

Before moving on, it is important to mention a circumstance that turned out decisive in directing Einstein's efforts towards a geometric theory of gravitation. A geometric formulation of special relativity had been developed a couple of years after the *Annalen* article by Hermann Minkowski [75], a former teacher of Einstein at the Zürich Polytechnic. In fact Minkowski followed in the steps of Henri Poincaré, who had anticipated the basic idea in his paper of 1906 [84], but he studiously omitted to mention the great French scientist.

Euclidean geometry can be studied, formally speaking, with any number of dimensions: what makes dimensions 2 and 3 so special is that they allow for a simple physical interpretation of the axioms and theorems, and therefore for a natural visualization. The axiomatic presentation of Euclidean geometry as can be found in classical textbooks is somewhat cumbersome. In fact it is equivalent to a more logically transparent (if less intuitive) presentation putting at centre-stage the axioms of the set of real numbers $\mathbb{R}$ and the assumption of the existence of an inner product[4] on the space of translations of $\mathbb{R}^n$ (the space of n-tuples of real numbers). The main difference between Euclidean and Minkowskian 4-dimensional geometry is that in the former the inner product of two vectors $u = (u^1, u^2, u^3, u^4)$ and $v = (v^1, v^2, v^3, v^4)$ has the appearance:

$$\langle u, v\rangle = u^1v^1 + u^2v^2 + u^3v^3 + u^4v^4,$$

which when $u = v$ gives the well-known Pythagorical form:

$$\langle u, u\rangle = \left(u^1\right)^2 + \left(u^2\right)^2 + \left(u^3\right)^2 + \left(u^4\right)^2,$$

while in Minkowskian geometry the inner product is

$$g(u, v) = -u^1v^1 - u^2v^2 - u^3v^3 + c^2u^4v^4, \tag{2}$$

where c is the speed of light in empty space (or, formally, any positive constant). An inner product with this *signature* $(-, -, -, +)$ is called *Lorentzian*.[5] The Minkowski inner product can also be written as

$$g(u, v) = \sum_{\mu,\nu} \eta_{\mu\nu}u^\mu v^\nu, \tag{3}$$

where:

$$(\eta_{\mu\nu}) = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & c^2 \end{pmatrix}. \tag{4}$$

An immediate consequence of (2) is that $g(u, u)$ can be positive, negative or zero – and when it is zero this does *not* imply that u is the zero vector, contrary to Euclidean intuition. In fact one can classify all nonzero vectors according to whether $g(u, u) > 0$ (these vectors are called *timelike*), $g(u, u) < 0$ (they are called *spacelike*), or $g(u, u) = 0$ (they are called *ligthlike* or *null*).[6] The timelike vectors fill two solid cones, symmetric with respect to the zero vector, and the lightlike vectors together with the zero vector

[4] An inner product is a non-degenerate symmetric bilinear form on a real vector space. Details on the concepts of this section can be found in textbooks on linear algebra, e.g. [68].

[5] Several authors (e.g. Pauli [82]) use the signature $(+, +, +, -)$, also called Lorentzian. Here the opposite convention is adopted throughout to conform to Einstein's usage. Of course in the passage from one to the other convention many definitions have to be slightly modified.

[6] The zero vector is decreed to be spacelike.

form the boundary of the set of all timelike vectors. The set of all lightlike vectors is called the *lightcone*.

A curve whose 4-dimensional velocity vector (or *4-velocity*) is always either timelike or lightlike is called a *worldline*. It represents any ordinary physical process, and the condition on the 4-velocity means that its ordinary 3-dimensional velocity, *according to every inertial system*, does not exceed the speed of light. Suppose the timelike worldline Γ from p to q is parametrized as $x^\mu = x^\mu(s)$, with $\mu = 1, 2, 3, 4$ and s in $[a_1, a_2]$; then its Lorentzian 'length' is

$$T = \int_{a_1}^{a_2} \sqrt{\sum_{\mu,\nu} \eta_{\mu\nu} \frac{dx^\mu}{ds} \frac{dx^\nu}{ds}} ds, \tag{5}$$

and is called the *proper time* between p and q along Γ; the physical interpretation of this quantity is that T represents the time measured by a clock whose worldline is Γ. As expected, the timelike line segments are critical points of the proper-time functional, symbolically:

$$\delta\left(\int_p^q d\tau\right) = 0,$$

but they do not minimize proper time: they *maximize* it. This is one important point where Lorentzian geometry is sharply at variance with Euclidean geometry.

The fact that all physical processes known to us are endowed with a specific, objective time-direction is represented geometrically by selecting one of the timelike cones – the *future* cone – together with the corresponding lightlike vectors, and prescribing that the correct time order along every worldline is that defined by any parametrization with 4-velocity vector always in the future cone. This means that a *time orientation* is fixed.

Thus special relativity can be reformulated as the theory according to which the space of all events – physical space–time – is represented by a *Minkowski space*, i.e. an affine space of dimension 4, with an inner product on its associated vector space, and a time orientation; a space–time orientation is usually also added to the definition.

5. The Rotating Disk and the Physical Meaning of Coordinates

As to the effects of rotation, Einstein discussed in several papers, starting in 1912, the case of an uniformly rotating circular platform of radius R [22]. He had begun to think about it much earlier, as documented by a letter to Sommerfeld of 1909 [97].

What is the length L of the rim of this disk according to the rotating observer, as compared with the length L' measured by an inertial observer? Clearly $L' = 2\pi R$, since it is assumed that the platform is circular according to the latter observer. On the other hand, if the speed of a point on the rim with respect to the centre is v, the length of a rotating element of arc ds should suffer a Lorentz contraction according to the inertial observer, who will measure it as:

$$ds' = \sqrt{1 - \frac{v^2}{c^2}} ds.$$

Thus, by integrating all over the rim, one obtains

$$2\pi R = L' = \sqrt{1 - \frac{v^2}{c^2}}\, L,$$

that is, the length L of the rim of the disk according to the rotating observer is *bigger* than $2\pi R$. Since the radius of the disk suffers no contraction (because it is orthogonal to the velocity), its length must be R even according to the rotating observer; it follows that for this observer the ratio of the circumference to the radius is more than 2π: in other words, a basic theorem of Euclidean geometry fails.

Similarly, because of time dilatation, a clock on the rim of the disk will appear to run slower to an observer at the centre, and therefore one cannot assume that all clocks on the platform have the same rate, independent of position (as was assumed on special relativity).

This argument was taken by Einstein to imply that the coordinates in a noninertial frame cannot be interpreted as directly associated with measurement operations:

> In the general theory of relativity, space and time cannot be defined in such a way that differences of the spatial coordinates can be directly measured by the unit measuring-rod, or differences in the time coordinate by a standard clock [31, p. 117].

In his Princeton lectures Einstein used the example of the disk together with the principle of equivalence to infer that, as rotation can produce a non-Euclidean physical space, so also "the gravitational field influences and even determines the metrical laws of the space–time continuum" and that "in the presence of a gravitational field the geometry [of ideal rigid bodies] is not Euclidean" [38, p. 61]. In the 'context of discovery' this is perhaps the crucial link – though paradoxically based on a conceptual confusion, as we shall see (§14) – from the initial speculation about freedom of coordinate choice to geometrization of gravitation.

The rotating disk argument has been much discussed; in fact, perhaps surprisingly, the physical status of rotating coordinate systems in general relativity is still not a matter of universal agreement among scholars.[7] However, some firm points can be established.

First of all, if all coordinate systems are allowed, then no guarantee exists that the coordinates of an arbitrary system possess, individually, *any* physical meaning at all. For instance, in special relativity it is also very easy to define coordinate systems where no coordinate can be interpreted as either 'timelike' or 'spacelike'. Notice that, even in terms of such unphysical coordinate systems, one can define expressions corresponding to measurable physical quantities: in a sense this is the whole point of introducing tensors and the absolute calculus (§7). The upshot is that in general relativity all measurable physical quantities should be expressed in tensor form (or by using a generalization of tensors, called *geometric objects*, cf. [91]). For the most general quantities this has been done (with important exceptions, cf. §13), but use of coordinates can hardly be avoided in concrete examples, and controversy over the physical meaning of different coordinate systems has raged among relativists from the beginning.

Second, in special relativity all the inertial observers agree upon the Euclidean character of space geometry (though they definitely disagree on the numerical values of lengths and angles!), but the introduction of non-inertial, even physically reasonable, co-

[7] See the book review [70].

ordinate systems inevitably leads to a variety of non-Euclidean space geometries – *even in special relativity*. Given our first point, this means not that one is forced to admit non-Euclidean space geometries, but that space geometry (as contrasted to *space–time* geometry) should not be given a big importance in general relativity.

Finally, if space–time is *not* of the Minkowskian type, then *for no possible choice of the coordinate systems* can one have *all* lengths and times (along coordinate lines) to be measured as the differences of the coordinates of points. Here the situation can be compared properly with the one holding on an ordinary sphere, where the longitude-latitude coordinate system (for instance) allows one to measure distances between same-longitude points as the difference of their latitudes, but not to measure distances between same-latitude points as the difference of their longitudes (except at the equator). In fact on the sphere there exist *no* coordinate systems which make it possible to interpret in this direct metrical sense the differences of *both* coordinates. This is connected with the issue of the curvature of a surface, as we shall see in the next section.

The fact that arbitrary coordinate systems cannot be endowed with a direct physical meaning was the main stumbling block in Einstein's path to general relativity, which delayed him until 1912, according to his own recollection: "I was much bothered by this piece of knowledge, for it took me a long time to see what coordinates at all meant in physics. I did not find a way out of this dilemma until 1912 [...]" [41, p. 316].

6. Gauss' Theory of Surfaces

Einstein had been taught the differential geometry of surfaces in Zürich [81, p. 212], and now that he had to deal with both non-Euclidean geometries and general coordinate systems, this was the place he was to look for tools and inspiration.

The German mathematician Carl Friedrich Gauss (1777–1855) had established a basic distinction between those properties of a surface which depend on the way it lies in the Euclidean 3-space, and those which pertain to its *intrinsic geometry*. The fundamental intrinsic property is the length of a curve lying on the surface.

If $x^a = x^a(u)$ (with $a = 1, 2, 3$ and $u = (u^1, u^2)$ varying in an open subset of $\mathbb{R}^2$) is a parametrization of a piece of a surface,[8] and $u^1 = u^1(s)$, $u^2 = u^2(s)$ with s varying in the interval $[a_1, a_2]$ are the parametric equations of this curve in the surface coordinates u^1, u^2, then the length of the curve is given by the integral:

$$L = \int_{a_1}^{a_2} \sqrt{\sum_{i,j=1}^{2} g_{ij}\big(u(s)\big)\frac{du^i}{ds}\frac{du^j}{ds}}\,ds,$$

where

$$g_{ij} = \sum_{a=1}^{3} \frac{\partial x^a}{\partial u^i}\frac{\partial x^a}{\partial u^j}.$$

[8] Notice that it is not always the case that a single parametrization suffices to cover the whole surface. For instance the ordinary sphere requires at least two.

Now the g_{ij} can be computed from data which are all available to a dweller on the surface, so can L. In the traditional infinitesimal language of calculus one can say that the basic metric information about a surface is given by the distance of two 'infinitesimally close' points, whose square is:

$$ds^2 = \sum_{i,j=1}^{2} g_{ij}(u)\, du^i\, du^j. \tag{6}$$

This defines a quadratic form (or, equivalently, a scalar product) on each tangent space, and is called a *metric* on S.

By definition, all intrinsic properties of a surface are those which, in any given coordinate system, can be expressed in terms of the g_{ij}. Two surfaces that can be put into a one-to-one correspondence which preserves the length of curves are called *isometric*, and such a correspondence is called an *isometry*. Clearly, two isometric surfaces, different as they may be in other ways, have the same intrinsic geometry.

For instance, if we cut a 'square' out of a cylinder, this is in intrinsic terms indistinguishable from a square (with the same side) cut out of a plane, even though, as is obvious, the 'square' extracted from the cylinder is 'bent', and the other is not. The fact is that the bending of the cylindric 'square' has to do with its extrinsic, not with the intrinsic, geometry: indeed, the two 'squares' are isometric surfaces. In ordinary terms, this can be reformulated as follows: it is possible to have a perfect 'geographic' map of the cylindric square. On the other hand such a map is impossible for any region, no matter how small, of a *sphere*. So when we see that the geographic maps of a sufficiently big portion of the Earth's surface give curiously disproportionate features to some regions (e.g. Greenland), we are shown a phenomenon which can be proven to occur, to varying degrees, even for the smallest portions of a spherical surface. We shall see in a moment how this can be proven.

A natural question that can be asked for two points p, q on a connected surface is whether some of the curves linking them *minimize* the length. The problem was first given a solution in 1728 by the great Swiss mathematician Leonhard Euler [51]. Such curves, if any exist, will be (some of the) critical points of the length functional; in the formalism of the calculus of variations the condition defining the length-critical curves is symbolically indicated by:

$$\delta\left(\int_p^q ds\right) = 0,$$

which can be shown to be equivalent, in an arbitrary system of coordinates, to the second-order system of 2 ordinary differential equations:

$$\frac{d^2u^i}{ds^2} + \sum_{j,k=1}^{2} \Gamma^i_{jk} \frac{du^j}{ds}\frac{du^k}{ds} = 0 \quad (i = 1, 2), \tag{7}$$

where the functions Γ^i_{jk} are called *Christoffel symbols* and turn out to have the following expression:

$$\Gamma^i_{jk} = \frac{1}{2} \sum_{r=1}^{2} g^{ir} \left(\frac{\partial g_{rk}}{\partial u^j} + \frac{\partial g_{rj}}{\partial u^k} - \frac{\partial g_{jk}}{\partial u^r} \right), \tag{8}$$

where (g^{ir}) is the inverse matrix of (g_{ir}). Notice that $\Gamma^i_{jk} = \Gamma^i_{kj}$, that is, this 3-indices family of functions is symmetric with respect to the lower indices.

Curves satisfying the system of Eqs (7) are called *geodesics* of the surface. Well-known examples are the straightlines (and their segments) in the plane, and the arcs of the big circles on a sphere. As is clear from (8), the geodesics depend on the g_{ij} only, hence they belong to the intrinsic geometry of the surface.

Gauss introduced a concept which was to prove crucial in subsequent developments. It is a number which is defined at each point of a surface, called the *total* (or *Gaussian*) *curvature*, or simply the *curvature* of the surface at that point. Though its original definition did not suggest it, Gauss proved what he himself called "a most excellent theorem", which can be stated as follows: if there is an isometry between two surfaces, then corresponding points have equal curvature. In other words, the total curvature is an intrinsic geometric quantity.

Comparison of curvature functions provides one easy way of proving our previous statement on the non-existence of perfect maps of spherical regions, no matter how small. In fact it can be shown that a sphere of radius R has constant curvature equal to $1/R^2$, while the Euclidean plane has constant curvature equal to zero (it is *flat*); therefore no isometry can exist between any region of the plane and any region of the sphere, as claimed.

The fact that the curvature of the Euclidean plane vanishes everywhere suggests the natural question whether the converse implication holds: are flat surfaces isometric to the plane? The answer is positive, with an obvious restriction: if on the surface Σ the curvature vanishes, then for each point of Σ there exists *in a neighbourhood of that point* an Euclidean coordinate system (that is, such that $g_{ij} = \delta_{ij}$).[9] In other terms, the curvature is *the* local obstruction to the existence of Euclidean coordinates or, equivalently, the obstacle to the existence of a coordinate system with everywhere vanishing Christoffel symbols. As we shall see, it is this formal characterization of curvature that lends itself most directly to generalization.

7. The "Absolute Calculus"

In order to have a space–time analogue of Gaussian surfaces the concepts we have described in Sect. 6 had to be generalized to dimensions higher than two. This required, first of all, to develop the concept of an n-dimensional space, or *manifold*, which was to surfaces what n-dimensional Euclidean space is to the plane. Half a century before Einstein started to think about gravitation, in 1854, this task was accomplished by the German mathematician Bernhard Riemann (1826–1866) in a famous lecture entitled "On the hypotheses which lie at the basis of geometry" and published in 1868, posthumously.[10] This lecture is by common agreement one of the high points in the history of mathematical thought, and it provided the foundation for the development of differential geom-

[9] As usual, δ_{ij} is 1 if $i = j$, 0 if $i \neq j$.

[10] Though Riemann has been rightly qualified as "the man who more than any other influenced the course of modern mathematics" [99, p. 157], he was also deeply interested and knowledgeable in physics and philosophy, and this circumstance had an influence on his mathematical work. The history of relativity is the work of a number of scholars for whom the formal distinctions between different disciplines, which are so dear to bureaucrats, had little or no relevance.

etry over the next hundred years, including the tensor calculus (or "absolute calculus", where 'absolute' means 'coordinate-independent') to which the Italian geometers gave a decisive contribution (cf. [85]).

Thus, contrary to Newton's case, the mathematics required for general relativity was not Einstein's creature. It was ready-made when he needed it, and he had to ask assistance from his friend, the mathematician Marcel Grossmann, to reach a reasonable mastery of it. Indeed, he declared his theory to be "a real triumph of the general differential calculus as founded by Gauss, Riemann, Christoffel, Ricci, and Levi-Civita" [27], and the theory itself was to have an influence on the development of differential geometry [105]. Let us review here the main concepts and notation, for future reference.[11]

Essentially, an n-dimensional manifold is a set which can be covered with compatible coordinate systems with values in $\mathbb{R}^n$. Once the concept of a n-dimensional manifold is clarified, it is easy to generalize all the notions we have defined in the case of surfaces to spaces of higher dimension: it is sufficient to substitute everywhere n to 2. For instance, the basic concept of the metric is simply generalized by defining:

$$ds^2 = \sum_{\mu,\nu=1}^{n} g_{\mu\nu}(x)\, dx^\mu\, dx^\nu. \tag{9}$$

From now on we shall adopt the convention (introduced by Einstein) that when an index occurs twice in some expression (normally once in an upper, the other in a lower position), the sum over it is automaticaly understood: for Latin indices $(a, b, \ldots, i, j, \ldots)$ from 1 to 3; for Greek indices $(\alpha, \beta, \ldots, \mu, \nu, \ldots)$ from 1 to 4. Thus (9) can be re-written as:

$$ds^2 = g_{\mu\nu}(x)\, dx^\mu\, dx^\nu.$$

In modern terms, this is a symmetric rank 2 covariant tensor, such that $(g_{\mu\nu})$ is a positive definite matrix at each point in the domain of the coordinate system. This is also called a *Riemannian metric*, and a manifold with a Riemannian metric is a *Riemannian manifold*.

The equations of geodesics can be written in a similar way:

$$\frac{d^2u^\alpha}{ds^2} + \Gamma^\alpha_{\mu\nu}\frac{du^\mu}{ds}\frac{du^\nu}{ds} = 0, \tag{10}$$

where the $\Gamma^\alpha_{\mu\nu}$ are the Christoffel symbols, which can be proven to be expressed exactly as in the case of surfaces (cf. (8)):

$$\Gamma^\alpha_{\mu\nu} = \frac{1}{2}g^{\alpha\beta}\left(\frac{\partial g_{\beta\nu}}{\partial x^\mu} + \frac{\partial g_{\beta\mu}}{\partial x^\nu} - \frac{\partial g_{\mu\nu}}{\partial x^\beta}\right). \tag{11}$$

The concept of a tensor generalizes that of a vector and is fundamental in the differential geometry of manifolds. Tensors (or multilinear functions) enjoy the important property that their multi-indices expression (or *set of components*) in every coordinate

[11] Most textbooks on general relativity present a more or less detailed account of differential geometry and tensor calculus (e.g. [38,60]). Two useful textbooks are [14] and (more comprehensive) [18]; a wide-ranging account of pseudo-Riemannian geometry is contained in [79]. Weyl's masterpiece [104] is still well worth studying.

system is dictated by their very nature, in the sense that knowing the set of components of a tensor in a given coordinate system allows one to derive its set of components in any other coordinate system. Tensors can be multiplied, contracted and, on a Riemannian manifold, differentiated.

Generalizing Gaussian curvature to n-dimensional Riemannian manifolds is a little more tricky, but not exceedingly so. As we know, the curvature of a surface can be considered as that function which must vanish if a local coordinate system exists with vanishing Christoffel symbols. Thus we can ask under which condition a local coordinate system $(x^1, \ldots, x^n)$ on an n-manifold exists such that the Christoffel symbols vanish identically. A computation shows that a necessary condition, which in fact also turns out to be a sufficient one, is the vanishing of the following entity:

$$R^{\alpha}_{\mu\nu\beta} = \frac{\partial \Gamma^{\alpha}_{\mu\beta}}{\partial x^{\nu}} - \frac{\partial \Gamma^{\alpha}_{\mu\nu}}{\partial x^{\beta}} + \Gamma^{\alpha}_{\nu\tau}\Gamma^{\tau}_{\mu\beta} - \Gamma^{\alpha}_{\beta\tau}\Gamma^{\tau}_{\mu\nu}. \tag{12}$$

This object, with 4 indices, is another, very important example of a tensor: the *Riemann–Christoffel* (or *curvature*) *tensor*. It expresses the local deviation of the manifold from Euclidean n-space; when it vanishes everywhere the manifold is called *flat* (in agreement with the usage introduced above in the case of surfaces). Riemann first defined it exactly the way we have just shown, in a paper he submitted in 1861 for a competition at the French Academy of Sciences (the paper did not win, and was eventually published, posthumously, in 1876).

The *Ricci tensor*[12] is defined as the contracted tensor

$$R_{\mu\nu} = R^{\alpha}_{\mu\nu\alpha}.$$

By using the formula:

$$\Gamma^{\alpha}_{\alpha\mu} = \frac{\partial}{\partial x^{\mu}} \log\sqrt{|g|}, \quad \text{where } g = -\det(g_{\mu\nu}), \tag{13}$$

the Ricci tensor can be expressed in the following useful way:

$$R_{\mu\nu} = -\frac{\partial \Gamma^{\alpha}_{\mu\nu}}{\partial x^{\alpha}} + \Gamma^{\alpha}_{\nu\beta}\Gamma^{\beta}_{\mu\alpha} + \frac{\partial^2(\log\sqrt{|g|})}{\partial x^{\mu}\partial x^{\nu}} - \Gamma^{\alpha}_{\mu\nu}\frac{\partial(\log\sqrt{|g|})}{\partial x^{\alpha}}, \tag{14}$$

which clearly displays the Ricci tensor as a *symmetric* tensor (i.e. $R_{\mu\nu} = R_{\nu\mu}$). Formula (14) was one of the most frequently used in earlier computations in general relativity. It also explains why Einstein consistently assumed that the coordinate changes satisfy the condition $g = -1$; in such a case, indeed, we obtain a rather simple-looking formula (which, nonetheless, is the sum of 20 terms!):

$$R_{\mu\nu} = -\frac{\partial \Gamma^{\alpha}_{\mu\nu}}{\partial x^{\alpha}} + \Gamma^{\alpha}_{\nu\beta}\Gamma^{\beta}_{\mu\alpha}. \tag{15}$$

[12] Einstein used the following notation: $B^{\alpha}_{\mu\nu\beta}$ instead of $R^{\alpha}_{\mu\nu\beta}$, $G_{\mu\nu}$ instead of $R_{\mu\nu}$, and G instead of S for the scalar curvature; moreover he used the symbols $\Gamma^{\sigma}_{\mu\nu}$ to indicate the *negative* of what is today normally so denoted. Other authors today define the curvature tensor as the *negative* of that here defined. Here and in the following chapter we shall stick to Einstein's conventions (not to his symbols) except for the Christoffel coefficients. When comparing different accounts of relativity, readers are warned to check carefully both the notation and the definitions of the main mathematical entities; they might find it useful to consult the "Table of Sign Conventions" in [76], where thirty-odd works by different authors are classified.

When the Ricci tensor vanishes, the manifold is called *Ricci-flat*. It is clear that a flat manifold is also Ricci-flat, but the converse fails (this technical fact will turn out to be of the utmost importance in the debate on the extent to which general relativity can be considered to satisfy Mach's principle).

A further invariant quantity can be obtained by contracting the Ricci tensor itself; it is called *scalar curvature* and is defined as

$$S = g^{\mu\nu} R_{\mu\nu}. \tag{16}$$

Clearly in a Ricci-flat manifold, we have also $S = 0$.

On a Riemannian manifold it is possible to define several differential operators acting on tensors. These are all built from the coordinate independent version of the partial derivative, which is called a *covariant derivative*, and on a vector field V acts in the following way:

$$\nabla_\mu V^\nu = \frac{\partial V^\nu}{\partial x^\mu} + \Gamma^\nu_{\mu\lambda} V^\lambda. \tag{17}$$

One tensor operator which is very important in the formalism of general relativity is the *divergence*. In ordinary Euclidean 3-space and in orthonormal coordinates, the divergence of a vector is defined in the familiar way:

$$\operatorname{div} \vec{v} = \frac{\partial v^1}{\partial x^1} + \frac{\partial v^2}{\partial x^2} + \frac{\partial v^3}{\partial x^3} = \frac{\partial v^a}{\partial x^a}.$$

If general coordinates are used, the concept of covariant derivative allows us to discover immediately the correct divergence expression, for any dimension of the space, and for any Riemannian metric ds^2: $\operatorname{div} V = \nabla_\mu V^\mu$, which can be rewritten in a more transparent fashion by exploiting (13) again:

$$\operatorname{div} V = \frac{1}{|g|^{1/2}} \frac{\partial(|g|^{1/2} V^\mu)}{\partial x^\mu}. \tag{18}$$

One can define similarly the divergences of a tensor of any order.

8. Spaces of Constant Curvature and the Relativity of Position

The curvature tensor is related to the Gaussian curvature as follows: if we select two independent tangent vectors A, B at a point p in M, we can consider the set of all points of M which are joined to p by a geodesic issuing from p with velocity vector lying in the subspace Π generated by A, B. In the neighbourhood of p this is a surface $\Sigma(\Pi)$, and we can compute its Gaussian curvature $K(\Pi)$ in p. This is given in an arbitrary coordinate system by:

$$K(\Pi) = -\frac{g_{\lambda\mu} R^\lambda_{\nu\rho\sigma} A^\mu B^\nu A^\rho B^\sigma}{(g_{\mu\rho} g_{\nu\sigma} - g_{\mu\sigma} g_{\nu\rho}) A^\mu B^\nu A^\rho B^\sigma}.$$

If $K(\Pi)$ is independent of the tangent 2-plane Π, then $K(\Pi)$ is called the (sectional) curvature at p and is denoted simply by $K(p)$. It is an important theorem that if this independence holds for all points p (and $n > 2$, of course), then $K(p)$ is also independent of p, that is, M is a manifold *with constant curvature*.

In such a manifold, the curvature K and the scalar curvature S are proportional according to the formula:

$$S = -n(n-1)K,$$

where n is the manifold's dimension. (For instance for $n = 3$ we have $S = -6K$, and for $n = 4$, $S = -12K$.) Note that, vice versa, from the fact that S is constant (for instance, $S = 0$) it does *not* follow that K also is.

Let us illustrate these concepts with some examples which are of crucial importance in cosmology: *the simplest Riemannian manifolds of dimension 3 and constant curvature*.

The metric of the 3-sphere of radius R_0, that is of the subspace of $\mathbb{R}^4$ defined by the equation:

$$(\xi^1)^2 + (\xi^2)^2 + (\xi^3)^2 + (\xi^4)^2 = R_0^2, \tag{19}$$

can be expressed in spherical coordinates as:

$$d\sigma^2 = R_0^2\big(d\theta^2 + \sin^2\theta\big(d\phi^2 + \sin^2\phi\, d\chi^2\big)\big),$$

and a standard computation shows that $K = 1/R_0^2$. If we introduce the new coordinate $r = R_0 \sin\theta$, we can rewrite the metric as

$$d\sigma^2 = \frac{dr^2}{1 - Kr^2} + r^2\big(d\phi^2 + \sin^2\phi\, d\chi^2\big). \tag{20}$$

Now this formula happens to give a metric of constant curvature for *any* value of K, including negative numbers and zero, so it turns out to be quite useful in computations (cf. Chap. 6).

Interestingly, the issue of physical geometry is relevant to the issue of absolute vs. relative space, and thus to the relativity of motion. In fact the argument that it is inherently impossible to know one's own position in absolute space has been used to deprive this notion of any empirical content, and as a consequence of any scientific relevance. In his booklet on the foundations of mechanics, *Matter and Motion*, Maxwell wrote, anticipating in 1876 the kind of analysis made familiar fifty years later by the logical positivist school:

> All our knowledge, both of time and place, is essentially relative. When a man has acquired the habit of putting words together, without troubling himself to form the thoughts which ought to correspond to them, it is easy for him to frame an antithesis between his relative ignorance and a so-called absolute knowledge, and to point out our ignorance of the absolute position of a point as an instance of the limitation of our faculties. Any one, however, who will try to imagine the state of a mind conscious of knowing the absolute position of a point will ever after be content with our relative knowledge [74, p. 12].

However, this argument, including its final touch of mockery, relies heavily on an assumption on the nature of physical geometry of which Maxwell was apparently unaware. The assumption is that physical space is *homogeneous*, that is, that for any two points p, q there is an isometry of the whole space sending p to q; in particular the sectional curvature must be constant. Thus any two points in Euclidean 3-space or on the 3-sphere are, indeed, geometrically indistinguishable, but the same would not be true for two points, if such exist, of physical space where curvature assumes different values.[13]

It is perfectly conceivable that physical space has no isometries other than the identity map. In this case it would be justified to speak of space as an absolute entity. This shows that the homogeneity of space is in fact a basic principle of relativity.

A final fact which deserves to be emphasized is that two spaces which are both of constant curvature K can be proven to be isometric in the small, but not necessarily in the large, where *topological properties* enter. For instance, a plane and a cylinder are locally isometric (§6), but not (globally) isometric. The main instance where this difference becomes important is in the study of the large-scale structure of the universe, where topological notions must be given special attention.

9. The Generalized Space–Time

When Minkowski introduced his geometric approach he neatly set the stage for further generalization by Einstein, since he expressed the inner product as the pseudo-distance between two 'infinitesimally close' points:[14]

$$ds^2 = -dx^2 - dy^2 - dz^2 + c^2dt^2 = \eta_{\mu\nu}\, dx^\mu\, dx^\nu,$$

where we have put $x^1 = x$, $x^2 = y$, $x^3 = 3$, $x^4 = t$. This expression generalizes very naturally to a Riemannian metric in the form:

$$ds^2 = g_{\mu\nu}\, dx^\mu\, dx^\nu.$$

What was unusual in the mathematics of general relativity, even to mathematicians, was the *signature* of the metric of the relativistic space–time, which was Lorentzian rather than Euclidean (§4). However, the definitions we have briefly listed for Minkowski space–time transfer with no great difficulty to the case of Lorentzian manifolds.

In a general-relativistic space–time there are several (local) time functions, and therefore several (local) spatial sections. As we have emphasized (§5), even in Minkowski space–time these spaces are not all Euclidean, and need not share the same geometry. So in general relativity, just as in Riemann's geometry, a plurality of possible space geometries has to be admitted from the start; we shall see in Chapter 6 some striking and relevant examples.

Although Einstein is often identified as the man who introduced the non-Euclidean geometries in physics, this is far from true. Other scientists before him investigated the possibility that the geometry of the physical space could be non-Euclidean, and thought that it was partly an empirical matter to establish which was the true physical geometry [37]. In fact both claims had been discussed in the nineteenth century, notably by Lobachevsky, Riemann and Clifford; Gauss himself, who had been involved in geodetic surveys, was not alien to this line of thought. It has been estimated [77, p. 73] that at least 80 papers were published in the fifty years before 1915 on the physics valid in a world where the true geometry is non-Euclidean!

[13]This refutation of Maxwell's argument was presented by W.K. Clifford in [15, pp. 193–204].

[14]One should not take the square in 'ds^2' literally, as if it meant that Lorentzian distance may sometimes be an imaginary number: it only means that ds^2 is a quadratic form.

The British mathematician William Kingdom Clifford, who translated Riemann's famous lecture, made some bold conjectures, and in his popular *The Common Sense of the Exact Sciences*, posthumously published in 1885, stated:

> We may conceive our space to have everywhere a nearly uniform curvature, but that slight variations of the curvature may occur from point to point, and themselves vary with the time. These variations of the curvature with the time may produce effects which we not unnaturally attribute to physical causes independent of the geometry of our space. We might even go so far as to assign to this variation of the curvature of space 'what really happens in that phenomenon which we term the motion of matter' [15, pp. 202–3].

In 1900 Karl Schwarzschild, an astronomer and applied mathematician of the school of Göttingen, discussed the possibility of identifying the geometry of physical space by measurements of stellar parallaxes, and also fixed numerical bounds on the curvature, based on the then available data ([93]; cf. [87]).

On the other hand, two years later in his *Science and Hypothesis* [83], Poincaré famously argued that, though different sets of geometric axioms could be assumed in physics, none could be held to be the 'true' one: geometry was to a large extent a matter of reasonably choosing one's conventions. This sobering approach, however, was accompanied by a risky prediction, namely, that for the foreseeable future physicists would have stuck to Euclidean geometry, not for any confidence in the 'certainty' of this geometry, but because of its superior simplicity. However, the actual development of physics took a different direction – a good cautionary tale for historical prophecy.

10. What is Space–Time, Really?

In 1913 Einstein coauthored with his old friend Marcel Grossmann an essay, *Outline of a Generalized Theory of Relativity and of a Theory of Gravitation* [42], which was divided into one "Physical Part", by Einstein, and a "Mathematical Part", by Grossmann. In this work, we find several anticipations of things to come and basic ingredients which were to stay in the subsequent versions, all embedded in a fog of misunderstandings and also mathematical errors.

One decisive point which was to remain was the form of the field equations. In Sect. 5 Einstein suggested that his intended adaptation of Poisson's equation:

$$\Delta\phi = 4\pi G\rho, \tag{21}$$

which is the differential form of the Newtonian attraction law,[15] would likely have the form:

$$\kappa T^{\mu\nu} = G^{\mu\nu},$$

where $G^{\mu\nu}$ was to be a rank 2 contravariant tensor derived from $g_{\mu\nu}$ and containing up to the second-order derivatives of the $g_{\mu\nu}$. The tensor also had to be symmetric, since $T^{\mu\nu}$ was the "stress-energy tensor of the material flow"; the form of $T^{\mu\nu}$ was derived in the case of "continuously distributed incoherent masses" – that is, a *dust* – as:

$$T^{\mu\nu} = \rho\frac{dx^\mu}{ds}\frac{dx^\nu}{ds}, \tag{22}$$

[15] More about this point will be said in Chap. 6, §§1 and 4.

and it was shown that, under the assumption of infinitesimal conservation of the energy-momentum, it satisfied a law which Einstein postulated for the stress-energy tensor of *any* matter distribution:[16]

$$(\operatorname{div} T)^{\mu} = \nabla_{\lambda} T^{\lambda\mu} = 0. \tag{23}$$

A major setback was that at this stage Einstein had come to be convinced that the general covariance principle (§1) had to be given up, and that one should be satisfied with a theory in which the admissible coordinate changes are *linear*. He explicitly wrote: "[. . .] we have no basis whatsoever for assuming a general covariance of the gravitational equations". This opinion was subsequently strengthened, when he stumbled upon the following conceptual difficulty, first presented in Sect. 12 of the article, which bears the unambiguous title "Proof of a Necessary Restriction in the Choice of Coordinates".

The argument, known as 'the hole problem' and which has spanned over the last two decades a vast historical and epistemological literature ([78], cf. [71]), can be formulated as follows. Suppose that space–time is identified with a certain region S of R^4, and that we change coordinates only inside a bounded domain D, leaving the standard coordinates unchanged outside. In modern terms this is equivalent to defining a diffeomorphism Φ of S, which sends D into itself and is equal to the identity everywhere except in D. Now consider two distributions of matter on S, as represented by the stress-energy tensors T and T', with T vanishing in D and $T' = \Phi_*(T)$, i.e. T' is the result of transforming T by means of Φ. Suppose $g = g_{\mu\nu}\, dx^{\mu}\, dx^{\nu}$ is the space–time metric, and let $g' = \Phi_*(g)$. Then it is clear that T', *as a function of the standard coordinates*, must be identical with T, while g' is certainly not the same as g. The unavoidable consequence seems to be that the stress-energy tensor does not, and cannot, determine by itself the space–time metric. Or one may say, with Einstein, that there are *two* different solutions for the $g_{\mu\nu}$ in the *same* coordinate system, which coincide on the boundary of D: "In other words, *the course of events in this domain cannot be determined uniquely by general-covariant differential equations*" ([25]; italics in the original).

Einstein discussed this problem with several colleagues, and received crucial help from them (cf. for instance [64]). At first, as we have seen, he decided to restrict coordinate changes only to linear transformations. Within this special class of coordinate changes it is clear that a diffeomorphism such as Φ cannot exist.

Eventually, however, he came back to general covariance and solved his difficulty by admitting that the triples (S, T, g) and (S, T', g') must be considered as representing *one and the same space–time*. In other words, space–time as a mathematical model of the physical space–time is uniquely determined only up to isometry class: so it is not a single space, but an equivalence class of spaces.[17] On the other hand, the physical space–time, as the set of all events (or physical coincidences), can obviously be given an intrinsic meaning; as Einstein wrote to Ehrenfest at the end of his journey to general relativity (December 26, 1915):

> The physically real in the world of events (in contrast to that which is dependent upon the choice of a reference system) consists in *spatiotemporal coincidences* [footnote: "and in nothing else!"] (cit. from [64, p. 37]).

[16] What Einstein actually wrote would be, in our notation: $|g|^{1/2}(\operatorname{div} T)_{\mu} = 0$.

[17] This point is emphasized in modern textbooks, such as [60, p. 56] and [90, p. 27].

It was in the autumn of 1915 Einstein that came to the conclusion that the field equations he had written down in his paper with Grossmann [24] were incorrect, and that the restriction to linear coordinate changes was misconceived. The final work took place from October to November – "one of the most exciting and strenuous times of my life, but also one the most successful ones", as he wrote to Sommerfeld [81, p. 250]. On 25 November he presented to the Royal Prussian Academy the paper containing the revised field equations with matter. During the month of November he submitted in all 4 papers to the Academy (on the 4th, 11th, 18th, 25th), all quite short (two dozens pages in all), and with a number of changes and second thoughts from one to the other. In a letter of January 1916 to Lorentz, Einstein honestly described them as follows: "The series of my papers about gravitation is a chain of false steps [Irrwegen] which nevertheless by and by led to the goal" [81, p. 271]. It is clear that Einstein was under pressure not to be outdone in the search for the field equations by other scientists he knew were working on the same problem, and particularly by David Hilbert, one of the greatest mathematicians of his age.[18]

11. The Newtonian Limit

We shall not dwell on the several "false steps" documented in the first two papers [27,28] of this breathless series, and shall concentrate on the two final papers.

The paper of 18 November [29] deals with two applications of the new field equations to the study of the planetary motion, and corrects by a factor of 2 formula (1), that is, the prediction on the deflection of light rays made four years before in [23]. Notice that at this stage he has arrived at an incomplete version of the final field equations:

$$R_{\mu\nu} = \kappa T_{\mu\nu}. \tag{24}$$

Einstein uses (24) to show that

> this most fundamental theory of relativity [...] explains qualitatively and quantitatively the secular rotation of the orbit of Mercury (in the sense of the orbital motion itself) which was discovered by Leverrier and which amounts to 45 sec of arc per century.

This he calls "an important confirmation" of the theory. He adds:

> Furthermore, I show that the theory has as a consequence a curvature of light rays due to gravitational fields twice as strong as was indicated in my earlier investigation.

Einstein links this correction of (1) directly to the hypothesis that the trace of the matter tensor must vanish, an assumption he is making at this time and which implies, given (24), *that the scalar curvature of the metric must be zero*. This hypothesis was discarded just a few days later, in the fourth paper.

The method used to solve the planetary problem is an approximation scheme, which takes the metric to be of the form:

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu},$$

where $\eta_{\mu\nu}$ is the Minkowski metric (with units such that $c = 1$), and $h_{\mu\nu}$ is "small" with respect to 1.

[18] On the 'race' Einstein–Hilbert see the update [17].

Moreover it is assumed that 1) the metric $g_{\mu\nu}$ is time-independent and time-orthogonal (the time coordinate being $t = x^4$); 2) it reduces to the Minkowski metric at spatial infinity; and 3) the spatial part of the metric is spherically symmetric. This gives the following approximate form of the metric:

$$ds^2 = -\left(\delta_{ab} + \alpha \frac{x^a x^b}{r^3}\right) dx^a\, dx^b + \left(1 - \frac{\alpha}{r}\right) dt^2. \tag{25}$$

The field Eqs (24) are used *under the hypothesis of empty space*, so that they can be written simply as:

$$R_{\mu\nu} = 0. \tag{26}$$

The main point of the mathematical development is to arrive at an expression for the potential which contains a term in r^{-3}, so that the equations of motion for a planet can be written, in second approximation, as:

$$\frac{d^2 x^a}{ds^2} = \frac{\partial \Phi}{\partial x^a}, \quad \text{with } \Phi = -\frac{\alpha}{2r}\left(1 + \frac{B^2}{r^2}\right).$$

It is in fact the cubic term which is responsible for the perihelion advance effect, as had already been shown twenty years before by one Paul Gerber, who had also given the very same final formula obtained by Einstein for the advance:

$$\epsilon = \frac{24\pi^3 a^2}{T^2 c^2 (1 - e^2)}, \tag{27}$$

where e is the excentricity, a is the semi-major axis and T is the period of the unperturbed orbit. This circumstance, and even more the sense in which general relativity "explains quantitatively and qualitatively the secular rotation of the orbit of Mercury" deserves a digression.

12. Einstein and Mercury's Perihelion

It is certainly remarkable that a German high-school teacher (because such was Gerber (1854–?)), had proposed in a renowned physics magazine in 1898 a different potential for the gravitational interaction which gave, in approximation, precisely the "r^{-3}" correction (with the right coefficient!), and provided a striking connection between the gravitational interaction and the speed of light. The result did not go unnoticed at the time, notwithstanding the obscurity of the author, as is proved by the fact that Ernst Mach mentioned it in his *Mechanik*, for the first time in 1901, in the fourth edition. Interestingly, Mach stressed that "only Paul Gerber" had succeeded in connecting the propagation of gravitation with the speed of light.[19]

Doubts as to the correctness of the relativistic derivations were already being voiced by some eminent contemporaries, such as the mathematicians C. Burali-Forti,

[19] "Only Paul Gerber [and here reference is made to [55]], studying the motion of Mercury's perihelion, which is 41 [sic] seconds of arc per century, did find that the speed of propagation of gravitation is the same as the speed of light. This speaks in favour of the aether as the medium of gravity" (Mach 1901, p. 199).

S. Zaremba, C.L. Poor, between 1922 and 1923. This was not a side issue, since the lack of rigour in the derivation was the main argument used by Einstein to discredit Gerber's priority when he was challenged on this score:

> The other way [to explain Mercury's perihelion advance without the theory of relativity] is to quote a paper by Gerber, who gave the correct formula for the movement of the Mercury perihelion before I did. Yet experts not only agree that Gerber's derivation is faulty from beginning to end, but also that the formula cannot be obtained as a consequence of the assumptions from which Gerber started out. The paper of Herr Gerber is, therefore, completely worthless, a miscarried and irreparable theoretical attempt. I state that the general theory of relativity provided the first real explanation of the movement of the perihelion of Mercury. I had not mentioned the Gerber paper because I was not aware of it when I wrote my paper of the movement of the Mercury perihelion; but even if I had known about it, there would have been no reason to mention it [36].

So for Einstein the fact that the right formula (27) had been derived from Gerber's theory could not count in favour of that theory as long as the derivation was logically objectionable: indeed it was not enough to give that theory even a modicum of respectability (cf. "completely worthless").[20]

The question of the problematic aspects of the relativistic derivation of Mercury's advance surfaced even at the level of the Nobel committee in 1921–1922: one of the experts, A. Gullstrand, from Uppsala, "a scientist of very high distinction", had objected that "other, long-known deviation from the pure two-body Newtonian law should be re-evaluated with general relativistic methods before there could be even an attempt to identify the residual effect to be explained" [81, p. 509].[21] John L. Synge, who was the author of one of the classic reference books on relativity [101], wrote half a century after Einstein's first formulation of general relativity:

> [...] when one examines some proofs in the Neo-cartesian spirit, too often they seem to dissolve completely away, leaving one in a state of wonder as to whether the author really thought he had proved something. Or is the reader stupid? It is hard to say. *In any case I am still waiting for a rational treatment of the dynamics of the solar system according to Einstein's theory* [[100, p. 14] italics added].

Things are made even more confusing by the circumstance that the total observed secular advance is $5599''.74 \pm 0.41$, most of which is accounted for in terms of inertial effects (i.e. the Earth not being an inertial frame).[22] Moreover the general relativistic derivation requires that the Sun is strictly a sphere, and a deviation from the spherical form is enough to destroy the ideal Keplerian ellipsis and to produce a perihelion advance even in Newtonian physics.

Einstein's winning opinion that his derivation, relying so heavily on the Newtonian explanation of all but the residual anomalous advance of about $43''$ – an argument, not surprisingly, termed "intellectually repellent" by Synge – gave general relativity

[20] In referring to the "experts", Einstein is alluding to Seeliger, who in 1906 had advanced the accepted Newtonian explanation of Mercury's anomalous advance. When in 1917 Gerber's 1902 paper was republished, this time in *Annalen der Physik*, Seeliger claimed that in his calculations Gerber had made an elementary mistake. But the mistake was Seeliger's, according to [88, p. 139].

[21] Pais – who, incidentally, does not mention Gerber's work – comments that this objection is "not very weighty"; but he does not justify in any way this dismissive remark, which seems to me not just cavalier but incorrect.

[22] As explained in [101, p. 296, note 4]; see also [76, p. 1113, Box 40.3].

such an exalted status, while Gerber's anticipation, which had been appreciated by Mach, did not even deserve to be cited, is surely an interesting topic for sociologists of science.

13. The 'Final' Field Equations

The final November 1915 [30] paper begins with a pithy account of the roundabout development of the field equations in the first two papers:

> Historically they evolved in the following sequence. First, I found equations that contain the Newtonian theory as an approximation and are also covariant under arbitrary substitutions of determinant 1. Then I found that these equations are equivalent to generally-covariant ones if the scalar [i.e. the trace] of the energy tensor of "matter"vanishes. The coordinate system could then be specialized by the simple rule that $\sqrt{-g}$ must equal 1, which leads to an immense simplification of the equations of the theory. It has to be mentioned, however, that this requires the introduction of the hypothesis that the scalar of the energy tensor of matter vanishes.

Now Einstein says he has found how to avoid "this hypothesis about the energy tensor of matter", and the method is "merely by inserting it into the field equations in a slightly different manner". The field equations for the vacuum, which were used in the previous paper, are untouched by this change, which leads to the final version:

$$R_{\mu\nu} = -\kappa\left(T_{\mu\nu} - \frac{1}{2}T g_{\mu\nu}\right), \tag{28}$$

which in today's textbooks is more frequently found in the (equivalent) form:

$$R_{\mu\nu} - \frac{1}{2}S g_{\mu\nu} = -\kappa T_{\mu\nu}. \tag{29}$$

The left side is known today as the *Einstein tensor*. (We shall come back in Chapter 6, §11, to the rationale of this expression.)

The Newtonian limit is elaborated in general in [31, Sect. 21]. The assumptions are the following: $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$ with $h_{\mu\nu}$ "small compared with 1" (in ordinary units this means 'small compared with the speed of light c'), and quantities of second and higher order are to be neglected; at spatial infinity the $h_{\mu\nu}$ must go to zero; the 3-velocity of the particles is "small"; and the time derivatives of the $g_{\mu\nu}$ must be negligible. In the case of matter modelled as a dust, we obtain the Poisson's equation in the form:

$$\Delta\phi = \frac{\kappa c^4}{2}\rho \tag{30}$$

whence it follows, by comparison with the ordinary Poisson's Eq. (21), that $\kappa c^4/2 = 4\pi G$, that is,

$$\kappa = \frac{8\pi G}{c^4} = \frac{1.87 \times 10^{-27}\ \text{cm/g}}{c^2}, \tag{31}$$

where G is the gravitational constant. As is clear, the 'coupling constant' κ between 'geometry' and 'matter' is very small.

Even in the last 1915 version, the stress-energy tensor of matter is *postulated* to have zero divergence. This is evidence enough that Einstein ignores the so-called *(second) Bianchi identities*, which imply that the left-hand side of (28), and thus also the energy tensor, has zero divergence for *any* metric. It is interesting to mention that the Bianchi identities had been discovered in 1880, but until the end of 1915 Einstein was not aware of them, and, surprisingly, neither was Hilbert.[23]

By using what to Einstein is an independent assumption, i.e. $\operatorname{div} T = 0$, a "conservation theorem of matter and gravitational field" is obtained, in the form:

$$\frac{\partial}{\partial x^{\mu}}\left(T_{\nu}^{\mu} + t_{\nu}^{\mu}\right) = 0. \tag{32}$$

It must be pointed out, however, that t_{ν}^{μ}, which Einstein calls "the 'energy tensor' of the gravitational field" in [30], is *not* a tensor, as he acknowledges in [31]; today it is called the *pseudotensor* of the gravitational field. In particular, it is always possible to choose a coordinate system such that, at a given point, it vanishes. The fact is that in general relativity *one cannot derive genuine integral conservation laws*, notwithstanding all the emphasis placed on the Lagrangian formalism by Einstein and his contemporaries. There is not even a coordinate-independent notion of 'inertial mass' of a system. At most, one can interpret the vanishing of the stress-energy tensor as a *local* conservation law, which can be applied in regions where the space–time curvature happens to be negligible. Given the importance of the (global!) laws of conservation in the historical development of physics, one might have expected a more widespread dissatisfaction amongst physicists as regards these features of general relativity, but eventually most of them accepted even this departure from classical physics.[24]

The conclusion of the paper is that, though

> the postulate of relativity in its most general formulation (which makes space–time coordinates into physically meaningless parameters) leads with compelling necessity to a very specific theory of gravitation that also explains the movement of the perihelion of Mercury [...]

nonetheless that postulate

> cannot reveal to us anything new and different about the essence of the various processes in nature that what the special theory of relativity taught us already.

Except for gravitation itself, apparently, if "compelling necessity" is to be taken seriously. In a sense this statement represents an anticipated, if puzzling, answer to Kretschmann's criticism of the principle of general covariance. In the next section we shall discuss at some length the subtle and important issues involved.

The paper in which Einstein formulated for the first time in a systematic fashion the foundations of his general relativity only appeared in 1916 in *Annalen der Physik*, and with its sixty-odd pages was the size of a booklet (in fact it was also published as such,

[23] [81, p. 275; 89]. In the best tradition of mathematical misattributions, the Bianchi identities were not discovered by Bianchi. It seems that the first to find them was Ricci, who communicated his result to E. Padova, who published it in 1889; however, the *contracted* Bianchi identities, which are sufficient for the proof that the Einstein tensor always vanishes, had been published by Aurel Voss *nine* years earlier.

[24] Of course objections were already raised at the time, among others by Lorentz, Levi-Civita and Schrödinger. A recent criticism of general relativity from the viewpoint of its failure of admitting genuine laws of conservation is contained in [69]; see also the discussion in [9].

with added "Einleitung" and "Inhalt"). It consists of 22 sections grouped into five parts; the longest one is the second (more than one-third), which provides a (very essential) introduction to Riemannian geometry and tensor calculus. This part was certainly not easy going for whoever was not already acquainted with the basic mathematical concepts. It was precisely the novelty of the mathematics – not in itself, but as to its use in a physical theory – which explains the awe which many of Einstein's colleagues felt at their first encounter with the equations of the theory. Though many a mathematician could boast to know better and handle differential geometry more skilfully than Einstein,[25] the same could not be said of the average physicist. It is this circumstance that earned Einstein, widely though incorrectly, the fame of being a mathematician more than a physicist.

On the other hand, apart from Grossmann's crucial collaboration, Einstein had been supported and helped by the mathematicians of Göttingen, including Felix Klein, Hilbert and Hermann Weyl. He stayed there between June and July 1915, and also gave a series of six lectures on the still developing general relativity; these lectures were very favourably received. It is interesting that the eminent algebraist Emmy Noether wrote in 1915 that she was in a team *performing difficult computations for Einstein*, and that "none of us understands what they are good for" [81, p. 276]. He corresponded on technical difficulties also with other mathematicians and physicists, including Lorentz [65], Tullio Levi-Civita [12] and Paul Hertz [64]. So much for the enduring myth of the genius who works his way in solitude.

14. Was the Copernican Controversy a Pseudo-Problem?

One widespread interpretation of general relativity, and the one that may be most responsible for its fame among non-specialists, was that it gave the final verdict on the Copernican controversy. The verdict – it is reported – is that the controversy was rooted in a conceptual confusion: there is no fact of the matter as to whether the Earth or the Sun is at rest, since all depends on one's (arbitrary) choice of coordinate system. This earned Einstein the honour of being compared to Copernicus, and many popular books appeared which traced the development of physics "from Copernicus to Einstein" (cf. [86]).

This interpretation was to some extent endorsed by Einstein himself, in the book he coauthored in 1938 with Leopold Infeld [43, p. 212]:

> Can we formulate physical laws so that they are valid for all CS [= coordinate systems], not only those moving uniformly, but also those moving quite arbitrarily, relative to each other? The struggle, so violent in the early days of science, between the views of Ptolemy and Copernicus would then be quite meaningless. Either CS would be used with equal justification. The two sentences, 'the sun is at rest and the earth moves', or 'the sun moves and the earth is at rest', would simply mean two different conventions concerning two different CS. Could we build a real relativistic physics valid in all CS; a physics in which there would be no place for absolute, but only for relative motion? This is indeed possible!

Similarly Max Born wrote: "Thus from Einstein's point of view Ptolemy and Copernicus are equally right. What point of view is chosen is a matter of expediency" [8, p. 345].

[25] Cf. Felix Klein's comments quoted at pp. 33, 36 of [64].

Though one is entitled to be a little surprised by the rash approach to the history of science that these passages reveal, one should not in general under-rate the importance of such oversimplifications in getting public and professional recognition for a certain scientific theory. Historical deformation is not the smallest factor in the attraction a new scientific theory exerts on lay persons and professionals alike.[26] However, in this case the trouble is not just with the way general relativity was advertised: it had to do with Einstein's own perception of the contribution he had given by adopting "general covariance". On several occasions he stated that in his theory all coordinate systems are on an equal footing, and that *therefore* he had succeeded in generalizing the special principle of relativity to all accelerated systems.

In order to clarify this point, it is useful to remember that Newtonian mechanics has a very effective means for dealing with non-inertial systems. For instance, if we have an inertial system $(\vec{r}, t)$ and a uniformly rotating system with same origin $(\vec{r}\,', t)$, the coordinate change is given by $\vec{r}\,' = A\vec{r}$, where A is a function of time with values in the group of the special orthogonal matrices. It can be shown, by computation, that if the second principle of dynamics $\vec{F} = m\vec{a}$ is satisfied for a given particle of mass m, then $\vec{F}' = m\vec{a}'$ also holds, with

$$\vec{F}' = A\vec{F} + \vec{F}_1 + \vec{F}_2, \tag{33}$$

where $\vec{F}_1$ is the Coriolis force and $\vec{F}_2$ is the centrifugal force, both depending on the angular velocity of the rotating system with respect to the inertial one. So one might say that the second principle of dynamics is satisfied also in rotating systems, at least if one is willing to consider (33) as a reasonable transformation law for a force. Why did this undisputed fact not stop the search for a theory without absolute space and time? The answer, of course, is that the *form* of the forces in the two systems is *not* preserved: for instance, if $\vec{F}$ depends just on the distance of the particle from the origin, then it is clear that the same property does not, and cannot, hold for $\vec{F}'$. In other words, the *notational similarity* of the original and the primed force laws cannot be taken as indicating *law invariance* in the spirit of the principle of (Galilean or special) relativity.

So there is no doubt that the mere acceptance of 'all coordinate systems' in the formulation of the physical laws cannot be read as a generalization of the principle of relativity – otherwise classical mechanics should also be considered 'generally relativistic', at least in the sense of giving equal mathematical treatment to inertial and non-inertial systems.

In 1917 mathematician Erich Kretschmann [66] made a similar point, arguing that "general covariance" in Einstein's sense was not a substantive physical principle, since it could be adopted in the formulation of virtually *any* physical theory. Einstein replied by saying that he concurred, adding only that the principle "carries considerable heuristic weight", and explaining what he meant as follows [33]:

> Among two theoretical systems, both compatible with experience, one will have to prefer the one that is simpler and more transparent from the point of view of absolute differential calculus. One just should bring the mechanics of Newtonian gravitation in the form of absolute-covariant equations (four-dimensional) and one will certainly become convinced that princi-

[26]It is interesting that one of the first times Einstein was likened to Copernicus was by no less than Max Planck, in his (successful) recommendation of Einstein in 1910 for a professorship in Prague [81, p. 192].

ple (a) [i.e. general covariance] excludes this theory, not on theoretical grounds, but on practical ones!

It is clear that Einstein meant this as a rhetorical challenge. In fact just a few years later the first generally covariant four-dimensional formulations of Newtonian gravitation were published [11,53]. If 'simplicity' and 'transparency' are to be judged, it is at least doubtful if the covariant 4-dimensional version of Newtonian physics should be considered as more obscure or difficult to handle than general relativity.[27]

Coming back to the main physical issue which was at stake in the Copernican controversy, does general relativity make any difference to the distinction between locally inertial and non-inertial (e.g. rotating) systems? Not much. As two Italian physicists pointed out in 1929 [58], in general relativity a locally rotating system, for instance, is as easily distinguishable from a locally inertial system as it is in classical mechanics. In other words, inertial effects exist also in general relativity![28] If this was not the case, general relativity would simply contradict ordinary empirical evidence.[29]

However, it can be argued that as to absolute motion, general relativity does introduce a novelty in the spirit of Mach's criticism: the locally inertial systems are *determined* by the metric (they are the freely falling systems), and the metric, in turn, is *linked* to the energy–matter distribution through the field equations, so the inertial effects turn out to be *linked* to that distribution, rather than to absolute motion. However, general relativity is not a completely 'Machian' theory, for the simple but inescapable mathematical reason that in the preceding statement 'linked' cannot be changed to 'determined'. This is proven, for instance, by the existence of nontrivial (i.e. different from Minkowski space–time) solutions of the vacuum field Eqs (26) (see next section). The debate on Mach's arguments against the basic Newtonian notions is still going on, and papers and books adopting a 'Machian' standpoint, sometimes seriously departing from general relativity and going back to classical notions, appear from time to time [7,6,56].

Our conclusion is that, though repeatedly and authoritatively endorsed, the claim that general relativity finally 'solved' the Copernican controversy by showing that Copernicus and Ptolemy were 'both right', is based on several mistakes, concerning both the gist of their disagreement and the scope of 'general covariance'. However such a claim was probably decisive in winning for Einstein's theory a cultural and philosophical place second to no other physical theory of his time.[30]

15. The First Exact Solution

The metric (25) used by Einstein to give his explanation of the anomalous advance of Mercury's perihelion was not an exact solution of his equations. The problem of finding the exterior metric corresponding to the field produced by a spherically symmetric body

[27] For a more recent outline, see [21].

[28] The mathematical tool by which this can be most easily shown was introduced by the 21-year-old Enrico Fermi in 1922 [52].

[29] Incidentally, it was this basic reason, rather than "a great and unfortunate talent for creating difficulties for himself" [81, p. 232] that made Max Abraham [4,5], as well as many other physicists sceptical about the very programme of generalizing the principle of relativity.

[30] This privilege was later stolen by quantum mechanics, for similarly unwarranted claims concerning its radical philosophical implications. For more on the topic discussed in this section [57] is worth reading.

was solved within weeks from the publication of [29] by Schwarzschild (cf. §9), who was then serving as a volunteer on the Russian front, and knew Einstein's third paper, not the fourth. The metric he found ([94], communicated to the Prussian Academy by Einstein himself on January 13), was a solution of the *vacuum* field Eqs (26), and this was a happy circumstance, since the full field equations of 18 November (24) do not coincide with the full final Eqs (28).

This is the metric:

$$ds^2 = -\frac{dr^2}{1-\alpha/R} - R^2\big(d\theta^2 + \sin^2\theta d\phi^2\big) + \left(1 - \frac{\alpha}{R}\right)c^2\,dt^2, \tag{34}$$

where (r, θ, ϕ) are standard spherical coordinates, R is an auxiliary variable, defined as $R = (r^3 + \alpha^3)^{1/3}$, and α is a positive expression containing the mass of the body $(\alpha = 2Gm/c^2)$.

Slightly more than a month later, and a few months before dying aged 43 of an illness, Schwarzschild submitted a second paper ([95], also communicated by Einstein, on 24 February), where he described the exact solution for the interior of a homogeneous sphere of incompressible fluid with (constant) density ρ_0. The metric was:

$$\begin{aligned} ds^2 = &-\frac{3}{\kappa\rho_0}\big(d\chi^2 + \sin^2\chi d\theta^2 + \sin^2\chi\,\sin^2\theta d\phi^2\big) \\ &+ \left(\frac{3\cos\chi_a - \cos\chi}{2}\right)^{-2} c^2\,dt^2, \end{aligned} \tag{35}$$

where the subscript a indicates that the quantity is computed at the surface of the sphere. The conclusion of the second article is:

> For an observer measuring from outside [...] a sphere of a given gravitational mass [...] cannot have a radius measured from outside smaller than: $P_o = \alpha$. For a sphere of incompressible fluid the limit will be $9\alpha/8$. (For the Sun α is equal to 3 km, for a mass of 1 gram is equal to $1.5 \cdot 10^{-28}$ cm.)

In fact at the centre of a sphere having a radius equal to the limit value, the pressure of the fluid would become infinite.

Schwarzschild's contributions to general relativity rank as some of the most important ever. Among other things, his exterior solution (34) made it possible to derive the three empirical consequences in a more satisfactory way than Einstein had been able to do. Notice, however, that these predictions are not enough in themselves to give general relativity a unique status. Those who think that it is very unlikely that two different theories may reproduce the same empirical predictions (at a certain historical time) should consider that (34) is also a solution of all field equations having at the left side an expression formed from the derivatives, of *any* degree, of the Ricci tensor $R_{\mu\nu}$; this elementary remark implies that the three famous predictions are also made by all theories postulating such field equations.[31]

Notwithstanding their rôle in fixing some of the loose ends of Einstein's arguments, Schwarzschild's results were puzzling in other ways. First of all, here was another solution (after Minkowski space–time) for the vacuum field equations, something which

[31] This example is of more than merely methodological interest; see [73].

seemed inconsistent with Mach's principle. Second, (34) fails for $R = \alpha$, which corresponds in Schwarzschild's coordinates to the centre of the sphere ($r = 0$). While this could be accepted as the analogue of the singular point of the Newtonian potential for a point mass, still it was not very clear what meaning could be attached to a singularity *of* space–time (rather than *in* space–time).

These perplexities could both be simply dissolved, at that time. The metric (34) does not describe the whole universe: it can be considered to be valid only for an isolated system; and Mach's principle is a topic for cosmology, rather than for the astronomy of the solar system (more will be said about it in Chap. 6). As for the second objection, Schwarzschild did not have to worry about the $R = \alpha$ singularity, because (34) was only valid for the *exterior* of a spherical mass, and he had proved that such a mass had *always* a radius bigger than α (inside the mass it is (35) which must be applied). In 1923 French physicist Marcel Brillouin gave an explicit formulation of this argument, in very clear terms [10].

Only a few decades later, after discussions in which very different opinions emerged [46,48], the exterior Schwarzschild solution (34) was taken to represent the field of an utterly collapsed star; and $R = 0$ (which, obviously, simply does not make sense in Schwarzschild's original derivation) was interpreted as a true singularity (in opposition to $R = \alpha$, which was re-interpreted as a coordinate singularity).[32] Neither Einstein nor Eddington (arguably "the most distinguished astrophysicist of his time" – [13, p. 93]) ever accepted that such stars could exist, or that general relativity implies that they should exist. Indeed Einstein suggested in 1935 how to extend (34) to obtain a solution where the region $R < \alpha$ is deleted,[33] and in 1939 he advanced a (suggestive, rather than demonstrative) argument implying that "the 'Schwarzschild singularity' does not appear for the reason that matter cannot be concentrated arbitrarily. And this is due to the fact that otherwise the constituting particles would reach the velocity of light" [39].

An important step in the achievement of a new consensus was an article published in 1960 and describing a simple construction of a maximal extension for Schwarzschild's original space–time [67]. Constructing an extension of a singular space–time is useful to understand the nature of its singularities. However, it is sobering to consider that "at least *eleven* completely inequivalent extensions" of Schwarzschild's solution have been found.[34]

The name 'black holes' was introduced in 1969. Notice that the idea that there could be bodies from which no radiation could be emitted (because their escape velocity is bigger than c) had been advanced much earlier by Newtonian physicists, such as John Michell (1783) and Laplace, who called them "dark bodies" (cf. Appendix A in [60]); the first paper to discuss this possibility in connection with Schwarzschild's solution appeared in 1929 [80]. However, the relativistic black hole is a much weirder entity than the classical one. For instance, after crossing the hypersurface $R = \alpha$ what was previously a radial coordinate becomes a *time* coordinate! This is not strange in itself as far as mathematics goes, but one may legitimately question whether it is physically reasonable, and the same may be asked for other curious features of the modern interpretation of

[32] For more on singularities, see Chap. 6, §18.

[33] It is the idea of the "Einstein–Rosen bridge" ([45]; see also the 1936 essay "Physics and Reality", in [41], particularly pp. 353–4).

[34] [96, p. 839]; in this essay one can find a description of these extensions.

the Schwarzschild solution: for example whether they are consistent with taking it as an answer to the physical problem of determining the field of a *static* spherical mass.

In the last thirty years black holes have also been studied from a quantum-mechanical point of view, a new field of inquiry in which British theoretical physicist Stephen Hawking's contributions have been seminal. However, the whole field is still highly controversial, including its basic prediction: "black hole radiation" [62].

By now black holes (described by several other metrics, too) have become almost a byword for Einstein and relativity, and a huge literature on them is today available. However, it is not clear that, should the not overwhelming observational evidence for their existence be explained away, one should have less respect for general relativity, let alone reject it. Surely general relativity does not, in itself, dictate the customary physical interpretation of the Schwarzschild exterior solution.[35]

Many other exact solutions of the Einstein's equation have been found in the last ninety years. Their study has grown into a small academic industry, whose contributions, it must be admitted, are more often of a mathematical than of a physical nature.[36] This is one of the ways general relativity has enriched the field of differential geometry. However, a crucial difference between general relativity and Lorentzian geometry is that a Lorentzian metric, to make sense physically, must be accompanied by a realistic matter model described by a stress-energy tensor such that *together* they satisfy the field equations. Just finding a Lorentzian metric and then formally *defining* a stress energy tensor as suitably proportional to the Einstein tensor of that metric does not normally work.

16. Einstein's Dissatisfaction with His Theory

One point that deserves to be emphasized is that Einstein never considered general relativity as the last word on the subject of gravitation. In fact he started proposing changes in his field equations very soon. In 1917, as we shall see in the next chapter, he introduced a new term in them in order to allow for what then seemed to him a physically reasonable cosmological solution:

$$R_{\mu\nu} - \lambda g_{\mu\nu} = -\kappa\left(T_{\mu\nu} - \frac{1}{2}T g_{\mu\nu}\right).$$

Just two years later he suggested another, 'trace-less' version:

$$R_{\mu\nu} - \frac{1}{4}S g_{\mu\nu} = -\kappa T_{\mu\nu} \tag{36}$$

in order to deal with the possibility that the gravitational fields play "an essential part in the structure of the particles of nature" [35]. In his popular book on relativity he wrote in a footnote (never deleted in subsequent editions):

> The general theory of relativity renders it likely that the electrical masses of an electron are held together by gravitational forces [40, p. 52].

[35] For defenses of Schwarzschild's original interpretation of his exterior solution and criticism of black hole theory see [2,3]; cf. [96, p. 757]; for a participant history of black hole research, interspersed with portraits of the main characters see [103].

[36] For a sample, analysed in depth, see Chapter 5 of [60] (pp. 117–79), and for a virtually complete survey see the encyclopaedic work [98].

His reasons for trying to find a unified theory of gravitation and electricity, following the first proposals of Weyl and Eddington, were rooted in his perception of the field equations as fundamentally defective insofar as they relegated to the stress-energy tensor all forces that could not be, at the moment, analysed in geometrical terms.

In 1936 he compared his field equations to "a building, one wing of which is made of fine marble (left part of the equation), but the other wing of which is built of low-grade wood (right side of equation)" [41, p. 342]. He elaborated this point in his *Autobiographical Notes*:

> The second member on the left side [i.e. $-Sg_{\mu\nu}/2$] is added because of formal reasons; for the left side is written in such a way that its divergence disappears identically in the sense of the absolute differential calculus. *The right side is a formal condensation of all things whose comprehension in the sense of field-theory is still problematic*. Not for a moment, of course, did I doubt that *this formulation was merely a makeshift* in order to give the general principle of relativity a preliminary expression. For it was essentially not anything *more* than a theory of the gravitational field, which was artificially isolated from a total field of as yet unknown structure.[37]

Einstein's search for a unified theory was pursued during the last thirty years of his life. The prevailing, though not universal, opinion today on these efforts is that they were not only unsuccessful, but also that they were basically flawed because of his unwillingness to use quantum mechanics as an essential ingredient of the unification.

17. Conclusive Remarks

Up to the early 1950s, general relativity was a little-frequented subject, amongst physicists – a theory that had to be praised, but that could be safely ignored. As one of Einstein collaborators, Peter Bergmann, told Abraham Pais: "You only had to know what your six best friends were doing and you would know what was happening in general relativity" [81, p. 268]. Its supporting evidence was sparse, questionable, and unstable: essentially it reduced to the changing experimental verdicts on the three notorious tests. (In the next chapter we shall deal with a field of physics on which the impact of general relativity has been very big; of course a full appraisal of general relativity also requires considering its cosmological applications.)

The three empirical predictions were about the only ones that general relativity could refer to for about forty years and, on the whole, the confirming evidence gathered in those decades was not of the highest quality [47]. Though the deflection of light, after Eddington's and Crommelin's celebrated eclipse expeditions in 1919, instantly made Einstein a world-famous star, it was a controversial piece of evidence from the beginning, and was not even consistently confirmed by the next few eclipse observations. True, the gravitational redshift earned Einstein the Royal Society's medal in 1926, because of the authoritative data on Sirius B presented by Walter Adams of the Mount Wilson observatory; it is just a pity that in the meantime these data have dissolved into thin air, and that there is more than a suspicion that the eminent astronomer just saw what he "expected to find, even if it didn't exist" [63, pp. 65–72]. And, as to Mercury perihelion, it had never been a very sound piece of evidence in the first place.

[37] [92, p. 75]; all italics, except for the last, are mine.

Gravitational waves might have been, and in a sense are, another great subject for empirical research.

Unfortunately, on the theoretical side, notwithstanding Einstein's two early papers on them [32,34], the very question of whether or not general relativity really predicts that such waves exist is a tricky one (cf. [69]). Eddington was against them. Einstein himself for a time, in 1937, was convinced that he had a proof that gravitational waves do not exist. Pais in his 1982 biography of Einstein abstains from taking sides on how valid the fundamental quadrupole formula is [81, p. 281].

Incidentally, those unfamiliar with the way physics develops may think that, in contrast to whether or not a given prediction has been confirmed, the issue of whether or not a given theory does make that prediction is a logical question on which there can be no serious disagreement. The rarely acknowledged truth is that the links between an ordinary scientific theory and its predictions have varying degrees of strength, and it is completely normal for different experts to differ in their estimates of how strong or loose a certain link is.

On the empirical side, the detection of gravitational waves has been announced several times since the 1960s, but it is fair to sum up the situation by saying that all positive results so far have been contradicted. Today the field is thriving, in the sense that big international projects centred on gigantic instruments like LIGO (Laser Interferometer Gravitational-Wave Observatory) and Virgo are running, but still no hard evidence of gravitational waves has been found.[38] It has been written [108]:

> The discovery and study of the formation of a black hole by means of gravitational waves would provide a stunning test of relativistic gravity.

This test, however, has still to come to fruition.

More recent tests of general relativity have been performed, with outcomes generally considered satisfactory (cf. Chap. 7). But empirical confirmations should be neither exaggerated nor over-rated; in some fields of science (like astrophysics or cosmology, as opposed to, say, medicine) they are not in themselves more sound or important than theoretical achievement, as evaluated in terms of consistency, mathematical rigour and conceptual transparency. In his later years Einstein insisted on this point on several occasions. Of course even these qualities may be difficult to evaluate and controversial. However, it is safe to say that, concerning the aesthetic appeal of general relativity, there has been a wider consensus on its beauty than on that of any other physical theory [13].

References

[1] Abiko S. 2000: "Einstein's Kyoto address: 'How I created the theory of relativity' ", *Historical Studies in the Physical Sciences*, **31**, Part 1, pp. 1–35.

[2] Abrams L. S. 1979: "Alternative space–time for the point mass", *Phys. Rev. D*, **20**, pp. 2474–2479; "Erratum", *Phys. Rev. D*, **21**, p. 2438.

[3] Abrams L. S. 1989: "Black holes: the legacy of Hilbert's error", *Can. J. Phys.*, **67**, pp. 919–26.

[4] Abraham M. 1914: "Die neue Mechanik", *Scientia*, **15**, pp. 8–27 (French translation: pp. 10–29).

[38] The fact that a sector of the scientific community can prosper for decades with no real success on record – a phenomenon which, needless to say, is far from being restricted to physics! – is in itself an interesting historical and sociological topic. An instructive, recent account of the sociology of the gravitational waves community is contained in [16].

[5] Abraham M. 1914: "Sur le problème de la relativité", *Scientia*, **15**, pp. 101–3.
[6] Assis A. K. 1999: *Relational Mechanics*, Montreal, Apeiron.
[7] Barbour J. B., Pfister H. 1995: *Mach's Principle: From Newton's Bucket to Quantum Gravity*, Boston – Basel – Berlin, Birkhäuser.
[8] Born M. 1962: *Einstein's Theory of Relativity*, New York, Dover.
[9] Bozhkov Y., Rodrigues Jr. W. A. 1995: "Mass and Energy in General Relativity", *General Relativity and Gravitation*, **27**, pp. 813–9.
[10] Brillouin M. 1923: "Les points singuliers de l'univers d'Einstein", *J. de Physique*, 6th ser., **23**, pp. 43–6 (an English translation by S. Antoci is available on "arXiv:physics/0002009").
[11] Cartan E. 1923–1924: "Sur les variétés à Connexion Affine et la Théorie de la Relativité Généralisée", *Annales Scientifiques de l'Ecole Normale Supérieure*, **40** (1923), pp. 325–412; **41** (1924), pp. 1–25.
[12] Cattani C., De Maria M. 1989: "The 1915 Epistolary Controversy between Einstein and Tullio Levi-Civita", pp. 175–200 of *Einstein and the History of general relativity* (Howard D., Stachel J., eds.), Boston, Birkhäuser.
[13] Chandrasekhar S. 1987: *Truth and Beauty. Aesthetics and Motivations in Science*, Chicago – London, University of Chicago Press.
[14] Clarke C. 1979: *Elementary General Relativity*, London, Edward Arnold.
[15] Clifford W. K. 1955: *The Common Sense of the Exact Sciences*, New York, Dover.
[16] Collins H. M. 2003: "Lead into gold: the science of finding nothing", *Stud. Hist. Phil. Sci.*, **34**, pp. 661–91.
[17] Corry L., Renn J., Stachel J. 1997: "Belated Decision in the Hilbert–Einstein Priority Dispute", *Science*, **278**, pp. 1270–3.
[18] D'Inverno R.: *Introducing Einstein's Relativity*, Cambridge University Press.
[19] Eddington A. S. 1918: *Report on the Relativity Theory of Gravitation*, London, Fleetway Press.
[20] Eddington A. S. 1920: *Space, Time and Gravitation*, Cambridge University Press.
[21] Ehlers J. 1991: "The Newtonian Limit of General Relativity", pp. 95–106 of *Classical Mechanics and Relativity: Relationship and Consistency* (ed. by Ferrarese G.), Napoli, Bibliopolis.
[22] Einstein A. 1907: "Relativitätsprinzip und die aus demselben gezogenen Folgerungen", *Jahrbuch der Radioaktivität*, **4**, pp. 411–62; **5**, pp. 98–9.
[23] Einstein A. 1911: "Über den Einfluss der Schwerkraft auf die Ausbreitung des Lichtes, *Ann. Phys.*, **35**, pp. 898–908.
[24] Einstein A. 1913: "Zum gegenwärtigen Stande des Gravitationsproblems", *Physikalische Zeitschrift*, **14**, pp. 1249–62.
[25] Einstein A. 1914: "Zum Relativitäts-problem", *Scientia*, **15**, pp. 337–48 (French translation: 139–50).
[26] Einstein A. 1914: "Die formale Grundlage der allgemeinen Relativitätstheorie", *Königlich Preussische Akademie der Wissenschaften* (Berlin), *Sitzungsberichte*, pp. 1030–85.
[27] Einstein A. 1915: "Zur allgemeinen Relativitätstheorie", *Königlich Preussische Akademie der Wissenschaften* (Berlin), *Sitzungsberichte*, pt 2, pp. 778–86.
[28] Einstein A. 1915: "Zur allgemeinen Relativitätstheorie (Nachtrag)", *Königlich Preussische Akademie der Wissenschaften* (Berlin), *Sitzungsberichte*, pt 2, pp. 799–801.
[29] Einstein A. 1915: "Erklärung der Perihelbewegung des Merkur aus der allgemeinen Relativitä tstheorie", *Königlich Preussische Akademie der Wissenschaften* (Berlin), *Sitzungsberichte*, pt 2, pp. 831–9.
[30] Einstein A. 1915: "Die Feldgleichungen der Gravitation", *Königlich Preussische Akademie der Wissenschaften* (Berlin), *Sitzungsberichte*, pt 2, pp. 844–7.
[31] Einstein A. 1916: "Grundlage der allgemeine Relativitätstheorie", *Ann. Phys.*, **49**, pp. 769–822.
[32] Einstein A. 1916: "Näeherungsweise Integration der Feldgleichungen der Gravitation", *Königlich Preussische Akademie der Wissenschaften* (Berlin), *Sitzungsberichte*, pt 1, pp. 688–96
[33] Einstein A. 1918: "Prinzipielles zur allgemeinen Relativitätstheorie", *Annalen der Physik*, **55**, pp. 241–4.
[34] Einstein A. 1918: "Über Gravitationswellen", *Königlich Preussische Akademie der Wissenschaften* (Berlin), *Sitzungsberichte*, pp. 154–67.
[35] Einstein A. 1919: "Spielen Gravitationsfelder im Aufber der materiellen Elementarteilchen eine wesentliche Rolle?", *Königlich Preussische Akademie der Wissenschaften* (Berlin), *Sitzungsberichte*, pt 1, pp. 349–56 (transl. in [44], pp. 189–98).
[36] Einstein A. 1920: "Meine Anwort. Ueber die anti-relativitätstheorie G.m.b.H.", *Berliner Tageblatt*, August 27.

[37] Einstein A. 1921: “Geometrie und Erfahrung”, trans. in Einstein 1994, pp. 254–68.
[38] Einstein A. 1922: *The Meaning of Relativity*, Princeton, Princeton University Press.
[39] Einstein A. 1939: “Stationary system with spherical symmetry consisting of many gravitating masses”, *Ann. Math.*, **40**, pp. 922–36.
[40] Einstein A. 1954: *Relativity. The Special and the General Theory*, 15th ed., Routledge, London and New York.
[41] Einstein A. 1994: *Ideas and Opinions* [1954], New York, The Modern Library.
[42] Einstein A., Grossmann M. 1913: *Entwurf einer verallgemeinerten Relativitätstheorie und einer Theorie der Gravitation*, Leipzig, Teubner.
[43] Einstein A., Infeld L. 1938: *The Evolution of Physics*, New York, Simon and Schuster.
[44] Einstein A., Lorentz A. A., Weyl H., Minkowski H. 1932: *The Principle of Relativity*, Dover.
[45] Einstein A., Rosen N. 1935: “The Particle Problem in the General Theory of Relativity”, *Phys. Rev.*, **48**, pp. 73–7.
[46] Eisenstaedt J. 1982: “Histoire et Singularités de la Solution de Schwarzschild”, *Archive for History of Exact Sciences*, **27**, pp. 157–98.
[47] Eisenstaedt J. 1986: “La relativité générale à l’étiage: 1922–1955”, *Archive for History of Exact Sciences*, **32**, pp. 115–85.
[48] Eisenstaedt J. 1987: “Trajectoires et impasses de la solution de Schwartzschild, 1925–1955”, *Archive for History of Exact Sciences*, **37**, pp. 275–357.
[49] Engel A., Shucking E. (eds.) 1997: *The Collected Papers of Albert Einstein, vol. 7, The Berlin Years: Writings, 1914–1917, English Translation*, Princeton–Oxford, Princeton University Press.
[50] Engel A., Shucking E. (eds.) 2002: *The Collected Papers of Albert Einstein, vol. 7, The Berlin Years: Writings, 1918–1921, English Translation*, Princeton–Oxford, Princeton University Press.
[51] Euler L. 1732: “De linea brevissima in superficie quacumque duo quaelibet puncta jungente”, *Comment. Acad. Petropol.*, **3** (ad annum 1728), pp. 110–24.
[52] Fermi E. 1922: “Sopra i fenomeni che avvengono in vicinanza di una linea oraria”, *Rend. Lincei*, **31**, pp. 21–3, 51–2, 101–3.
[53] Friedrichs K. 1927: “Eine invariante Formulierung des Newtonschen Gravitationsgesetzes und des Grenzübergunganges vom Einsteinschen zum Newtonschen Gesetz”, *Mathematische Annalen*, **98**, pp. 566–75.
[54] Gauss C. F. 1828: *Disquisitiones generales circa superficies curvas* (in *Werke 4*, pp. 217–58).
[55] Gerber P. 1898: “Über die räumliche und zeitliche Ausbreitung der Gravitation”, *Zeitschrift für Mathematik und Physik*, **43**, pp. 93–104.
[56] Graneau P., Graneau N. 2003: “Machian Inertia and the Isotropic Universe”, *General Relativity and Gravitation*, **35**, pp. 751–70.
[57] Ginzburg V. L. 1983: “The Heliocentric System and the General Theory of Relativity (from Copernicus to Einstein)”, pp. 254–308 in Gribanov *et al.* 1983.
[58] Giorgi G., Cabras A. 1929: “Questioni relativistiche sulle prove della rotazione terrestre”, *Rend. Lincei*, **9**, pp. 513–7.
[59] Gribanov *et al.* 1983: *Einstein and the Philosophical Problems of 20th-Century Physics* [1979], Moscow, Progress Publishers.
[60] Hawking S. W., Ellis G. F. R. 1973: *The Large-Scale Structure of Space–Time*, Cambridge University Press.
[61] Held A. (ed.) 1980: *General Relativity and Gravitation. One Hundred Years After the Birth of Albert Einstein*, New York and London, Plenum Press.
[62] Helfer A. D. 2003: “Do black holes radiate?”, *Rep. Prog. Phys.*, **66**, pp. 943–1008.
[63] Hetherington N. S. 1988: *Science and Objectivity*, Ames, Iowa State University Press.
[64] Howard D., Norton J. D. 1993: “Out of the Labyrinth? Einstein, Hertz, and the Göttingen Answer to the Hole Argument”, pp. 30–62 of *The Attraction of Gravitation: New Studies in the History of General Relativity* (Earman J., Janssen M., Norton J. D., eds.), Boston–Basel–Berlin, Birkhäuser.
[65] Kox A. J. 1988: “Hendrik Antoon Lorentz, the Ether, and the General Theory of Relativity”, *Arch. Hist. Exact Sci.*, **38**, pp. 67–78.
[66] Kretschmann E. 1914: “Über den physikalischen Sinn der Relativitätspostulate, A. Einsteins neue und seine ursprünglicher Relativitätstheorie”, *Annalen der Physik*, **53**, pp. 575–614.
[67] Kruskal M. D. 1960: “Maximal Extension of Schwarzschild Metric”, *Phys. Rev.*, **119**, pp. 1753–5.
[68] Lang S. 2004: *Linear Algebra*, New York – Heidelberg – Berlin, Springer-Verlag.

[69] Logunov A., Mestvirishvili M. 1989: *The Relativistic Theory of Gravitation*, Moskow, Mir Publishers.
[70] Lusanna L. 2004: "Book Review", *Found. Phys.*, **34**, pp. 1281–2.
[71] Macdonald A. 2001: "Einstein's hole argument", *Am. J. Phys.*, **69** (2), pp. 223–5.
[72] Mach E. 1960: *The Science of Mechanics. A Critical and Historical Account of its Development*, La Salle, Open Court.
[73] Mannheim P. D. 2000: "Attractive and Repulsive Gravity", *Found. Phys.*, **30**, pp. 709–46.
[74] Maxwell J. C. 1920: *Matter and Motion* [1976], ed. with notes and appendices by J. Larmor, New York, Dover 1991.
[75] Minkowski H. 1909: "Raum und Zeit", *Jahresbericht der deutschen Mathematiker-Vereinigung*, **18**, pp. 75–88, *Physikalische Zeitschrift*, **10**, pp. 104–11.
[76] Misner C. W., Thorne K. S., Wheeler J. A. 1973: *Gravitation*, New York, Freeman.
[77] North J. D. 1990: *The Measure of the Universe* [1965], New York, Dover.
[78] Norton J. 2004: "The Hole Argument", *Stanford Encyclopedia of Philosophy*, http://plato.stanford.edu/entries/space-time-holearg/.
[79] O'Neill B. 1983: *Semi-Riemannian Geometry With Applications to Relativity*, New York, Academic Press.
[80] Oppenheimer J. R., Snyder H. 1939: "On Continued Gravitational Contraction", *Phys. Rev.*, **56**, pp. 455–459.
[81] Pais A. 1982: *"Subtle Is the Lord...". The Science and Life of Albert Einstein*, New York, Oxford University Press.
[82] Pauli W. 1958: *Theory of Relativity*, New York, Dover.
[83] Poincaré H. 1902: *La Science et l'Hypothèse*, Paris, Flammarion.
[84] Poincaré H. 1906: "Sur la dynamique de lélectron, *Rendiconti del Circolo Matematico di Palermo*, **21**, pp. 129–75.
[85] Ricci G., Levi-Civita T. 1901: "Méthodes de calcul différentiel absolu et leurs applications, *Mathematische Annalen*, **54**, pp. 125–201.
[86] Reichenbach H. 1942: *From Copernicus to Einstein* [1928], trans. from the German, New York, Dover 1980.
[87] Robertson H. P. 1949: "Geometry as a branch of physics", in [92], pp. 315–32.
[88] Roseveare N. T. 1982: *Mercury's Perihelion from Le Verrier to Einstein*, Oxford, Clarendon Press.
[89] Rowe D. E. 2002: "Einstein's Gravitational Field Equations and the Bianchi Identities", *The Mathematical Intelligencer*, **24** (4), pp. 57–66.
[90] Sachs R. K., Wu H. 1977: *General Relativity for Mathematicians*, New York – Heidelberg – Berlin, Springer-Verlag.
[91] Salvioli S. E. 1972: "On the Theory of Geometric Objects", *J. Differential Geometry*, **7**, pp. 257–78.
[92] Schilpp P. A. (ed.) 1949: *Albert Einstein: Philosopher–Scientist*, Evanston (Ill.), Library of Living Philosophers.
[93] Schwarzschild K. 1900: "Über das zulässige Krümmungsmaass des Raumes", *Vierteljahrsschrift der astronomischen Gesellschaft*, **35**, pp. 337–47 (English transl. in *Class. Quant. Grav.*, **15** (1998), pp. 2539–44).
[94] Schwarzschild K. 1916: "Über das Gravitationsfeld einer Massenpunktes nach der Einsteinschen Theorie", *Königlich Preussische Akademie der Wissenschaften* (Berlin), *Sitzungsberichte*, pp. 189–96 (an English translation by S. Antoci and A. Loinger is available on "arXiv:physics/9905030").
[95] Schwarzschild K. 1916: "Über das Gravitationsfeld einer Kugel aus inkompressibler Flüssigkeit nach der Einsteinschen Theorie", *Königlich Preussische Akademie der Wissenschaften* (Berlin), *Sitzungsberichte*, pp. 424–34 (an English translation by S. Antoci and A. Loinger is available on "arXiv:physics/9912033").
[96] Senovilla J. M. M. 1997: "Singularity Theorems and Their Consequences", *General Relativity and Gravitation*, **29**, pp. 701–848.
[97] Stachel J. 1980: "Einstein and the Rigidly Rotating Disk", pp. 1–15 of Held 1980.
[98] Stephani H., Kramer K., MacCallum M., Hoenselaers C., Herlt E. 2003: *Exact Solutions of Einstein's Field Equations*, 2nd ed., Cambridge University Press.
[99] Struik D. 1987: *A Concise History of Mathematics*, 4th ed., New York, Dover.
[100] Synge J. L. 1966: "What is Einstein's Theory of Gravitation?", pp. 7–15 of B. Hoffmann (ed.), *Perspectives in Geometry and Relativity*, Bloomington – London, Indiana University Press.

[101] Synge J. L. 1971: *Relativity: The General Theory*, Fourth Printing, Amsterdam, North-Holland (First Edition: 1960).

[102] Tavakol R., Zalaletdinov 1998: "On the Domain of Applicability of General Relativity", *Found. Phys.*, **28**, pp. 307–31.

[103] Thorne K. S. 1994: *Black Holes & Time Warps*, New York – London, Norton.

[104] Weyl H. 1922: *Space–Time–Matter*, transl. of the 4th (1921) ed., with a new preface (1950), New York, Dover.

[105] Weyl H. 1949: "Relativity theory as a stimulus in mathematical research", *Proc. Am. Math. Soc.*, **93**, pp. 535–41.

[106] Whitrow G. J., Morduch G. E. 1965: "Relativistic theories of gravitation. A comparative analysis with particular refernce to astronomical tests", *Vistas in Astronomy*, **6**, pp. 1–67.

[107] Whittaker E. 1951–53: *A History of the Theories of Aether and Electricity*, New York, Dover 1989.

[108] Will C. M. 1999: "Gravitational Radiation and the Validity of General Relativity", *Physics Today*, Oct., pp. 38–43.